

Emittance Preservation in a Plasma Wakefield Accelerator

by

Robert Ariniello

B.S., Colorado School of Mines, 2015

M.S., University of Colorado, 2018

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Physics
2022

Committee Members:

Michael Litos, Chair

John Cary

Tobin Munsat

Henry Kapteyn

Robert Ergun

Ariniello, Robert (Ph.D., Physics)

Emittance Preservation in a Plasma Wakefield Accelerator

Thesis directed by Prof. Michael Litos

Plasma wakefield accelerators have the potential to drastically reduce the size and cost of conventional particle accelerators. In these devices, a laser or charged particle beam excites a density wave in a plasma. The strong fields created by the density wave can then be used to accelerate a trailing beam to high energy. The accelerating electric field can be three to four orders of magnitude larger than that found in a typical radio frequency accelerator. Particle accelerators, however, are tools for producing the beams required by particle colliders and light sources, and these applications place strict requirements on the quality of the beam. Here, beam quality refers to the beam emittance, a measure of the transverse momentum spread of the particles in the beam. Unfortunately, the beam emittance tends to grow considerably in plasma wakefield accelerator experiments.

The primary mechanism behind emittance growth, chromatic phase spreading, increases the emittance when the divergence of the beam is not matched to the focusing force of the plasma. This process can be mitigated by using a plasma source with density ramps. The focusing force of the plasma gradually increases in the ramp, reducing the beam size adiabatically. This dissertation covers the development of a laser ionized plasma source intended to demonstrate emittance preservation. We start by developing a general, analytic theory of beam evolution in an adiabatic plasma ramp. Using this theory, in combination with particle in cell and particle tracking simulations, we enumerate the requirements a plasma source must meet in order to preserve the beam's emittance. Based on these requirements, we propose an optical technique capable of producing a suitable laser ionized plasma source. Finally, we demonstrate the optical technique experimentally.

Acknowledgements

I've had the opportunity to build experimental setups at both the University of Colorado and SLAC. Getting to work at both institutions introduced me to a lot of people, from whom I have learned tremendously. I'm grateful to my advisor Mike Litos who has supervised me along the way and, because he knows everyone, involved me in many collaborations. He showed that with endless enthusiasm and optimism, you can turn an empty room into a, sometimes, fully functioning laser lab. John Cary has provided insight on how to write a good paper and get to the root of a physics problem, he also provided our PIC software. Chris Doss has been with me in lab from the beginning. Keenan Hunt-Stone provided many physics discussions and survived the cliffs. Valentina Lee has kept the laser running and the questions too. Many people helped build the lab and keep the laser operating: Kaikai Zhang, Tim Lim, and Markus Raschke. Andreas Becker lent us his grad student Joel Venzke to do TDSE calculations.

At SLAC, Mark Hogan introduced me to FACET by turning me loose on a massive project with minimal guidance and having confidence that I would figure it out. Luckily, I had Henrik Ekerfelt, Doug Storey, Spencer Gessner, Christine Clarke, Brendan O'Shea, Claudio Emma, James Allen, Elias Gerstmayr, and Zhijiang Chen to keep my company in the tunnel. Then there are the many collaborators who have trusted me to design equipment for them: Sébastien Corde, Max Gilljohann, Ken Marsh, Chaojie Zhang, Sebastian Meuren, Andrew Sutherland, and Bernhard Hidding.

Finally, my family has supported me through the entire process and my girlfriend Audrey has patiently listened and provided advice through this long journey.

Contents

Chapter

1	Introduction	1
1.1	Outline	5
2	Basics of Plasma Wakefield Acceleration	7
2.1	Plasma	8
2.1.1	What is a Plasma?	8
2.1.2	Plasma Parameters	9
2.1.3	Linear Plasma Waves	11
2.2	Plasma Wakefield Acceleration	19
2.2.1	Plasma Wakefield Acceleration in the Linear Regime	22
2.2.2	The Nonlinear Blowout Regime	29
2.2.3	Fields in the Nonlinear Regime	36
2.3	Accelerator Physics	42
2.3.1	Transverse Motion	43
2.3.2	Courant-Snyder Parameters	47
2.3.3	The Beam Distribution and Emittance	51
2.3.4	Gaussian Beam Distribution	58
2.3.5	Beam Self Fields	59
2.3.6	Longitudinal Effects	61
2.4	Particle in Cell Simulation	64

3	Evolution of the Witness Beam	66
3.1	Transverse Dynamics in a Uniform Plasma	68
3.1.1	Evolution of the Beta Function	68
3.1.2	Emittance Growth	69
3.2	Transverse Dynamics in a Nonuniform Plasma	70
3.2.1	Adiabatic Evolution	71
3.2.2	Emittance Growth of the Matched and Unmatched Beam	75
3.2.3	Non-Adiabatic Evolution	78
3.2.4	Perturbative Plasma Ramp	80
3.2.5	Summary of the Effects of the Plasma Ramps	82
3.3	Acceleration and Wake Loading	83
3.3.1	Acceleration and Loading for the FACET-II Beam	84
3.3.2	Plasma Width	87
3.4	A General Approach Including Acceleration	89
3.4.1	Single Particle Motion	89
3.4.2	Evolution of the Beam Moments and Emittance	92
3.4.3	Solution in a Uniform Plasma	97
3.4.4	Evolution of the Slice Moments and Slice Emittance	99
3.4.5	Emittance Growth of a Beam with Imperfect Wake Loading	99
3.4.6	Emittance Evolution in a Plasma Source with Density Ramps	109
3.5	Additional Effects	111
3.5.1	Feedback on the Wake	112
3.5.2	Ion Motion	115
3.6	Plasma Source Requirements	117
3.6.1	Example Ramp Design	118
3.6.2	Uniform Plasma Ramp	122

4	Laser Ionized Plasma Sources	124
4.1	Ionization Models	125
4.1.1	ADK Tunneling Ionization	127
4.1.2	PPT Ionization Theory	129
4.1.3	Ionization Formulas in Practical Units	133
4.1.4	SAE-TDSE Ionization Calculations	135
4.2	Lasers and Optics	136
4.2.1	Basics of Light	136
4.2.2	The Laser	139
4.2.3	Fourier Optics	140
4.2.4	Nonlinear Optics	144
4.3	Bessel Beams	147
4.3.1	Wave Equation in Cylindrical Coordinates	148
4.3.2	The Axicon Lens	151
4.3.3	Shaping the On-axis Intensity	153
4.4	Laser Propagation Through a Plasma	155
4.4.1	Index of Refraction in a Plasma	156
4.4.2	Split Step Fourier Simulation	157
4.4.3	Optimizing the Plasma Width	158
4.5	Tandem Lens Optics	160
4.5.1	Phase of Each Lens	161
4.5.2	Fabrication	163
4.5.3	E301 Plasma and Optics	168
4.6	The Problem of the Electron Beam	170
4.7	Shaping the Gas Density	171
4.7.1	Elongated Gas Jets	173
4.7.2	Laser Ionization of a Lithium Vapor Oven	177

5	Experimental Setup	180
5.1	University of Colorado Laser Lab	181
5.1.1	The Laser System	181
5.1.2	The Vacuum System	183
5.1.3	DAQ	185
5.1.4	Setup for Testing Lenses	186
5.2	FACET-II Accelerator	187
5.2.1	Brief Linac Overview	189
5.2.2	Beam Parameters	193
5.2.3	IP Area	193
5.2.4	Diagnostics	198
5.3	FACET-II Laser System	202
5.3.1	CPA System	203
5.3.2	Deformable Mirror and Transport	204
5.3.3	Laser System in the Tunnel	207
6	Experimental Results	209
6.1	Intensity Profile Extraction	209
6.2	Characterization of Lens A	210
6.3	Characterization of Lens B	212
6.3.1	Alignment Sensitivity of the Tandem Lens	215
6.3.2	Phase Sensitivity of the Tandem Lens	217
6.3.3	Laser Collimation	219
6.3.4	Frequency Effects	220
7	Conclusion	221
7.1	Future Work	223

Bibliography	225
---------------------	------------

Appendix

A Green's Function for a Linear Plasma Wave	235
B Adiabatic Solution for the Beta Function	236
C Perturbative Solution for the Beta Function	239
D Particle Tracking Algorithm	242
E Beam Evolution in a Plasma Source with Adiabatic Ramps	244
F Versions of PPT	246

Tables

Table

3.1	FACET-II beam parameters	84
3.2	Emittance saturation lengths	105

Figures

Figure

2.1	Linear wake density	15
2.2	Three dimensional linear wakefields	20
2.3	Influence of the drive beam on the amplitude of the linear wakefield	24
2.4	Useful phase of the linear wakefields	27
2.5	Beam loading in the linear regime	28
2.6	Trade-off between transformer ratio and efficiency	29
2.7	The trajectory crossing condition for blowout	32
2.8	Nonlinear wake density	33
2.9	Nonlinear wake model	36
2.10	Nonlinear wakefields	41
2.11	Phase space evolution through a quadrupole and a drift	47
2.12	Courant-Snyder invariant ellipse	50
2.13	4D phase space distribution of an electron beam	52
2.14	Statistical phase space ellipse	55
2.15	Growth of the normalized emittance	58
2.16	Beam self fields	61
2.17	Adiabatic dampening	63
3.1	Single particle motion in an ion column	69

3.2	Evolution of the beta function in an ion column	70
3.3	Projected emittance growth due to chromatic phase spreading	70
3.4	Adiabatic plasma ramp	76
3.5	Emittance growth in an adiabatic plasma	78
3.6	Non-adiabatic plasma ramp	79
3.7	Perturbative plasma ramp	81
3.8	Drive-witness separation and bunch length	86
3.9	Transformer ratio and final energy spread	87
3.10	Plasma density for different plasma widths	88
3.11	Convergence with plasma width	89
3.12	Single particle motion	92
3.13	Emittance growth with mismatch and offset	93
3.14	Emittance growth without energy gain	98
3.15	Linear and quadratic beam loading	100
3.16	Longitudinal slice centroid oscillations	103
3.17	Growth of the projected and slice emittances	106
3.18	Energy slice spot sizes	108
3.19	Beam evolution in a ramped plasma source with acceleration	111
3.20	Comparison of plasma ramps	112
3.21	Head erosion	113
3.22	Hosing instability	114
3.23	Ion motion	116
3.24	Decelerating gradient	118
3.25	Gaussian ramp tolerances	119
3.26	Example ramp design	121
3.27	Uniform plasma ramp	122

4.1	Field ionization regimes	126
4.2	Laser ionization fraction of different gases	130
4.3	Experimental comparison of ionization theories	133
4.4	TDSE ionization rate for lithium	136
4.5	Fourier optics	142
4.6	Bessel beams	149
4.7	Axicon lens	151
4.8	Intensity of the axicon lens	152
4.9	Shaping the on-axis intensity	155
4.10	Split-step Fourier algorithm	159
4.11	Optimizing the plasma width	160
4.12	Tandem lens schematic	161
4.13	Phase of each tandem lens	164
4.14	Binary optics	165
4.15	Measured binary optic profile	167
4.16	Etch patterns	168
4.17	Design of the E301 plasma and lens	170
4.18	Beam evolution in the E301 plasma	171
4.19	Beam ionization in E301	172
4.20	Converging-diverging nozzles	175
4.21	Transverse density profile of various 2D conical nozzles	176
4.22	Ramp profile of the elongated gas jet	177
4.23	Design of the E300 plasma and lens	178
5.1	Centaurus 15 TW laser system	182
5.2	Typical laser profile	183
5.3	Vacuum system schematic	184

5.4	Tandem optics test setup	188
5.5	Overview of the FACET-II accelerator	189
5.6	Photocathode and rf gun	190
5.7	Overview of the main FACET-II linac	191
5.8	Final bunch compressor and experimental area	192
5.9	FACET-II beam parameters	194
5.10	Schematic of the experimental area	195
5.11	Schematic of the IP area	196
5.12	Primary vacuum chambers	197
5.13	Imaging spectrometer	199
5.14	Transverse deflecting cavity	200
5.15	Electro-optic sampling beam position monitor	202
5.16	Overview of the FACET-II experimental laser	204
5.17	Schematic of the laser transport	205
6.1	Intensity at lens B	211
6.2	Laser propagation after lens A	212
6.3	Transverse offsets of lens A	213
6.4	Bessel focus of the tandem lens	214
6.5	Bessel beam with wavefront errors	214
6.6	Longitudinal intensity profile of the tandem optics	215
6.7	Offsets of lens B	216
6.8	Impact of focus and coma aberrations on the tandem lens	217
6.9	Impact of astigmatism and trefoil aberrations on the tandem lens	218
6.10	A diverging beam through the tandem optics	219
6.11	Tandem lens focus at different frequencies	220

Chapter 1

Introduction

Particle accelerators are incredibly powerful scientific tools. They have been the primary research tool of high energy particle physics from the discovery of the quark to the measurement of the Higgs boson. As particle physics has advanced, it has demanded higher and higher beam energies from accelerators. At the same time, accelerators have proliferated as bright x-ray sources through synchrotron light sources and, more recently, as x-ray free electron lasers (FEL). The short, bright, coherent pulses produced by these machines can be used to probe the structures and dynamics of condensed matter, novel materials, and biological systems. With such a wide range of applications, x-ray beam time is competitive on these machines; their cost and size limit the number available.

There are two drivers for advancing accelerator technology. The first is to create higher energy colliders. The second is to reduce the cost and size of existing accelerators so that they are more accessible. One of the main limitations for producing higher energy electron accelerators is the size of the machine. With existing radio frequency (RF) technology, the size of a linear accelerator is set by the accelerating gradient, which is the rate of energy gain per unit distance. The accelerating gradient is proportional to the strength of the electric field accelerating the particles. In a RF accelerator, the particles are accelerated by a travelling electromagnetic wave in a RF cavity. The strength of this wave is ultimately limited by the material properties of the cavity. At some critical electric field, the force from the RF wave on the electrons in the walls of the cavity will exceed the binding energy of the electrons. They are then separated from the wall in a process

called dielectric breakdown, damaging the cavity. Gradients of up to 50 MeV/m have been reached in tests [1]. Production accelerators typically operate at lower fields; the copper linac at SLAC National Accelerator Laboratory operates at 22 MV/m, and the super conducting Linac Coherent Light Source II (LCLS-II), under construction at SLAC, is designed to operate at 16 MV/m [2]. In summary, conventional RF accelerators are limited by fundamental physics considerations to accelerating gradients of less than 100 MeV/m, and often fall far shy of that.

Instead of a linear accelerator, circular accelerators can be used. The particles travel around the accelerator many times, gaining energy each time they pass through the accelerating cavities. In these devices, the accelerating gradient is not as much of a limitation. The limiting factor depends on the type of particle being accelerated. Circular electron accelerators are limited by the synchrotron radiation emitted by the electron as it moves around the circle. As the electron energy increases, the energy lost per revolution scales as the relativistic factor to the fourth power γ_b^4 . At some point, the loss equals the energy gain per revolution and the maximum energy is reached. Protons have about 2000 times the mass of electrons so, for a given energy, γ_b is smaller by a factor of 2000 and synchrotron radiation losses are no longer an issue, having been reduced by a factor of 1.6×10^{13} . Instead, it is the strength of the bending magnets, which keep the beam travelling in a circle, that limit the maximum particle energy.

For both types of accelerators, going to higher energy with current technology requires increasing the size of the accelerator. An acceleration device with a much larger gradient could be used to produce higher energy linear accelerators in the same space as existing accelerators or shrink existing accelerators down to a much smaller footprint. The factor that limits the gradient in conventional RF cavities is the breakdown of the cavity material. Using a plasma to contain the accelerating field circumvents this problem because the plasma is already broken down. Plasma wakefield acceleration (PWFA) is a technique for producing very large accelerating gradients using a density wave excited in a plasma [3, 4, 5]. The density wave, or wake, is excited by a particle beam or laser driver traveling close to the speed of light. Energy is transferred from the driving beam into the plasma. A suitably located, charged particle beam can then extract energy from the

plasma and be accelerated to high energy over a very short distance. Research in plasma wakefield experiments has made great progress towards turning this idea into reality.

Experiments at SLAC’s Final Focus Test Beam demonstrated that accelerating fields of 52 GV/m could indeed be achieved in an electron beam driven plasma wake [6]. These experiments used a single beam to both drive the plasma and sample the accelerating field. SLAC’s Facility for Advanced Accelerator Experimental Tests (FACET) had the capability to send a second “witness” beam behind the first beam [7]. Dual bunch experiments demonstrated that a witness beam could be accelerated with high energy transfer efficiency, $> 30\%$, from the drive to the witness beam and a low energy spread of 0.7%, but the total energy gain of the witness beam was limited [8]. Later experiments showed that single stage gain of 9 GeV was possible while still maintaining modest energy spread [9]. The FLASHForward facility at DESY introduced more stable, better characterized drive bunches allowing it to achieve gradients of 1.3 GeV/m, efficiency of 42%, and energy spread of 0.13%, simultaneously.

While most beam driven plasma wakefield accelerator experiments operate by accelerating an external beam, laser driven accelerators typically produce the accelerated beam in the plasma from plasma electrons. Using powerful lasers, accelerating gradients of over 100 GeV/m and electron beams with 4.2 GeV have been created [10, 11]. Typically, the beams are lower charge and have larger energy spread than those accelerated in a beam driven accelerator.

Despite the promising experimental progress, collider and light source applications require more than just high energy and low energy spread. The beams must be of high quality and high brightness. In this case, quality is quantified by the transverse emittance of the beam, a measure of the amount of transverse momentum spread the particles have. Low emittance is essential to achieve the small focal spots required for high brightness in a collider, increasing the number of collision events. In an FEL, the lasing process will not occur if the beam’s emittance is not sufficiently low. In beam driven plasma wakefield experiments, the emittance tends to grow significantly while the beam is accelerated. Chaining multiple plasma stages together to create a high energy accelerator would result in catastrophic emittance growth with current technology. Emittance preservation has

yet to be demonstrated experimentally and is the next major goal for plasma accelerator research.

This dissertation is part of an experimental campaign to demonstrate emittance preservation in a beam driven plasma wakefield accelerator, preferably while maintaining a high accelerating gradient, a high efficiency, low energy spread, and significant energy gain. Emittance growth primarily results when the divergence of the beam is not matched to the strong focusing forces from the plasma. This requires the beam to be focused to a small size, typically smaller than can be achieved by conventional magnetic optics. Instead, a plasma source with a ramped plasma density has been proposed to focus the beam down to the required size [12, 13]. The beam dynamics in ramped plasmas have been previously studied [14, 15, 16, 17, 18, 19, 20]. However, there are still many unknowns about what the ramp should look like and how sensitive the emittance is to variations in the ramp and the beam. Transverse offset between the drive and witness beam can also cause emittance growth even if the beam is sufficiently focused, but existing literature has only briefly considered this mechanism [21, 22, 23, 24].

Emittance preservation is a major goal of upcoming experiments at the FACET-II facility at SLAC. FACET-II is an upgrade to FACET designed to provide the high current, high quality beams necessary for high gradient, large energy gain, and emittance preservation experiments. For these experiments to be successful two things need to be addressed. First, existing gaps in the understanding of beam dynamics in plasma ramps need to be understood in order to inform the design of the plasma source. Second, a technique for creating a plasma source with controlled density ramps needs to be developed. Only with new plasma sources will experiments at FACET-II be able to demonstrate preservation of the emittance.

This dissertation is about the design and creation of a plasma source for emittance preservation experiments at FACET-II. We start by developing a general approach to describe the evolution of the witness beam in a plasma wakefield accelerator with density ramps. Informed by our theoretical model, we develop requirements for the FACET-II plasma sources to preserve the emittance and maximize the performance of the accelerator. We then propose a laser ionization scheme to produce such a plasma and present optical tests that show our optical technique works.

1.1 Outline

The research presented in this dissertation works toward the goal of demonstrating emittance preservation in a plasma wakefield accelerator. The research broadly falls into two categories: theoretical investigations of beam dynamics and experimental plasma source development. The chapters are roughly separated along those lines. Ch. 2 and Ch. 3 are focused on beam dynamics while Ch. 4, Ch. 5, and Ch. 6 focus on the development of plasma sources for emittance preservation experiments. Ch. 4 refers back to the results of Ch. 3 to inform the plasma source design.

Chapter 2 provides the background information on plasma wakefield accelerators and accelerator physics necessary to understand the rest of the dissertation, particularly Ch. 3. Nothing in Ch. 2 is new work; instead, it is focused on explaining the key concepts and terms used within the plasma accelerator field with an emphasis on beam driven accelerators.

Chapter 3 works out the detailed beam dynamics, with a focus on chromatic effects, within a PWFA. The work is motivated by a desire to understand the impact the plasma source has on the quality of the beam. To this end, the chapter also applies the beam dynamics results to the specific case of FACET-II. Concrete requirements for the FACET-II plasma source are presented. Almost all of the content in this chapter represents new, novel work done as part of this dissertation.

Chapter 4 lays out a newly developed optical technique to produce custom tailored plasma sources with high precision. The first half of the chapter presents the background in field ionization and optics necessary to understand the technique. The second half develops the technique by combining several well-known ideas and applying them to the formation of laser ionized plasmas. The chapter ends by showing some detailed plasma source designs that use the optical technique to create plasmas that satisfy the requirements laid out in Ch. 3.

Chapter 5 describes the experimental setups used for plasma and optics testing at the University of Colorado Boulder (CU) and for plasma wakefield accelerator experiments at SLAC. A significant portion of the work done for this dissertation was spent designing, building, and commissioning hardware in both labs.

Chapter 6 describes the results of tests performed to characterize a pair of custom diffractive optics. The optics were designed as a part of this dissertation work to produce a laser ionized plasma source. The experiments were carried out at CU to demonstrate that the optical technique presented in Ch. 4 performs as expected.

Finally, Ch. 7 summarizes the dissertation and discusses the next steps required to experimentally demonstrate emittance preservation in a plasma wakefield accelerator.

Chapter 2

Basics of Plasma Wakefield Acceleration

A plasma wakefield accelerator is created by exciting a density wave in a plasma. The electric and magnetic fields associated with this wave are immense, capable of accelerating particles to high energy over a very short distance. These immense fields are not limited to the longitudinal direction—the direction of acceleration. Transverse fields exert focusing forces on the particles that are several orders of magnitude stronger than those produced by the focusing magnets used in conventional accelerators. These focusing forces cause the particles to oscillate back and forth. When a large number of particles, called a bunch or beam, are loaded into the wave, the dynamics of the system can become complex. Taking advantage of these dynamics to produce high quality beams from a plasma-based accelerator requires an understanding of two things: the structure of the wave and the dynamics of the beam placed within it. The first topic falls under the umbrella of plasma physics. The second is within the purview of accelerator physics.

This chapter contains the background information in plasma physics and accelerator physics necessary to understand Ch. 3. The first half of the chapter is about plasma physics, starting with the basic definition of a plasma and working up to an analytic description of highly nonlinear waves and wakefields. In the process, important concepts in the physics of plasma accelerators are introduced. The second half covers accelerator physics, introducing the concepts used to describe a particle beam and calculate the beam's dynamics under the influence of external fields.

2.1 Plasma

Plasma is often called the fourth state of matter, the first three being solid, liquid, and gas (the condensed matter physicists like to add a couple more). The obvious way to move up through the list is to increase the temperature. Ice melts into water at 0°C , water evaporates into steam at 100°C , and, if the temperature keeps increasing, steam eventually disassociates and ionizes into plasma. In the case of melting and evaporation, it is the bonds between atoms and molecules which are broken as the temperature increases. For a plasma, it is the bonds within the atom that break, stripping one or more electrons away from the nucleus. Both the electrons and ions are charged but unbound; this is at the heart of what makes plasma special. Gases interact through direct collisions with their neighbors. The charged particles in a plasma interact collectively through the much longer range electromagnetic force.

2.1.1 What is a Plasma?

A plasma is a fluid comprised of free positively and negatively charged particles. The fluid can contain a variety of ions of different species or different ionization states, as well as neutral particles and electrons. The particles in the fluid are only free if their thermal energy is greater than the potential energy from their nearest neighbors. This means plasmas are typically hot, with temperatures of at least 1 eV ($11,600^{\circ}\text{K}$) [25, 26].

A plasma is dominated by electromagnetic interactions rather than collisions. Too many collisions between charged particles and neutral particles would cause the fluid to behave like a gas. In addition, plasma is typically assumed to be quasineutral¹ in the stationary state, meaning that it has about the same amount of positive and negative charge in a given volume. In the next section we will introduce the various parameters that describe a plasma. Using these parameters, we will present the three plasma criteria, which formalize the requirements for an ionized gas to behave as a plasma.

¹ Not all plasma are quasineutral; for example, a charged particle beam has only one sign of charge. Many results in plasma physics, however, rely on the quasineutral assumption.

2.1.2 Plasma Parameters

The two basic plasma parameters are [27]:

- (1) particle number density, n_e for the electrons, n_i for ions, and n_n for neutrals,
- (2) temperature of each constituent, T_e for electrons and T_i for ions.²

The rest of the plasma parameters are derived from these two.

Plasmas interact primarily through collective, electromagnetic interactions. To derive the natural timescale for these interactions, consider the displacement of a slab of electrons by δx relative to the ions. The charge separation will result in an electric field given by Gauss' Law $E_x = en_e \delta x / \epsilon_0$, where e is the fundamental charge and ϵ_0 is the permittivity of free space. The force on an electron is $F_x = -eE_x$ and the equation of motion for the electron, with mass given by m_e , is $\ddot{\delta x} = -\frac{n_e e^2}{m_e \epsilon_0} \delta x$, which is the equation for a simple harmonic oscillator with frequency

$$\omega_p = \sqrt{\frac{n_e e^2}{m_e \epsilon_0}}. \quad (2.1)$$

The quantity ω_p is the electron plasma frequency, commonly referred to as the *plasma frequency*. The ion motion evolves on a different time scale given by $\omega_{pi}^2 = n_i Z^2 e^2 / (m_i \epsilon_0)$, where Z is the charge state of the ion. ω_{pi} is smaller than the electron plasma frequency by a factor of $\sqrt{Z^2 m_e / m_i}$. Because the ion mass is typically much larger than the electron mass, the ion motion can often be ignored.³

Collective, electromagnetic effects will dominate over collisions if the plasma period is much shorter than the average time, τ_n , between electron-neutral collisions. This leads to the first plasma criterion

$$\omega_p \tau_n > 1. \quad (2.2)$$

² Plasmas are rarely in thermal equilibrium. They often have non-Maxwellian velocity distributions, and the thermodynamic temperature is not well defined. In plasma physics, temperature is typically defined as $T_e = \frac{m_e}{3k_B n_e} \int (\mathbf{v} - \mathbf{v}_b) \cdot (\mathbf{v} - \mathbf{v}_b) f(\mathbf{v}) d^3 v$, where $f(\mathbf{v})$ is the velocity distribution and $\mathbf{v}_b$ is the bulk flow velocity [25].

³ Ion motion has little impact on the formation of the plasma wake. The small ion motion that develops near the rear of the wake, however, can have a significant impact on the dynamics of the witness beam, see Sec. 3.5.2.

The electron-neutral collision time is given by $\tau_n = \frac{1}{n_n \sigma_n} \sqrt{\frac{m_e}{2k_B T_e}}$, where $\sigma_n \approx 10^{-19} \text{ m}^2$ is the neutral atomic cross section and k_B is the Boltzmann constant [25, 26]. To ensure the first plasma criterion is satisfied in a plasma wakefield accelerator, we need to introduce the typical scales present in such a device.

Plasma wakes are excited by charged particle beams or laser pulses, both of which travel at close to the speed of light. The natural length scale for the wake is then the plasma oscillation period times the speed of light c :

$$\lambda_p = 2\pi \frac{c}{\omega_p}, \quad (2.3)$$

called the *plasma wavelength*; although, the *plasma wavenumber* $k_p = \omega_p/c$ is more frequently used. The driver of the wakefield must be about the same size as the plasma wavelength to efficiently drive the wake.⁴ Electron bunches and laser drivers typically have length scales around $10 - 100 \mu\text{m}$, while proton drivers produce wakes with $\lambda_p \approx 1 \text{ mm}$. As a result, the plasmas used in wakefield acceleration have densities between 10^{15} cm^{-3} and 10^{19} cm^{-3} . For comparison, the density of air at room temperature is approximately $3 \times 10^{19} \text{ cm}^{-3}$. The corresponding plasma frequencies vary between 0.9 THz and 90 THz. The typical electron temperature before the wake passes ranges from 1 – 100 eV. We are concerned with plasmas that are partially ionized at the edges of the plasma source. Even when the ionization fraction is as low as 1%, $T_e = 100 \text{ eV}$, and with a gas density of 10^{19} cm^{-3} , $\omega_p \tau_n \approx 10$ and the first plasma criterion is satisfied.

Plasmas appear quasineutral on length scales greater than the Debye length, over which the electric field is screened by the motion of charges in the plasma. There are contributions to the Debye length from both the electrons and ions, but, in the plasmas considered here, the electron temperature is typically much higher than the ion temperature and the Debye length can be approximated as [25, 27, 26]

$$\lambda_D = \sqrt{\frac{\epsilon_0 k_B T_e}{n_e e^2}}. \quad (2.4)$$

A plasma must be large enough for shielding to occur in order for it to be quasineutral. The second

⁴ The exception is proton beam drivers which excite many wake periods over the length of the bunch using the self-modulation instability. The typical plasma wavelength is around 1 mm for these beams.

plasma criterion is thus

$$\lambda_D \ll L, \quad (2.5)$$

where L is characteristic the spatial dimension of the system. For wakefield accelerator plasmas, the Debye length varies from 2.4 nm for a 10^{19} cm^{-3} plasma at 1 eV up to $2.4 \mu\text{m}$ for a 10^{15} cm^{-3} plasma at 100 eV. In any case, the plasma dimensions are at least several λ_p , which is significantly larger than the Debye length; the second plasma criterion is thus satisfied.

As we have seen above, screening of the electric field occurs on the scale of the Debye length. The screening results from the movement of particles in response to potentials in the plasma; as such, there needs to be enough particles present to do the screening. This requirement is quantified by the number of particles within a Debye sphere [25, 27, 26]:

$$N_D = \frac{4}{3}\pi n_e \lambda_D^3. \quad (2.6)$$

The final plasma criterion is $N_D \gg 1$. For a 10^{19} cm^{-3} plasma at 1 eV, there is about half a particle in the Debye sphere; while at 100 eV, there are 544. To strictly meet the plasma criterion, our higher density plasmas must also be hotter. It turns out, however, that this is irrelevant for the plasma physics discussed here. As a final note, the $n_e \lambda_D^3$ part is called the plasma parameter $\Lambda = n_e \lambda_D^3$.

2.1.3 Linear Plasma Waves

In the previous section, we derived the oscillation frequency of a slab of electrons displaced from equilibrium. In this section, we move beyond the simple slab geometry and consider a general electrostatic wave. We will be limited to the linear regime where the variations in the electron density are small, allowing the magnetic field to be ignored. To connect this with wakefield acceleration, we will consider a source in the form of an ultrarelativistic electron beam moving through the plasma at approximately the speed of light, for which we can calculate the electron density and electric fields analytically. We follow the approach presented in Keinigs and Jones [28] with

additional details from Refs. [29, 30, 31]. The ions are assumed to be stationary, and the plasma density is initially uniform and spatially infinite.

We use a fluid approach to describe the evolution of the plasma. The evolution is constrained by the continuity equation

$$\frac{\partial n_e}{\partial t} + \nabla \cdot (n_e \mathbf{v}) = 0, \quad (2.7)$$

where $\mathbf{v}$ is the electron fluid velocity. The continuity equation ensures charge is conserved in the absence of sources or sinks of electrons in the plasma.⁵ The force on a fluid element is given by the Lorentz force law

$$m_e \frac{\partial(n_e \mathbf{v})}{\partial t} = -en_e (\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad (2.8)$$

where $\mathbf{E}$ is the electric field and $\mathbf{B}$ is the magnetic field. We have assumed the plasma electron velocity is non-relativistic. The fields respond to the plasma density according to Maxwell's equations:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (2.9)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.10)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.11)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}, \quad (2.12)$$

where the charge density is given by $\rho = e(n_i - n_e)$ and the current density is $\mathbf{J} = -en_e \mathbf{v}$. These six equations provide a self-consistent description of the plasma density and electromagnetic field dynamics.

In the linear—or small amplitude—limit, we treat the change in plasma density as a perturbation $n_e = n_0 + \delta n$, with $\delta n \ll n_0$ ($n_i = n_0$). Further, we assume the plasma currents are small enough that the magnetic field is negligible compared to the electric field ($v \ll c$). Expanding the continuity equation to first order in δn gives

$$\frac{\partial \delta n}{\partial t} + n_0 \nabla \cdot \mathbf{v} = 0. \quad (2.13)$$

⁵ Ionization is an example of an electron source. Recombination would be a sink.

We ignore the magnetic field in the Lorentz force law:

$$m_e \frac{\partial \mathbf{v}}{\partial t} = -e \mathbf{E}. \quad (2.14)$$

Similarly, in Maxwell's equations we remove the magnetic field and consider Gauss's law

$$\nabla \cdot \mathbf{E} = -\frac{e}{\epsilon_0} \delta n. \quad (2.15)$$

Taking the divergence of Eq. (2.14) gives

$$\frac{\partial}{\partial t} \nabla \cdot \mathbf{v} = -\frac{e}{m_e} \nabla \cdot \mathbf{E}, \quad (2.16)$$

inserting Gauss's Law for $\nabla \cdot \mathbf{E}$ and Eq. (2.13) for $\nabla \cdot \mathbf{v}$ gives the equation of motion for the plasma

$$\frac{\partial^2 \delta n}{\partial t^2} + \frac{n_0 e^2}{m_e \epsilon_0} \delta n = 0. \quad (2.17)$$

The wave is electrostatic, relying only on Gauss's law. The oscillation frequency is recognizable as the plasma frequency.

The plasma wave must be excited by an external source. We will consider an electron beam, called the *drive beam*. The resulting plasma density wave is called the *wake* of the beam and the electric and magnetic fields produced by the wave are called the *wakefields*. The charge density in Gauss's law is modified by the addition of the beam charge density n_b . The equation of motion gains a forcing term:

$$\frac{\partial^2 \delta n}{\partial t^2} + \omega_p^2 \delta n = -\omega_p^2 n_b. \quad (2.18)$$

Assume the electron beam is relativistic with a Lorentz factor $\gamma_b = 1/\sqrt{1 - v_b^2/c^2}$ and is travelling in the positive z direction. The beam velocity is given by $\beta = v_b/c$. Because the beam is relativistic, the particles have an effective mass of $\gamma_b m_e$ and, therefore, evolve slowly compared to the wake. We can, therefore, make the *quasistatic approximation*: The beam dynamics are much slower than the plasma dynamics so the beam distribution can be treated as static in the calculation of the plasma response.⁶

⁶ The beam dynamics will be treated in detail in Section 2.3. The length scale for the beam evolution in a highly nonlinear wave is $k_\beta = k_p \sqrt{2\gamma_b}$, which is a factor of $\sqrt{2\gamma_b}$ longer than the dynamics of the plasma electrons. In the linear waves presented here, the length scale for the beam dynamics is even longer.

It is convenient to transform to a coordinate system that co-moves with the beam.⁷ We define $\xi = v_b t - z$ as the co-moving coordinate and we transform from the coordinates (x, y, z, t) to $(x, y, \xi, s \equiv z)$. Rewriting the equation of motion in terms of ξ :

$$\frac{\partial^2 \delta n}{\partial \xi^2} + \frac{\omega_p^2}{v_b^2} \delta n = -\frac{\omega_p^2}{v_b^2} n_b. \quad (2.19)$$

The plasma density perturbation behaves like a driven, simple harmonic oscillator in the s direction. The transverse coordinates only enter through the driving term - each transverse position is decoupled and driven by the local beam at that position. We can use a Green's function approach to solve the problem for an arbitrary beam density at a given (x, y) position. The Green's function for Eq. (2.19) is

$$G(\xi, \xi') = -\frac{\omega_p}{v_b} \sin \left[\frac{\omega_p}{v_b} (\xi - \xi') \right] \Theta(\xi' - \xi), \quad (2.20)$$

where $\Theta(-\xi)$ is the Heavyside step function; physically, $\Theta(-\xi)$ means the electron beam cannot drive a wake in front of itself. The Green's function is verified in Appendix A. The density perturbation driven by the beam is given by

$$\delta n(x, y, \xi) = -\frac{\omega_p}{v_b} \int_{-\infty}^{\xi} d\xi' n_b(x, y, \xi') \sin \left[\frac{\omega_p}{v_b} (\xi - \xi') \right]. \quad (2.21)$$

From here on, we assume the drive beam is highly relativistic so we can approximate $v_b \approx c$ and $\omega_p/v_b = k_p$. The wavelength of the wake is then λ_p .

We can draw several immediate conclusion from Eq. (2.21). First, there are two major regimes depending on the length of the drive beam compared to λ_p .⁸ If the length of the drive beam is much longer than λ_p , we integrate over multiple periods of the sine function at large ξ , which tends to zero. Physically, the plasma electrons have sufficient time to move in order to shield the fields from the beam; minimal density perturbations persist after the beam has passed. This shielding is shown in Fig. 2.1(a) for a beam with a Gaussian longitudinal profile of RMS width σ_ξ . If, on the

⁷ The co-moving coordinate system is not a transformation to a different frame. We are still in the lab frame but using a more convenient set of coordinates.

⁸ Strictly speaking, it is the characteristic lengths of features in n_b rather than the total beam length. For example, a long beam with a very sharp drop off at the tail will excite a wake in the short beam regime because the sharp drop off has a characteristic length shorter than λ_p .

other hand, the beam is shorter than λ_p , it will drive a sinusoidal density wave that moves along with the beam and the oscillations will persist after the beam has passed, as shown in Fig. 2.1(b). In our simple model, the wave never decays; in a real plasma there are various dampening mechanisms that will cause the wave amplitude to reduce over time.⁹ Second, the maximum value of δn is n_b if the beam is long and $k_p \int d\xi n_b$ if the beam is short. At the start of the derivation, we assumed $\delta n \ll n_0$. The wake will only be in the linear regime if the beam density is much less than the plasma density, $n_b \ll n_0$.

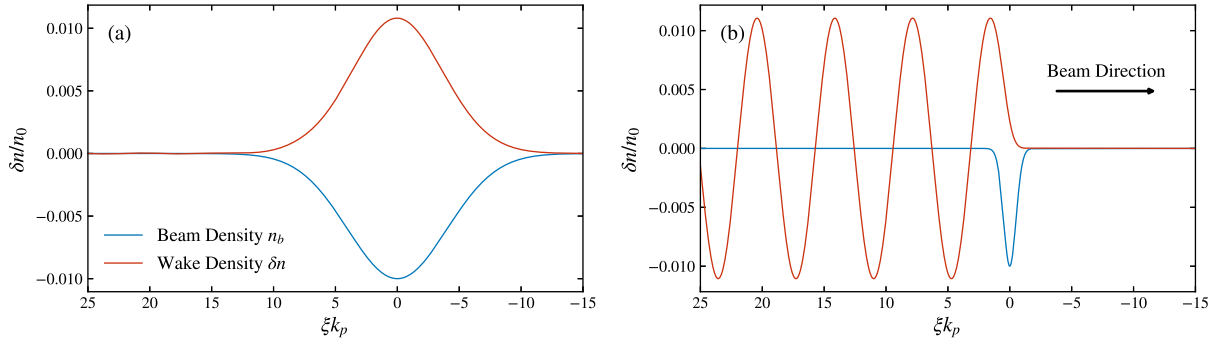


Figure 2.1: Plasma density waves along the z axis in the linear regime. (a) shows a plasma density wave driven by a long bunch with $\sigma_\xi k_p = 4$. The plasma reacts sufficiently fast to screen the beam fields. (b) shows a density perturbation driven by a short bunch ($\sigma_\xi k_p = 0.5$). Even after the beam has passed, the plasma density continues to oscillate. Adapted from Ref. [31].

Eq. (2.21) is the general solution for a linear density wave excited by an electron beam. The next step is to calculate the electric field within the wave. We start by presenting the derivation of the inhomogeneous wave equation [32]. Taking the curl of Eq. (2.11) we get

$$\nabla \times (\nabla \times \mathbf{E}) = -\frac{\partial}{\partial t} \nabla \times \mathbf{B}. \quad (2.22)$$

Using Eq. (2.12) to replace the $\nabla \times \mathbf{B}$ gives

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu_0 \frac{\partial \mathbf{J}}{\partial t} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (2.23)$$

⁹ Landau (collisionless) dampening and dampening due Coulomb collisions or collisions with neutrals will dissipate the wake. For more details see a plasma text i.e. [25, 26, 27].

We use the vector identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ to get

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu_0 \frac{\partial \mathbf{J}}{\partial t} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (2.24)$$

Finally, using Gauss's law to substitute for $\nabla \cdot \mathbf{E}$, we get the inhomogeneous wave equation

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{E} = \frac{1}{\epsilon_0} \nabla \rho + \mu_0 \frac{\partial \mathbf{J}}{\partial t}. \quad (2.25)$$

We can get a similar equation for the magnetic field by taking the curl of Eq. (2.25), using Faraday's Law, Eq. (2.11), to swap the electric field for the magnetic field, and integrating in time:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{B} = -\mu_0 \nabla \times \mathbf{J}. \quad (2.26)$$

The sources are the same as above $\rho = -e(\delta n + n_b)$ and $\mathbf{J} = -e(n_0 \mathbf{v} + n_b c \hat{z})$. Making use of Eq. (2.14) to replace the time derivative of $\mathbf{v}$ gives

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - k_p^2 \right) \mathbf{E} = -\frac{e}{\epsilon_0} \nabla(\delta n + n_b) - \frac{e}{\epsilon_0 c^2} c \hat{z} \frac{\partial n_b}{\partial t}, \quad (2.27)$$

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - k_p^2 \right) \mathbf{B} = \mu_0 e c \nabla \times (n_b \hat{z}). \quad (2.28)$$

Transforming to the co-moving coordinate causes the z part of $\nabla^2 \mathbf{E}$ to cancel the $\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}$ term; the same thing happens in the equation for $\mathbf{B}$. This cancellation only occurs because we have assumed $v_b \approx c$. We note that a similar cancellation happens between ∇n_b and $\frac{\hat{z}}{c} \frac{\partial n_b}{\partial t}$. Again, this cancellation is due to the fact that we have assumed the electron beam is travelling at the speed of light. This is the well-known relativistic reduction (by a factor $1/\gamma_b^2$) of the longitudinal electric field.¹⁰ The radial fields are magnified and, thus, a highly relativistic electron beam only looks like a source in the transverse direction [32]. Finally, the quasistatic approximation allows us to drop the s derivative in $\frac{\partial}{\partial z}$ as $\frac{\partial}{\partial s} \ll \frac{\partial}{\partial \xi}$.

In the co-moving coordinates, the wave equations are given by

$$(\nabla_{\perp}^2 - k_p^2) \mathbf{E} = -\frac{e}{\epsilon_0} \nabla \delta n - \frac{e}{\epsilon_0} \nabla_{\perp} n_b, \quad (2.29)$$

$$(\nabla_{\perp}^2 - k_p^2) \mathbf{B} = \mu_0 e c \left(\frac{\partial n_b}{\partial y} \hat{x} - \frac{\partial n_b}{\partial x} \hat{y} \right). \quad (2.30)$$

¹⁰ By assuming the beam has speed $v_b = c$, we have assumed γ_b is infinite and we thus get perfect cancellation of the longitudinal electric field from the beam. It is exactly 0. See Sec. 2.3.5 for a derivation of the beam self fields.

The first source term on the right of Eq. (2.29) is the charge separation of the plasma, the second term is the radial electric field from the electron beam. The source term on the right of Eq. (2.30) is the beam current. The plasma current is present for both $\mathbf{E}$ and $\mathbf{B}$ in the k_p^2 term.

We are interested in the fields created by the wake, not the driving beam. The wave equations are linear, so the total fields are the sum of the fields from the plasma source terms and the fields from the beam source terms. Keeping only the plasma source terms we get

$$(\nabla_{\perp}^2 - k_p^2) \mathbf{E} = -\frac{e}{\epsilon_0} \nabla \delta n, \quad (2.31)$$

$$(\nabla_{\perp}^2 - k_p^2) \mathbf{B} = 0. \quad (2.32)$$

Trivially, the solution for the magnetic field is $\mathbf{B} = 0$; the wave is electrostatic after all. The equation for the electric field can be solved independently in the longitudinal and transverse directions. To simplify the solution, we assume the electron beam is cylindrically symmetric; δn is then also cylindrically symmetric. We start with the longitudinal direction where, in cylindrical coordinates, we have

$$(\nabla_{\perp}^2 - k_p^2) E_z = -\frac{e}{\epsilon_0} \frac{\partial \delta n}{\partial \xi}. \quad (2.33)$$

The radial electric field is given by

$$\left(\nabla_{\perp}^2 - \frac{1}{r^2} - k_p^2 \right) E_r - \frac{2}{r^2} \frac{\partial E_{\phi}}{\partial \phi} = -\frac{e}{\epsilon_0} \frac{\partial \delta n}{\partial r}, \quad (2.34)$$

where we have used the radial part of the vector Laplacian in cylindrical coordinates:

$$(\nabla^2 \mathbf{E})_r = \left(\nabla^2 E_r - \frac{E_r}{r^2} - \frac{2}{r^2} \frac{\partial E_{\phi}}{\partial \phi} \right). \quad (2.35)$$

To make further progress, we will assume the electron beam, and thus the plasma response, is cylindrically symmetric and separable: $E_z(r, \xi) = R(r)Z(\xi)$ and $n_b(r, \xi) = R_b(r)Z_b(\xi)$. If the beam density is separable, then the density perturbation is also separable according to Eq. (2.21): $\delta n(\xi, r) = R_n(r)Z_n(\xi)$. Inserting into Eq. (2.33), we get

$$Z (\nabla_{\perp}^2 - k_p^2) R = -\frac{e}{\epsilon_0} R_n \frac{\partial Z_n}{\partial \xi} \quad (2.36)$$

The longitudinal solution is trivial and is given by

$$Z(\xi) = \frac{e}{\epsilon_0 k_p^2} \frac{\partial Z_n(\xi)}{\partial \xi}. \quad (2.37)$$

The Laplacian in cylindrical coordinates is

$$\nabla_{\perp}^2 R = \frac{1}{r} \frac{\partial R}{\partial r} + \frac{\partial^2 R}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 R}{\partial \phi^2}, \quad (2.38)$$

with which we can write the radial equation as

$$\frac{\partial^2 R}{\partial r^2} + \frac{1}{R} \frac{\partial R}{\partial r} - k_p^2 R = -k_p^2 R_n, \quad (2.39)$$

which is the inhomogeneous Bessel's equation of zero order with a driving force $-\frac{e}{\epsilon_0} R_n$. The Green's function solution to this equation comes from Jackson Sec. 3.11 [32]

$$g(r, r') = 4\pi K_0(k_p r_{>}) I_0(k_p r_{<}), \quad (2.40)$$

where $r_{>}$ is the greater of r and r' and $r_{<}$ is the lesser. In this case the Green's function satisfies

$$\frac{\partial^2 g}{\partial r^2} + \frac{1}{g} \frac{\partial g}{\partial r} - k_p^2 g = -\frac{4\pi}{r} \delta(r - r') \quad (2.41)$$

so the radial solution is given by

$$R(r) = k_p^2 \int_0^{\infty} r' dr' R_n(r') K_0(k_p r_{>}) I_0(k_p r_{<}). \quad (2.42)$$

Combining everything together gives the longitudinal electric field as

$$E_z(r, \xi) = \frac{e}{\epsilon_0} \int_0^{\infty} r' dr' \frac{\partial}{\partial \xi} \delta n(r', \xi) K_0(k_p r_{>}) I_0(k_p r_{<}). \quad (2.43)$$

The radial electric field can be found using a similar process. The only major difference enters from the additional $1/r^2$ term in Eq. (2.34) which results in a Bessel function of the first order. The radial electric field is given by

$$E_r(r, \xi) = \frac{e}{\epsilon_0} \int_0^{\infty} r' dr' \frac{\partial}{\partial r'} \delta n(r', \xi) K_1(k_p r_{>}) I_1(k_p r_{<}). \quad (2.44)$$

These are only the fields generated by the plasma wave—the driving electron beam produces its own fields which are discussed in Section 2.3.5.

Fig. 2.2 shows the beam density, plasma density and electric fields from the plasma. The drive beam has a cylindrically symmetric, 3-dimensional Gaussian distribution given by

$$n_b(r, \xi) = n_{b0} e^{-\xi^2/(2\sigma_\xi^2) - r^2/(2\sigma_r^2)}. \quad (2.45)$$

In the figure, $\sigma_\xi = 0.5/k_p$, $\sigma_r = 1/k_p$, and $n_{b0} = -0.01n_0$. The electric field is normalized to $m_e c \omega_p / e$. We can explain why this is the appropriate factor using simple dimensional analysis. Because the wave is electrostatic, the electric field is the energy gained per unit length per unit charge of an electron in the wave. The plasma wave moves at speed c so the kinetic energy scale for an electron in the wave is order $m_e c^2$.¹¹ The natural length scale of this energy change is the plasma wavelength $1/k_p = c/\omega_p$. Dividing the energy scale by the length scale and the electron charge gives the normalization factor for the electric field.

A couple things to note from Fig. 2.2: The electric fields vary sinusoidally in the longitudinal direction behind the beam. The radial field is in phase with the density perturbation, while the longitudinal field is $\pi/2$ out of phase. The longitudinal field has a transverse structure similar to that of the plasma density. The field is peaked on-axis and falls off as r increases, but the falloff is much slower than the density falloff of the plasma wave. The magnitude of the radial electric field grows from zero on-axis up to some maximum value around the edge of the plasma wave, past which it decays towards zero. We will revisit these fields when we talk about the linear plasma wakefield accelerator in Sec. 2.2.1.

2.2 Plasma Wakefield Acceleration

Producing large electric fields in a particle accelerator requires two things:

- (1) a cavity or medium to support the electric fields long enough for the beam to pass through,
- (2) a source for the electric fields.

¹¹ The approximate kinetic energy is $\frac{\delta n}{n_0} m_e c^2$. We assumed the electrons are non-relativistic which means they must have energy much less than $m_e c^2$. Since we assumed $\delta n/n_0 \ll 1$ the non-relativistic assumption is automatically satisfied.

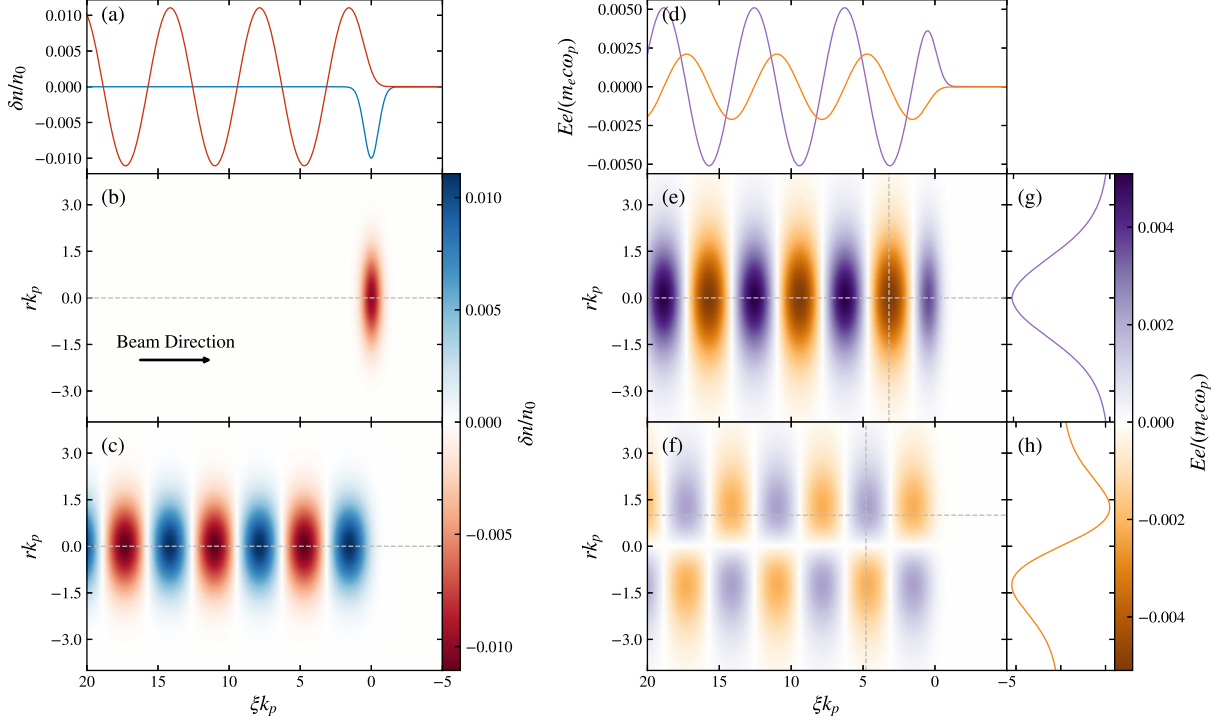


Figure 2.2: Beam and plasma number density of a linear plasma wave and the electrostatic fields generated by the wave. (a) shows the on-axis density of the driving electron beam and the resulting plasma wave. (b) is the density of the cylindrically symmetric Gaussian drive beam. (c) is the density of the excited plasma wave. (d) shows longitudinal lineouts of the fields. In purple is the on-axis longitudinal field and in orange is the off-axis radial field—the on-axis radial field is 0—notice that the two fields are $\pi/2$ out of phase. (e) and (g) show the longitudinal field and the transverse variation in the field, while (f) and (h) show the same thing for the radial field. Dashed lines denote the locations where the lineouts are taken

Conventional accelerators use copper or superconducting cavities to support the electric fields and are driven by RF sources with a frequency between 1 – 3 GHz. The problem with conventional accelerators is that the copper or superconductor can only support relatively weak fields. Copper can only handle electric fields up to about 50 MV/m before it is damaged by dielectric breakdown [1]. Plasma, on the other hand, has already been ionized and can support very large fields before further ionization occurs. This is the key difference that allows the plasma-based accelerator to achieve accelerating gradients three orders of magnitude larger than a conventional accelerator.

Using a plasma removes the main limitation: supporting the strong electric fields. RF sources,

however, are not suitable for driving high accelerating fields in a plasma. In 1979, Tajima and Dawson [3] realized that the pondermotive force of a laser pulse could excite electrostatic density waves with fields capable of trapping and accelerating charged particles to high energies. The plasma wakefield accelerator is composed of a driver beam (laser or particle) travelling at close to the speed of light that excites a density wave inside a plasma. Charged particles are then either launched or injected into the wakefields and accelerated to high energy. At the same time, the wake and wakefields act back on the driver reducing its energy. In this way, the wakefield accelerator acts as a very space efficient energy transformer—it transfers energy from the driver into the plasma and then from the plasma into the accelerated particles. Compare this to RF accelerators: they work by transferring energy from a RF source to the beam via metallic waveguides but do so over a much greater distance.

There are two main types of plasma wakefield accelerators depending on what drives the wake. As in Tajima and Dawson’s original suggestion, a laser can drive the wake in which case the accelerator is called a laser wakefield accelerator (LWFA). The second technique, originally proposed in Refs. [4, 5], is to use a charged particle beam, called a beam driven plasma wakefield accelerator (PWFA). The type of driver influences the shape of the wake and wakefields in the vicinity of the driver. If the driver is short, the wakefield structure in the rear of the wake is the same regardless of whether a laser or particle beam drives the wake. The only dependence on the details of the driver enters through the magnitude of the wake. As a result, the dynamics of the particles being accelerated are identical for the two acceleration techniques.

Many early PWFA experiments used very long beams to simultaneously drive the wake and sample the accelerating fields [33, 34, 6]. A practical plasma wakefield accelerator will operate in the two bunch configuration where a drive beam drives the wake in the plasma while the wake accelerates a second electron beam, called the *witness beam*. In an LWFA, the accelerated electrons are also called a witness beam. As long as the witness beam sits rearward of the laser fields, it experiences the same dynamics as in a PWFA.

In the preceding section, we discussed the formation of linear waves in a plasma. This is the

linear plasma wakefield accelerator; it has several less-than-ideal characteristics. We will discuss it briefly to introduce the concepts of beam loading and transformer ratio; we will also discuss the disadvantages of the linear regime. As the density of the particle driver or intensity of the laser driver increases, the linear approximations break down and the wave becomes nonlinear. With enough driving force, all of the electrons are evacuated from the center of the wake, and we enter the nonlinear blowout regime. After Section 2.2.1, the rest of this dissertation will assume the plasma-based accelerator is operating in the blowout regime.

2.2.1 Plasma Wakefield Acceleration in the Linear Regime

In Sec. 2.1.3 we discussed the formation of linear waves in a plasma. In particular, we considered waves driven by a relativistic electron beam. We were already dangerously close to describing a plasma wakefield accelerator; the only step remaining is to add a witness beam into the wave to be accelerated. We can include a second beam in our previous equations by adding the witness beam charge density into the total beam density $n_b = n_d + n_w$, where n_d and n_w are the charge densities of the drive and witness beam respectively. Because the wave is linear, the electron density and the resulting fields are simply the sum of the densities and fields from each individual beam. As such, we start by considering the effect of the wakefields on a single beam.

We already know that the beam must have features with length scales shorter than λ_p to generate a wake. The functional form of the distribution determines how E_z , the accelerating field, varies along the beam. We can combine Eqs. (2.21) and (2.43) to get [35]

$$E_z(r, \xi) = -k_p^2 \frac{e}{\epsilon_0} \int_0^\infty r' dr' \int_{-\infty}^\xi d\xi' n_b(r', \xi') \cos[k_p(\xi - \xi')] K_0(k_p r_>) I_0(k_p r_<). \quad (2.46)$$

We can take advantage of the cylindrical symmetry to extract the dependence of the field on the longitudinal beam profile, i.e., the beam current profile

$$E_z(r, \xi) = -\frac{e}{\epsilon_0} R(r) \int_{-\infty}^\xi d\xi' Z_b(\xi') \cos[k_p(\xi - \xi')], \quad (2.47)$$

where $R(r)$ is given by Eq. (2.42). The radial function $R(r) \approx R_b(r)$ when the beam is much wider than $1/k_p$. The fields are reduced for narrower beams; consult Katsouleas *et al.* Eq. (25)

for the expression of the reduction in field strength from a cylindrical beam [35]. It is convenient to separate n_b in such a way that $R_b(r)$ is unit-less and $R_b(0) = 1$. Consider a uniform cylindrical beam with charge Q , length $\Delta\xi$, and radius Δr ; the beam density is $n_{b0} = \frac{Q}{e\pi\Delta\xi\Delta r^2}$ and the beam profile is $Z_b(\xi) = n_{b0} \text{rect}\left(\frac{\xi}{\Delta\xi} - \frac{1}{2}\right)$. The field within the beam is given by

$$E_z(r, \xi) = -\frac{e}{\epsilon_0} R(r) \frac{1}{k_p} n_{b0} \sin(k_p \xi). \quad (2.48)$$

The front of the beam sees a smaller amplitude field than the rear; thus, the wakefield will extract more energy from the rear of the beam than the front. In addition, we see that for a uniform current distribution, $\Delta\xi < \lambda_p/2$ to efficiently drive a wake. The wakefield driven by the beam is

$$E_z(r, \xi) = -\frac{e}{\epsilon_0} R(r) \frac{1}{k_p} n_{b0} \sqrt{2 - 2\cos(k_p \Delta\xi)} \cos(k_p \xi + \phi), \quad (2.49)$$

where ϕ is the phase of the wake at the end of the beam:

$$\phi = \arctan \left[-\frac{1 - \cos(k_p \Delta\xi)}{\sin(k_p \Delta\xi)} \right]. \quad (2.50)$$

The amplitude of the wake can then be identified as

$$E_{z0} = -\frac{e}{\epsilon_0} R(0) \frac{1}{k_p} n_{b0} \sqrt{2 - 2\cos(k_p \Delta\xi)}. \quad (2.51)$$

Fig. 2.3(a) shows the longitudinal wakefield from several different drive beams all with the same charge, but different bunch lengths and thus different currents. For a given charge, the larger the beam current, the higher the amplitude of the wake function. Fig. 2.3(b) shows a similar figure except the beam current is held constant and the beam charge is varied. To maximize the strength of the wake, the beam should have as much charge in as short a length as possible.

Drive beams with uniform current profiles do not lose energy uniformly along their length. This makes them inefficient because the center of the beam runs out of energy before the ends. Optimally the decelerating field would be uniform, or close to uniform, along the full length of the beam; then, all the energy from the drive beam could be transferred to the plasma. Similarly, a uniform accelerating field along the extent of the witness beam is important to minimize the energy spread of the particles in the beam. To create uniform fields, we can change the shape of

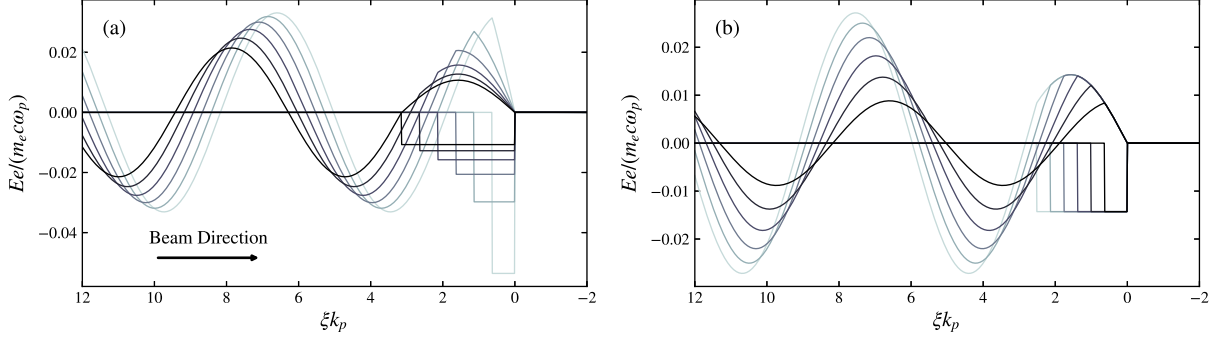


Figure 2.3: Longitudinal wakefields for constant current beams with different charges and peak currents. The rectangles show the current profile of the electron beam. In (a) the beam charge is held fixed while the bunch length, and thus current, is varied. Larger currents produce stronger wakefields for a given beam charge. (b) shows several beams all with the same current but varying charge. Unsurprisingly, more charge produces a stronger wakefield.

the current profile to change the form of the wakefields. This effect is called *beam loading* and it is a fundamental concept in plasma wakefield acceleration.

For the drive beam, it is impossible to maintain a constant field along the entire beam because the field must rise continuously from 0 at the front. A small uniform current beam, like that presented above, can be used to ramp the field up before the main beam. For both the main drive beam and the witness beam, a trapezoidal current profile can flatten the wake when properly shaped [35]. Assuming a beam profile of $Z_b(\xi) = a\xi + b$ between ξ_0 and $\Delta\xi$ and using Eq. (2.47), the field within the beam is found to be

$$E_z(r, \xi) = -\frac{e}{\epsilon_0} R(r) \frac{1}{k_p^2} \{a - a \cos[k_p(\xi - \xi_0)] + k_p(b + a\xi_0) \sin[k_p(\xi - \xi_0)]\}. \quad (2.52)$$

We will drop the radial dependence $R(r)$ for the rest of this discussion. The field created by the bunch summed with the field from the preceding bunch should be a constant. The longitudinal wakefield in the regions devoid of beam electrons is generically given by

$$E_{wave}(\xi) = E_{z0} \cos(k_p \xi + \phi) \quad (2.53)$$

Summing Eq. (2.53) and Eq. (2.52), and setting the sum equal to an arbitrary constant allows us to solve for a and b . We start by changing coordinates to $\zeta = \xi + \phi/k_p$ to simplify the expression.

The constants a and b are then given by

$$a = -\frac{E_{z0}k_p^2\epsilon_0}{e}\cos(k_p\zeta_0) \quad (2.54)$$

$$b = -\frac{E_{z0}k_p^2\epsilon_0}{e}\left[\frac{\sin(k_p\zeta_0)}{k_p} - \zeta_0\cos(k_p\zeta_0)\right] \quad (2.55)$$

The accelerating field within the beam is constant and given by

$$E_{beam} = E_{z0}\cos(k_p\zeta_0 + \phi). \quad (2.56)$$

We can work out the fields in the accelerator analytically for a “doorstep” drive beam and a triangular witness beam [36, 37, 38, 35]. The drive beam current profile is given by $Z_b(\xi) = n_{b0}$ for $0 < \xi \leq \pi/(2k_p)$ and $Z_b(\xi) = k_p n_{b0}\xi + n_{b0}(1 - \pi/2)$ for $\pi/(2k_p) < \xi \leq 2\pi N/k_p$, where N is an integer.¹² We will denote the initial uniform current section with the subscript “step” and the trapezoidal section with subscript “trap”. The wakefield rises from zero to $-\frac{en_{b0}}{\epsilon_0 k_p}$ following Eq. (2.48) inside the step and is described by Eq. (2.49) after the step:

$$E_{step}(\xi) = -\frac{en_{b0}}{\epsilon_0 k_p} \begin{cases} \sin(k_p\xi) & 0 < \xi \leq \frac{\pi}{2k_p} \\ \sqrt{2}\cos(k_p\xi - \pi/4) & \frac{\pi}{2k_p} < \xi \end{cases}, \quad (2.57)$$

where λ_0 is the beam current within the step. We already defined the trapezoidal part of the beam current above, but we can verify it flattens the wakefield by calculating a and b using Eqs. (2.54) and (2.55). This gives $a = k_p n_{b0}$ and $b = n_{b0}(1 - \pi/2)$. The field within the beam is given by Eq. (2.52) and the field after the beam is given by evaluating Eq. (2.47):

$$E_{trap}(\xi) = -\frac{en_{b0}}{\epsilon_0 k_p} \begin{cases} 1 - \sqrt{2}\cos(k_p\xi - \pi/4) & \frac{\pi}{2k_p} < \xi \leq 2\pi N \\ \left(\frac{\pi}{2} - 2 - 2\pi N\right)\sin(k_p\xi) & 2\pi N < \xi \end{cases}. \quad (2.58)$$

The N dependence means the wake amplitude can be driven arbitrarily high by increasing the length of the drive beam.¹³ Adding together E_{step} and E_{trap} in the region $\frac{\pi}{2k_p} < \xi \leq 2\pi N$ results

¹² A quarter period step is required for the beam current to be continuous while simultaneously maintaining a constant wakefield throughout the drive beam.

¹³ At some point, the wake amplitude will become so large the assumptions of linear theory will no longer be valid and the wake amplitude will saturate.

in the accelerating field experienced by the drive beam

$$E_{drive} = -\frac{en_{b0}}{\epsilon_0 k_p}, \quad (2.59)$$

which will reduce the energy of the beam.

The final step is to calculate the longitudinal wakefield excited by the witness beam. We have the freedom to choose the position where the witness beam starts. From Eq. (2.53), we can see that the field acting on the witness beam depends on the total wakefield from the drive beam $E_{wave} = E_{step} + E_{trap}$:

$$E_{wave}(\xi) = -\frac{en_{b0}}{\epsilon_0 k_p} \sqrt{1 + \left(1 - \frac{\pi}{2} + 2\pi N\right)^2} \cos(k_p \xi + \phi), \quad (2.60)$$

for $2\pi N < \xi$. Here ϕ is

$$\phi = \arctan\left(1 - \frac{\pi}{2} + 2\pi N\right). \quad (2.61)$$

We want the witness beam to be in the accelerating phase of the wake so

$$2\pi \left(N + M + \frac{1}{4}\right) < \xi_0 k_p + \phi < 2\pi \left(N + M + \frac{3}{4}\right), \quad (2.62)$$

where M is an integer describing which wake period the witness beam is placed in. We will use $M = 1$ from here on for simplicity. The field experienced by the witness beam can vary between 0 and $-\frac{en_{b0}}{\epsilon_0 k_p} \sqrt{1 + \left(1 - \frac{\pi}{2} + 2\pi N\right)^2}$. Define θ as the phase within a given wake period such that the field is proportional to $\sin \theta = \cos(k_p \xi + \phi)$. The witness beam needs to be placed between $\theta = \pi$ and $\theta = 2\pi$ to be in the accelerating phase of the wake. Recall that the focusing field of the wake is $\pi/2$ out of phase with the accelerating field; the witness beam is focused when $\pi/2 < \theta < 3\pi/2$. Fig. 2.4 shows the useful phase of the wake where the accelerating and focusing fields overlap.

The accelerating field is given in terms θ_0 , the value of θ at the start of the witness beam, by

$$E_{witness} = -\frac{en_{b0}}{\epsilon_0 k_p} \sqrt{1 + \left(1 - \frac{\pi}{2} + 2\pi N\right)^2} \sin(\theta_0). \quad (2.63)$$

As with the trapezoidal beam, we can solve for the current profile of the witness beam using Eqs. (2.54) and (2.55). The beam starts with a large current that reduces linearly in ξ . The

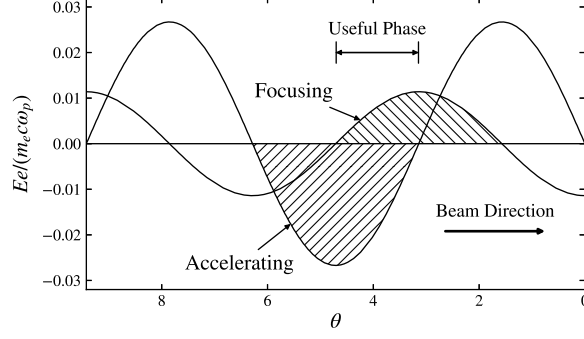


Figure 2.4: The transverse and longitudinal wake fields over one period in a plasma wakefield accelerator. The accelerator is only useful when the witness beam is placed in both the accelerating and focusing phase of the wake. Adapted from Ruth *et al.* [4].

maximum length of the beam is set by requiring that $Z_b > 0$, forming a triangular beam. The beam can be truncated before this point and will still see a uniform accelerating field. In Fig. 2.5, we show the beam current and longitudinal wakefield for a $N = 1$ drive beam with two different witness beams. By comparing the two beams in the figure, we can see that the larger the accelerating field within the witness beam, the shorter the beam length. The shorter beam corresponds to less charge being accelerated. Also, notice the difference in wakefield amplitude behind the witness beam—the amplitude is a measure of the energy remaining in the wake.

At this point we can define an important quantity in wakefield acceleration called the *transformer ratio*: the ratio of the accelerating field acting on the witness beam to the decelerating field acting on the drive beam $T = -E_{\text{witness}}/E_{\text{drive}}$. For our linear accelerator

$$T(\theta_0) = -\sqrt{1 + \left(1 - \frac{\pi}{2} + 2\pi N\right)^2} \sin(\theta_0). \quad (2.64)$$

The transformer ratio grows as the length of the drive beam is increased.¹⁴ A second important parameter is the *efficiency* of the accelerator, defined as the fraction of the energy lost by the drive beam that is transferred to the witness beam. To calculate the efficiency, we need to find the charge

¹⁴ It has been shown by Bane *et al.* [39] that a symmetric drive beam can at best produce a transformer ratio of 2. Our trapezoidal beam gets around this because it is asymmetric.

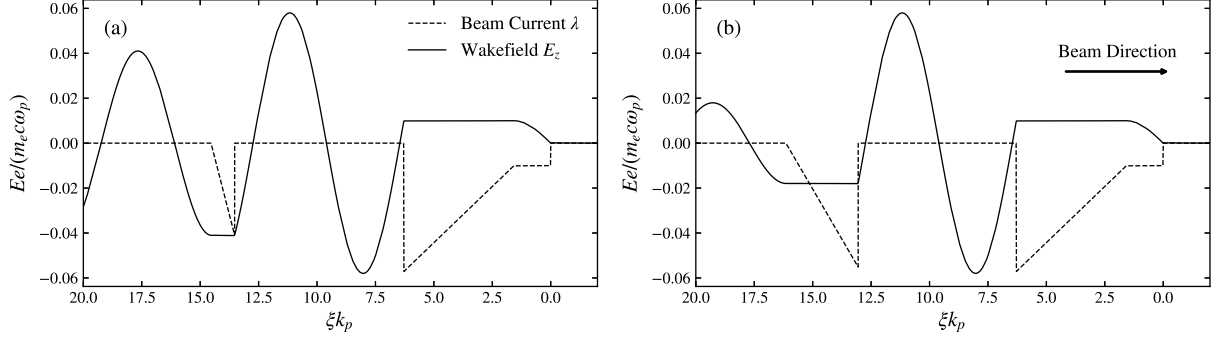


Figure 2.5: The longitudinal wakefield and beam current in a plasma wakefield accelerator. The drive beam has a “doorstep” current profile and the witness beam has a triangular current profile, both of which optimally load the wake. (a) the witness beam is placed at a larger accelerating field; consequently, the transformer ratio is large and the charge in the witness beam is low. (b) the witness beam is located at a small accelerating phase. The transformer ratio is small, but the accelerated charge is large. The efficiency can be estimated by looking at the residual wake after the witness beam. A large wake such as in (a) means a significant amount of energy remains in the wake and the efficiency is low.

in each beam. Integrating over the trapezoidal part of the drive beam current distribution gives

$$Q_{drive} = \frac{n_{b0}}{k_p} \left(\pi N + 1 - \frac{\pi}{4} \right) \left(2\pi N - \frac{\pi}{2} \right). \quad (2.65)$$

For the witness beam the charge is given by

$$Q_{witness} = -\frac{n_{b0}}{k_p} \sqrt{1 + \left(1 - \frac{\pi}{2} + 2\pi N \right)^2} \frac{\cos^2 \theta_0}{2 \sin \theta_0}. \quad (2.66)$$

The efficiency of the accelerator is given by

$$\eta(\theta) = -\frac{\int_{-\infty}^{\infty} Z_{witness}(\xi) E_z(\xi) d\xi}{\int_{-\infty}^{\infty} Z_{drive}(\xi) E_z(\xi) d\xi} = \cos^2 \theta, \quad (2.67)$$

where the integral over the witness beam is given by the product of Eqs. (2.63) and (2.66), and the integral over the drive beam requires integrating over the step and summing with the product of Eqs. (2.59) and (2.65). Fig. 2.6 shows the transformer ratio and efficiency of the accelerator as a function of θ_0 . There is always a trade-off between efficiency and transformer ratio.

The linear plasma wakefield accelerator has several less-than-ideal characteristics. First, the accelerating field depends on $R(r)$, particles closer to the axis experience stronger fields. To

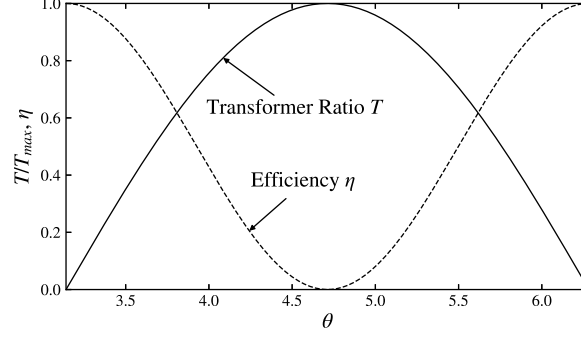


Figure 2.6: Transformer ratio and efficiency as a function of the witness beam’s phase in the wake.

minimize this effect, the witness beam needs to be much narrower than the drive beam, but then it becomes difficult to flatten the wake via beam loading for all r . Second, the focusing force is nonlinear in r resulting in degradation of the beam quality. Third, the linear regime is subject to the fundamental theory of beam loading which says that for a symmetric drive beam—the trapezoidal beams we considered earlier are asymmetric—the transformer ratio must be less than 2 [39]. This is a limitation because, in practice, it is hard to make doorstep beams, while symmetric Gaussian beams are common.

2.2.2 The Nonlinear Blowout Regime

As the density of the particle driver or intensity of the laser driver increases, the linear approximations break down, and the wave becomes nonlinear. With enough driving force, all the electrons are evacuated from the center of the wake and the nonlinear blowout regime is reached. We will assume we are in the blowout regime for the rest of this dissertation.

Linear theory breaks down when the density perturbation of the plasma wave becomes appreciable compared to the background plasma density. There are many advantages to operating in this regime. The most obvious—although arguably least important—is the strength of the accelerating electric field. In Sec. 2.1.3, we showed that the natural scale for the electric field is $m_e c \omega_p / e$ and in Sec. 2.2.1, we found fields with amplitudes of $E_{z0} \approx \frac{e \delta n}{\epsilon_0 k_p} = \frac{\delta n}{n_0} \frac{m_e c \omega_p}{e}$, where we have used the fact

that $\delta_n \approx n_{b0}$. The natural scale for the electric field is called the *wave breaking field* [40]

$$E_z = \frac{m_e c \omega_p}{e} = \sqrt{\frac{m_e c^2 n_0}{\epsilon_0}}, \quad (2.68)$$

which is the electric field that would result from complete charge separation in a wave with wavenumber k_p :

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \rightarrow k_p E_z = \frac{e n_0}{\epsilon_0}. \quad (2.69)$$

A highly nonlinear wake will produce fields on the order of the wave breaking field. For typical plasma densities in the range 10^{16} cm^{-3} to 10^{18} cm^{-3} , the peak accelerating field is on the order of 10 GV/m to 100 GV/m. The fields in the linear regime are, by definition, much less than the wave breaking field.

The primary advantages of operating an accelerator in the highly nonlinear regime are related to the uniformity of the fields. As in the linear case, we start by presenting the theory for the electron density before moving on to the wakefield structure. We follow the seminal work of Lu *et al.* [41, 42] with additional material from [43]. The full derivation is too involved for this dissertation, so we restrict ourselves to presenting the main results.

The nonlinear regime is reached when the beam density is sufficiently large that linear theory no longer applies. The highly nonlinear or blowout regime is reached when the center of the wake is devoid of any plasma electrons. Lu *et al.* [42] showed that trajectory crossing of plasma electrons is a necessary condition—but not sufficient—for a highly nonlinear wake to form. To understand what we mean by trajectory crossing, consider two electrons: one that starts at a small radius r_1 , and a second that starts at a larger radius r_2 ($r_2 > r_1$). As the drive beam passes, the electrons will be pushed outward before returning to the axis due to the attractive force of the ions. If at any point in time $r_1 > r_2$, then the electron trajectories will have crossed. The effect of trajectory crossing is that there is an innermost electron trajectory. All electrons that started with a smaller initial radius have crossed this innermost trajectory.¹⁵

¹⁵ Even in the blowout regime, electrons right on-axis are not deflected on large trajectories. Instead, the blowout can be thought of as a highly unstable equilibrium, particles only slightly off-axis will be rapidly deflected to a large trajectory.

The threshold for trajectory crossing is found by solving for the maximum radius of an electron r_m as a function of initial radius r_0 . Because the electrons all oscillate with approximately the same frequency, the maximum radius will occur at around the same value of ξ for all electrons. Thus, trajectory crossing occurs when $r_m(r_0)$ is no longer a monotonically increasing function. Dawson [44] worked out a simple ring model in the non-relativistic (electrostatic) limit. In the model, the plasma is treated as a series of concentric electron rings whose radius evolve according to

$$\frac{d^2 r}{d\xi^2} = -\frac{k_p r}{2} + \frac{k_p c(r_0, r, \xi)}{r}, \quad (2.70)$$

where the $k_p r/2$ term is the force from the ion column and $c(r_0, r, \xi) = \frac{1}{2}r_0^2 + \frac{1}{n_0} \int_0^r r' dr' n_b(r', \xi)$ is the force from the electrons at smaller radius, assuming no trajectory crossing. Solving this equation for $r_m(r_0)$ and finding the beam parameters that result in $\frac{dr_m}{dr_0} = 0$ gives the threshold for trajectory crossing. Fig. 2.7 shows the trajectories of electron rings driven by a uniform cylindrical (bi-flat-top) drive beam. For a uniform cylindrical beam of radius a , the cutoff for trajectory crossing is $n_b > 1.792n_0$ [42]. The exact value varies depending on the beam distribution. Trajectory crossing initially occurs for an electron that starts at $r_0 = a$, see Fig. 2.7(b). As the beam density increases, the portion of the plasma that experiences trajectory crossing expands. This portion corresponds to the domain of $r_m(r_0)$ where $\frac{dr_m}{dr_0} < 0$ with boundaries $r^* < r_0 < r_{0m}$ as shown in Fig. 2.7(c). The electron at $r_0 = 0$ is not deflected but all electrons between r^* and r_{0m} are deflected to a larger radius than the particle at r_{0m} . The minimum initial radius that experiences trajectory crossing r^* rapidly approaches 0 as the beam density is increased above the threshold. As a result, very few electrons are left in the center of the wake and a nearly perfect blowout forms.

The exact blowout threshold changes slightly depending on the details of the beam distribution, but a commonly used threshold is

$$n_b \gtrapprox 2n_0, \quad (2.71)$$

which typically agrees well with simulation.

We now know how to reach the highly nonlinear regime, but what does the wake look like?

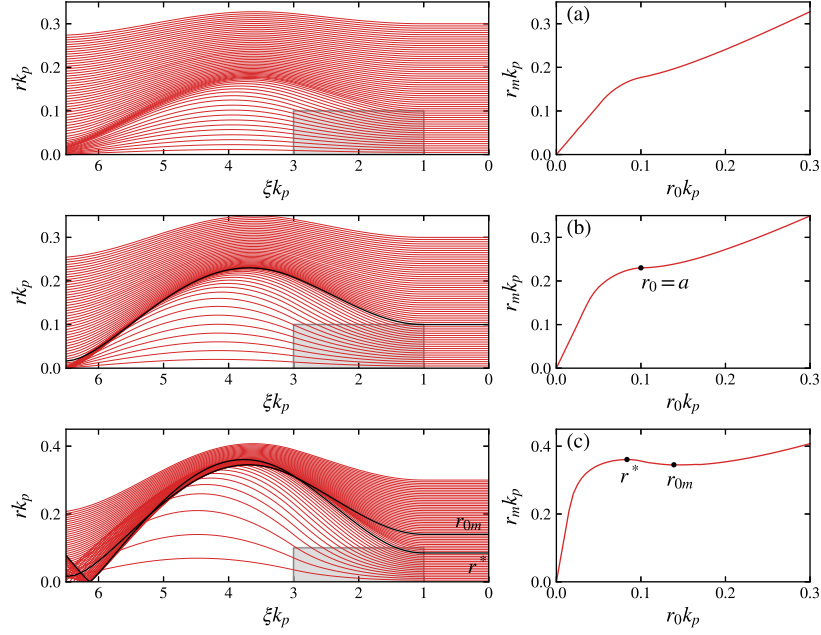


Figure 2.7: Electron trajectories for particles starting at different radial positions. The right plots show the maximum radial position of each trajectory. The shaded grey box is the electron drive beam. (a) shows trajectories for an electron beam with $n_b = n_0$, no trajectory crossing occurs. (b) is a beam with $n_b = 1.792 n_0$. This is the threshold for trajectory crossing: $\frac{dr_m}{dr_0} = 0$ at $r_0 = a$. (c) is a drive beam with $n_b = 4 n_0$. Trajectory crossing occurs for a range of initial electron positions. Adapted from Ref. [42].

The plasma density was originally investigated numerically using particle in cell (PIC) simulations—see Sec. 2.4 for details of the PIC method. It was found that the wake had three distinct regions: a central ion channel devoid of any plasma electrons, a narrow high density electron sheath around the ion column, and a region around the bubble where the plasma responds linearly. The electron density in a typical highly nonlinear wake is shown in Fig. 2.8(a). Unlike the linear case, the magnetic field is important in the nonlinear regime; thus, the structure of the plasma current in the wake is also important. Like the electron density, the plasma current is contained in a narrow sheath around the ion column. The plasma current structure is shown in Fig. 2.8(b). The sheath acts to screen the surrounding plasma from the strong fields present in the wake.

The boundary between the ion column and the sheath is extremely sharp. Compared to λ_p ,

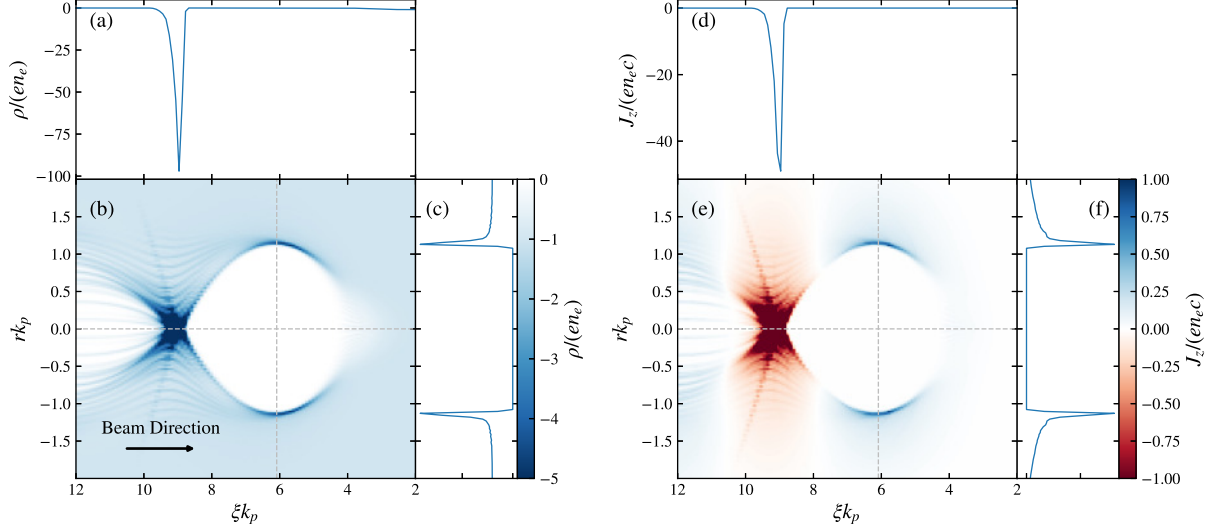


Figure 2.8: Structure of a highly nonlinear wake. (a-c) the charge density of the electrons in the wake. The center of the wake is devoid of plasma electrons which are concentrated in a narrow sheath. (d-f) the longitudinal plasma current, also concentrated in a narrow sheath.

the electron density is effectively discontinuous. As a result, the boundary can be described by a one dimensional function of the co-moving coordinate $r_b(\xi)$. When we were discussing trajectory crossing, we pointed out that there was an electron that started at r_{0m} that had the smallest maximum deflection from the axis.¹⁶ It ends up that this electron is the innermost electron along the majority of the ion column's length;¹⁷ as a result, its trajectory describes the channel boundary. The approach used by Lu *et al.* [41, 42] took advantage of this fact to find $r_b(\xi)$ by solving the single particle equation of motion for the innermost electron.

The derivation requires assuming some narrow profile for the plasma electron sheath. The only requirement on the profile is that the continuity equation is satisfied:

$$\frac{d}{d\xi} \int r dr \left(\rho - \frac{J_z}{c} \right) = 0. \quad (2.72)$$

In front of the blowout, the plasma is neutral and stationary so $\rho - J_z/c = 0$ and thus $\int r dr (\rho - J_z/c) = 0$ for all ξ . Assuming the ions are stationary, the source terms are $\rho = e(n_i - n_e)$ and

¹⁶ Excluding the small number of electrons very close to the axis.

¹⁷ We know it cannot be true at the start of the bubble because particles with $r_0 < r_{0m}$ take some distance to cross the r_{0m} trajectory.

$J_z = -en_e v_z$. The ion density is constant everywhere so the continuity equation requires that $\int r dr [n_0 + n_e(v_z/c - 1)] = 0$, which constrains the profile of the sheath. We will avoid discussing the single particle equation of motion as we have not talked about the electric and magnetic fields yet. Instead we present the result; the inner radius of the bubble is given by

$$A(r_b) \frac{d^2 r_b}{d\xi^2} + k_p B(r_b) r_b \left(\frac{dr_b}{d\xi} \right)^2 + k_p C(r_b) r_b = -\frac{k_p}{2\pi c e n_0} \frac{\lambda(\xi, r_b)}{r_b}, \quad (2.73)$$

where the source term is the beam current

$$\lambda(\xi, r_b) = 2\pi c \int_0^{r_b} \rho_b(\xi, r') r' dr'. \quad (2.74)$$

Both the drive and witness beams are included in the beam charge density $\rho_b = -en_b$. Eq. (2.73) then describes the impact of both beams on the sheath radius. The functional form of $A(r_b)$, $B(r_b)$, and $C(r_b)$ depend on the sheath profile assumed [45, 46] and the transverse density profile of the plasma [47, 48]. In a transversely uniform plasma column, the coefficients are given by

$$A(r_b) = 1 + k_p^2 \frac{r_b^2}{4} + k_p^2 \frac{r_b^2 \beta}{2} + k_p^2 \frac{r_b^3}{8} \frac{d\beta}{dr_b}, \quad (2.75)$$

$$B(r_b) = \frac{1}{2} + \frac{3\beta}{4} + \frac{3r_b}{4} \frac{d\beta}{dr_b} + \frac{r_b^2}{8} \frac{d^2 \beta}{dr_b^2}, \quad (2.76)$$

$$C(r_b) = \frac{1 + (1 + r_b^2 k_p^2 \beta/4)^2}{4(1 + r_b^2 k_p^2 \beta/4)^2}, \quad (2.77)$$

where $\beta(r_b)$ is a function that depends on the sheath profile. The results are insensitive to the exact profile used. Therefore, we will use a simple profile consisting of a uniform sheath of width Δ . The continuity equation determines the sheath density:

$$\rho - \frac{J_z}{c} = -en_0 \frac{r_b^2}{(r_b + \Delta)^2 - r_b^2}. \quad (2.78)$$

The derivation of β requires solving the Poisson-like equation for the wake potential. We will introduce the wake potential in the next section. For now we skip directly to the result

$$\beta(r_b) = \frac{(1 + \alpha)^2 \ln(1 + \alpha)^2}{(1 + \alpha)^2 - 1} - 1, \quad \alpha = \frac{\Delta}{r_b}. \quad (2.79)$$

Eq. (2.73) is still a bit cumbersome. For an accelerator, we want a strongly driven wake which means we can make the ultrarelativistic approximation $r_m \gg k_p^{-1}$ and $\Delta \ll r_b$, where r_m is the

maximum value of r_b . Further, we assume the sheath width is much smaller than the blowout radius $\Delta \ll r_b$. Expanding β in α gives $\beta \approx \alpha$ for a uniform sheath profile. For an exponential sheath profile, $\beta \approx 2\alpha$ [48]. In either case the coefficients can be simplified by dropping small quantities:

$$A(r_b) = k_p^2 \frac{r_b^2}{4}, \quad B(r_b) = \frac{1}{2}, \quad C(r_b) = \frac{1}{4}. \quad (2.80)$$

The differential equation for the sheath width then simplifies to

$$k_p r_b \frac{d^2 r_b}{d\xi^2} + 2 \left(\frac{dr_b}{d\xi} \right)^2 + 1 = -\frac{1}{2\pi c e n_0} \frac{4\lambda(\xi, r_b)}{r_b^2}. \quad (2.81)$$

Notice that r_b does not depend on Δ , and our assertion that the ion channel radius is insensitive to the details of the sheath profile is confirmed.

In finding Eq. (2.81), we assumed that $r_b \gg k_p^{-1}$, which is fine for most of the wake, but is invalid at the front and rear of the first wake period where the radius rises and falls to zero. Using an energy based argument, Golovanov *et al.* [43] showed that Eq. (2.81) is still applicable at the beginning of the bubble; however, it is difficult to solve numerically because the $\frac{d^2 r_b}{d\xi^2}$ term diverges at $r_b = 0$. To fix this, we can make the substitution $\psi_\xi(\xi) = \frac{r_b^2(\xi)}{4}$. The resulting differential equation is

$$\frac{d^2 \psi_\xi}{d\xi^2} + \left(\frac{2}{k_p} - 1 \right) \frac{1}{2\psi_\xi} \left(\frac{d\psi_\xi}{d\xi} \right)^2 = -\frac{1}{2\pi c e n_0} \frac{\lambda(\xi, 2\sqrt{\psi_\xi})}{2k_p \psi_\xi} - \frac{1}{2k_p}, \quad (2.82)$$

which can be numerically integrated with the initial conditions

$$\psi_\xi(\xi_0) = 0, \quad \left. \frac{1}{\psi_\xi} \left(\frac{d\psi_\xi}{d\xi} \right)^2 \right|_{\xi=\xi_0} = 0. \quad (2.83)$$

By definition, ψ_ξ is positive and the condition $-\frac{1}{2\pi c e n_0} \frac{\lambda}{2\psi_\xi} > \frac{1}{2}$ must be satisfied. This is not true at the head of the drive beam; the solution is to start integrating Eq. (2.82) at ξ_0 such that $\rho(\xi_0, 0) = 1/2$. Fig. 2.9 shows a comparison between the analytic solution of Eq. (2.82) and PIC simulations for increasingly relativistic wakes. At low drive beam density, the analytic predictions increasingly diverge from the simulation as ξ increases.

In the highly nonlinear blowout regime and under the ultrarelativistic limit, we now know how to calculate the radius of the ion column along the wake. Further, we know that $\rho - J_z$ for

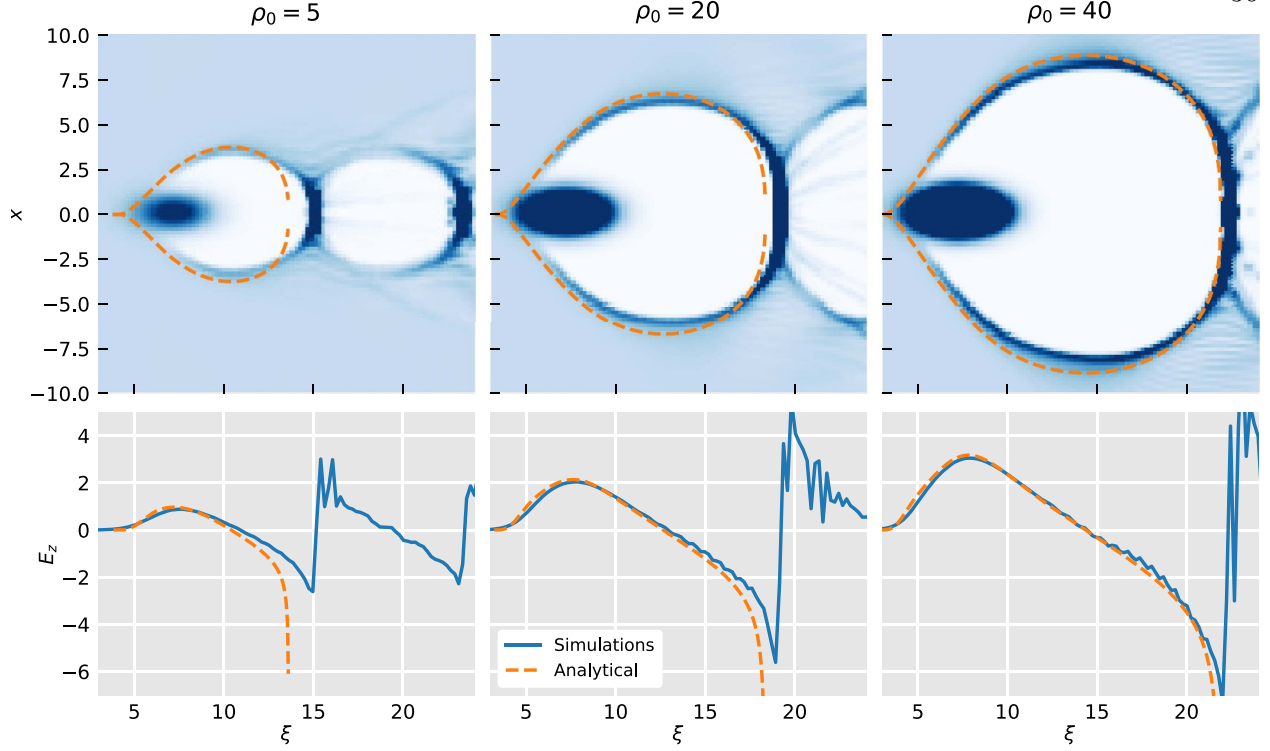


Figure 2.9: Theoretical predictions for the ion column radius in the ultrarelativistic limit compared with simulation. From right to left the beam density increases and the electrons in the wake become more relativistic. The top plots show the electron density including the beam. The bottom plots show the longitudinal electric field. Reproduced from Ref. [43].

the electrons is confined to a narrow sheath around the ion column. We do not know how to find the detailed form of the sheath or the relative contributions of the J_z and ρ . Although it may seem that we do not have enough information to find $\mathbf{E}$ and $\mathbf{B}$, we will show in the next section that knowledge of $r_b(\xi)$ is sufficient to find the wakefields in the ultrarelativistic limit.

2.2.3 Fields in the Nonlinear Regime

We were able to assume the magnetic fields in the linear regime were negligible because the kinetic energy of the plasma electrons was small compared to their rest energy. In the nonlinear regime, the fields are on the order of the wave breaking field and the character of the wake is fully electromagnetic. We start with the inhomogeneous wave equations for the potentials in the Lorentz

gauge:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) \varphi = -\frac{\rho}{\epsilon_0}, \quad (2.84)$$

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) \mathbf{A} = -\mu_0 \mathbf{J}, \quad (2.85)$$

where ϕ and $\mathbf{A}$ are the scalar and vector potentials, respectively. The Lorentz gauge is defined by requiring that the potentials satisfy the condition

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \varphi}{\partial t} = 0. \quad (2.86)$$

Transforming from (x, y, z, t) to co-moving coordinates (x, y, ξ, s) and making the quasistatic approximation causes the derivatives to transform as

$$\begin{aligned} \frac{\partial}{\partial z} &= -\frac{\partial}{\partial \xi} + \frac{\partial}{\partial s} \approx -\frac{\partial}{\partial \xi} \\ \frac{\partial}{\partial t} &= c \frac{\partial}{\partial \xi}, \end{aligned} \quad (2.87)$$

with which we can write the wave equations and Lorentz condition as

$$\nabla_{\perp}^2 \varphi = -\frac{\rho}{\epsilon_0}, \quad (2.88)$$

$$\nabla_{\perp}^2 \mathbf{A} = -\mu_0 \mathbf{J}, \quad (2.89)$$

$$\nabla_{\perp} \cdot \mathbf{A}_{\perp} = -\frac{1}{c} \frac{\partial}{\partial \xi} (\varphi - cA_z). \quad (2.90)$$

At this point it is useful to introduce the wake potential, defined as

$$\psi = \varphi - cA_z. \quad (2.91)$$

Subtracting the z component of Eq. (2.89) from Eq. (2.88) and using Eq. (2.91) to replace the potentials gives a 2D Poisson-like equation for the wake potential,

$$\nabla_{\perp}^2 (\varphi - cA_z) = \nabla_{\perp}^2 \psi = -\frac{1}{\epsilon_0} \left(\rho - \frac{J_z}{c} \right). \quad (2.92)$$

We can use ψ to find the longitudinal electric field. The electric field is given in terms of φ and $\mathbf{A}$ by

$$\mathbf{E} = -\nabla \varphi - \frac{\partial \mathbf{A}}{\partial t}. \quad (2.93)$$

Looking at only the z component and transforming to co-moving coordinates, we find

$$E_z = \frac{\partial \psi}{\partial \xi}. \quad (2.94)$$

If we know $\rho - J_z/c$, then we can solve the Poisson equation for ψ and use ψ to find the longitudinal electric field. The transverse fields are given by the normal relations:

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad (2.95)$$

$$\mathbf{E}_\perp = -\nabla_\perp \varphi - \frac{\partial \mathbf{A}_\perp}{\partial \xi}. \quad (2.96)$$

Now that we have expressions for the fields, it is a simple matter of finding the potentials from the charge and current densities.

In the previous section, we showed that $\rho - J_z/c$ for the electrons is concentrated in a narrow sheath around the constant density ion column. Within the ion column, $\rho = en_0$ and $J_z = 0$ because we have assumed the ions are stationary. We can solve Eqs. (2.88), (2.89) and (2.92) for the potentials within the bubble ($r < r_b(\xi)$). Assuming cylindrical symmetry we find

$$\varphi = \varphi_0(\xi) - \frac{r^2}{4} \frac{en_0}{\epsilon_0}, \quad (2.97)$$

$$A_z = A_{z0}(\xi), \quad (2.98)$$

$$A_r = -r \frac{1}{2c} \frac{d\psi_0}{d\xi}, \quad (2.99)$$

$$\psi = \psi_0(\xi) - \frac{r^2}{4} \frac{en_0}{\epsilon_0}, \quad (2.100)$$

where $\varphi_0(\xi) = \varphi(r=0, \xi)$, $A_{z0}(\xi) = A_z(r=0, \xi)$, and $\psi_0(\xi) = \psi(r=0, \xi)$.

We start by finding the wake potential along the axis. Far from the axis where the plasma is unperturbed, $\rho - J_z/c = 0$ and the wake potential is constant; we let $\psi(\xi \rightarrow \infty) = 0$.¹⁸

The wake potential on-axis is then found by integrating Eq. (2.92) and solving for ξ_0 . Using

$\nabla_\perp^2 \psi = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \psi}{\partial r} \right)$ we have

$$\begin{aligned} \frac{\partial \psi}{\partial r} &= -\frac{1}{r} \int_0^r r' dr' \frac{1}{\epsilon_0} \left(\rho - \frac{J_z}{c} \right), \\ \psi_0 &= \int_0^\infty \frac{dr}{r} \int_0^r r' dr' \frac{1}{\epsilon_0} \left(\rho - \frac{J_z}{c} \right). \end{aligned} \quad (2.101)$$

¹⁸ In the uniform sheath model, $\rho - J_z/c = 0$ outside of the sheath outer radius so the wake potential is 0 beyond that radius. The approach is justified even in a finite width plasma.

For the uniform sheath model from Ref. [42] (Eq. (2.78)), the integral is straightforward and the wake potential on-axis is

$$\psi_0(\xi) = \frac{en_0}{\epsilon_0} \frac{r_b^2(\xi)}{4} (1 + \beta(\xi)) \quad (2.102)$$

with $\beta(\xi)$ given by Eq. (2.79). In the ultrarelativistic approximation $\beta \approx \alpha$ and $\alpha \ll 1$. The wake potential becomes

$$\psi(\xi, r) = \frac{en_0}{\epsilon_0} \left(\frac{r_b^2(\xi)}{4} - \frac{r^2}{4} \right). \quad (2.103)$$

The longitudinal electric field is then given by Eq. (2.94):

$$E_z = \frac{en_0}{\epsilon_0} \frac{r_b}{2} \frac{dr_b}{d\xi}. \quad (2.104)$$

Knowledge of $r_b(\xi)$ is thus sufficient to solve for the longitudinal electric field. Notice that ψ_ξ is the ξ dependent part of ψ allowing us to easily find the accelerating electric field from ψ_ξ using Eq. (2.94) without ever having to calculate r_b . Examples of E_z for drive beams of increasing charge are given in the bottom of Fig. 2.9. The longitudinal field has a nearly constant slope of $\frac{en_0}{2\epsilon_0}$ in the center of the bubble and a spike in amplitude at the rear. Both the drive and witness beams can load the wake much like in the linear case.

It is trivial to show that $B_z = 0$ using Eq. (2.95). All that is left is to solve for the transverse fields. In particular, we are interested in the transverse force on an electron in the drive or witness beam. From the Lorentz force law, we see that the transverse force is

$$\mathbf{W}_\perp = \mathbf{E}_\perp + \mathbf{v} \times \mathbf{B}_\perp = \mathbf{E}_\perp + c\hat{z} \times \mathbf{B}_\perp, \quad (2.105)$$

where $\mathbf{W}_\perp$ is the focusing field defined such that the force on an electron is $\mathbf{F}_\perp = -e\mathbf{W}_\perp$. Evaluating Eq. (2.96) with φ from Eq. (2.97) and A_r from Eq. (2.99) gives the transverse electric field

$$\mathbf{E}_\perp = \frac{1}{2}r \left(\frac{en_0}{\epsilon_0} + \frac{1}{c} \frac{d\psi_0}{d\xi} \right) \hat{r}. \quad (2.106)$$

The $\frac{1}{2}r \frac{en_0}{\epsilon_0}$ term is the focusing field from the ion column space charge and the $\frac{1}{2c}r \frac{d\psi_0}{d\xi}$ term is from the radial plasma current. Similarly, combining Eqs. (2.95) and (2.99) gives the transverse magnetic field

$$\mathbf{B}_\perp = \frac{1}{2}r \frac{1}{c} \frac{d\psi_0}{d\xi} \hat{\phi}, \quad (2.107)$$

which is from the radial plasma current. The radial focusing field is $W_r = E_r - cB_\phi$, we see that the electric and magnetic fields from the radial plasma current cancel out, and the focusing field is equal to the electric field from the ion column space charge

$$W_r = E_{rion} = \frac{1}{2} \frac{en_0}{\epsilon_0} r = \frac{1}{2} \frac{m_e \omega_p^2}{e} r. \quad (2.108)$$

Importantly, the focusing force is linear in r and does not vary longitudinally within the bubble.

The wakefields are shown in Fig. 2.10 for a typical blowout wake. The highly nonlinear PWFA accelerator has several advantages over the linear PWFA. First, the longitudinal field is uniform in r , Fig. 2.10(g), so all particles at a given ξ are accelerated at the same rate. Second, the focusing force is linear in r , Fig. 2.10(h), and therefore, the beam electrons undergo simple harmonic motion in the transverse direction, which, we will show later, is important for preserving the beam quality. Third, the large E_z spike at the rear of the wake, Fig. 2.10(d), allows the transformer ratio to exceed 2 even for symmetric drive beams.

As in the linear case, beam loading occurs. The current profiles of the drive and witness beam modify the bubble radius according to Eq. (2.82), which then changes the longitudinal field following Eq. (2.104). It turns out that optimal loading of the wake in the blowout regime requires trapezoidal beams just like in the linear regime. In the nonlinear regime, the constants a and b have been worked out by Tzoufras *et al.* [49] and are given for the witness beam by

$$a = -\frac{E_{zs} k_p^2 \epsilon_0}{e}, \quad (2.109)$$

$$b = -\frac{E_{zs} k_p^2 \epsilon_0}{e} \left[\xi_s + \sqrt{\left(\frac{e E_{zs}}{m_e c \omega_p} \right)^4 + \left(\frac{\omega_p R_b}{2c} \right)^4} \right], \quad (2.110)$$

where $E_{zs} = E_z(\xi_s)$ is the longitudinal electric field at the start of the witness beam, ξ_s is the location of the start of the witness beam, and R_b is the maximum radius of the ion column.

The tradeoff between efficiency and transformer ratio is much less restrictive in the nonlinear regime. It is possible to achieve relatively high efficiency while maintaining a transformer ratio close to 1. As in the linear regime, the charge of the witness beam varies inversely to the strength of the accelerating field. The energy flux within the blowout at the maximum radius is given by

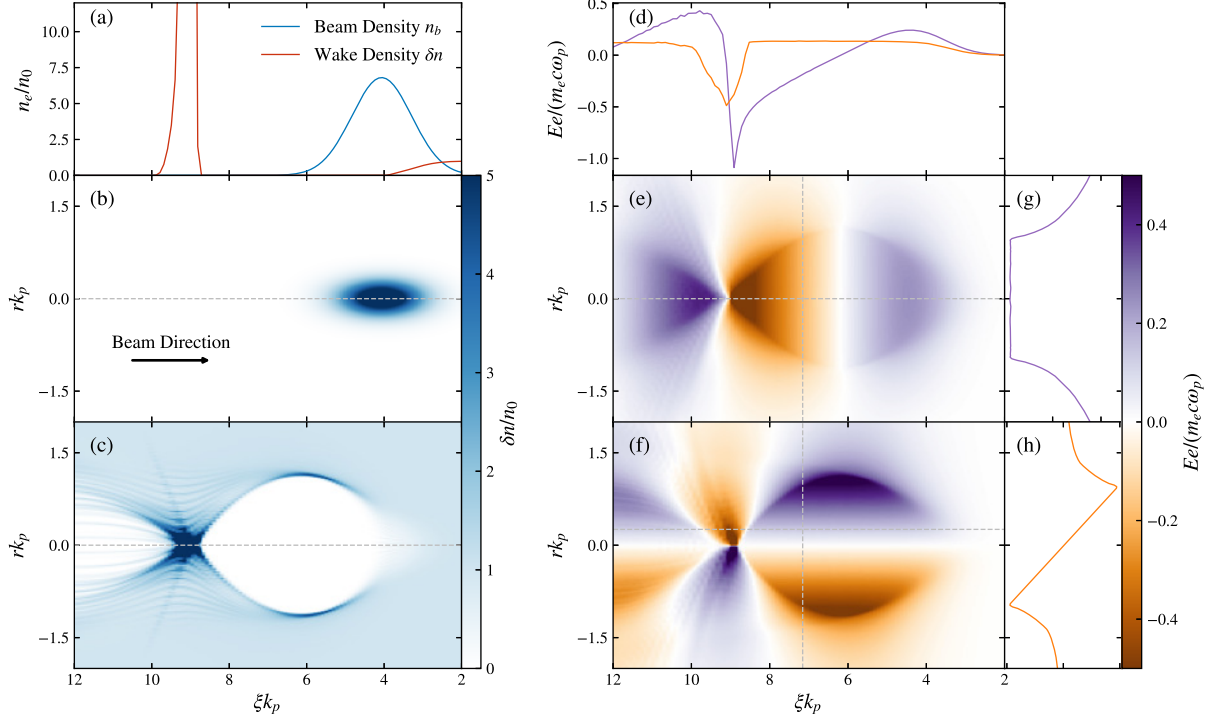


Figure 2.10: Beam and plasma density of a highly nonlinear plasma wave and the electric fields generated by the wave. (a) shows the on-axis density of the drive beam and the wake. (b) is the density of the Gaussian drive beam. (c) is the density of the excited plasma wave. (d) shows longitudinal lineouts of the fields. In purple is the on-axis longitudinal field and in orange is the off axis radial field—the on-axis radial field is 0. (e) and (g) show the longitudinal field and the transverse variation in the field while (f) and (h) show the same thing for the radial field.

[49, 43]

$$I_0 = \frac{m_e^2 c^5}{64 e^2} (k_p R_b)^4. \quad (2.111)$$

In the rest of the blowout the energy flux can be found using

$$\frac{dI_0}{d\xi} = -\frac{n_e e}{8\pi \epsilon_0^2} r_b \lambda(\xi, r_b) \frac{dr_b}{d\xi}. \quad (2.112)$$

The derivative term means any current in the decelerating phase of the wake will add energy to the wake, while any current in the accelerating phase of the wake will remove energy from the wake.

We finish our discussion of plasma waves by looking at the scaling laws for the highly nonlinear blowout regime. Recall that, in the linear regime we found that the strength of the wake scaled with both the current and the charge. Ref. [43] worked out the scaling laws for I_0 in a wake driven

by a cylindrical beam¹⁹ in the blowout regime. The energy flux available for the witness beam scales as

$$I_0 = 2\pi c^2 \sqrt{\frac{2m_e}{e}} \frac{Q^{3/2}}{\Delta\xi^{1/2}\lambda_p}, \quad (2.113)$$

where $\Delta\xi$ is the length of the beam and Q is the total charge. It is assumed that $\Delta\xi \ll \lambda_p$ and $r_B < \Delta\xi\sqrt{\rho_0}$, where r_B is the beam radius and ρ_0 is the beam density. As in the linear case, we see that increasing the charge and reducing the bunch length (increasing the peak current) maximizes the energy transfer rate into the wake, thus reducing the length of the accelerator. The blowout is approximately spherical with diameter of λ_p . From Eq. (2.111), we see that the wake sizes scales very weakly as the energy in the wake is increased; thus, even a strong wake has a size very close to λ_p .

2.3 Accelerator Physics

In the previous sections, we calculated the plasma waves and resulting wakefields created by a pair of electron beams in a PWFA. The beams themselves, however, were treated as static distributions by invoking the quasistatic approximation. Although the evolution is slow compared to the timescales of the plasma, it can be significant over the length of the PWFA. In this section, we introduce the concepts of accelerator physics, the study of the dynamics of a particle beam within an accelerator.

The PWFA is a single pass device, unlike a circular accelerator or storage ring, the particles only interact with the accelerator once. During this interaction, the particles in the drive beam lose energy, the particles in the witness beam gain energy, while both beams are focused by the transverse forces. We limit ourselves to the ultrarelativistic regime where inter-beam particle interactions are negligible (see Sec. 2.3.5 for details). This allows the evolution of the beam as a whole to be built up from the motion of the individual particles in external fields, in our case the wakefields. The beam evolution can of course feedback onto the wake and change the fields, but because the focusing force in a highly nonlinear wake depends only on the plasma density, the

¹⁹ For other beam distributions, the scaling laws will stay the same but the coefficients will vary.

feedback is typically not important.²⁰

We provide a cursory summary of beam dynamics in the simplified form relevant to a single stage PWFA, for more details see Ref. [50].

2.3.1 Transverse Motion

The coordinate system used in accelerator physics is defined in terms of a reference trajectory (called a reference orbit in a circular accelerator). The particles in both the drive and witness beams undergo small amplitude transverse motion around the reference trajectory. This transverse motion is called *betatron motion* after the type of accelerator it was originally studied in [50]. The reference trajectory in a linear accelerator is defined as the trajectory of a particle with zero betatron amplitude. The circular definition of the reference trajectory can be understood in a PWFA in terms of the focusing force within the wake. Recall that the focusing force varies in r and is zero on-axis for a cylindrically symmetric blowout wake. A particle placed on-axis will not undergo any oscillations; thus, the reference trajectory is the center of the wake. Because the wake itself follows the drive beam, the reference trajectory is defined by the drive beam trajectory.

Let $\mathbf{r}_0(s)$ be the reference trajectory where s is the path length along the trajectory. The coordinate system around the trajectory is called the Frenet-Serret coordinate system with $\hat{x}$ defined as perpendicular to the trajectory and in the tangential plane; $\hat{y}$ is also perpendicular to the trajectory while being perpendicular to the tangential plane. The final unit vector $\hat{s}$ is tangent to the trajectory. The unit vectors are given by

$$\hat{s}(s) = \frac{d\mathbf{r}_0(s)}{ds}, \quad \hat{x}(s) = \rho(s) \frac{d\hat{s}(s)}{ds}, \quad \hat{y} = \hat{s}(s) \times \hat{x}(s), \quad (2.114)$$

where $\rho(s) = \left| \frac{ds}{d\hat{s}} \right|$ is the radius of curvature of the trajectory. Typically $\mathbf{r}_0(s)$ is confined to the x - z plane, in which case, $\hat{y}$ is in the positive vertical direction. In a PWFA the reference trajectory is straight and oriented along the z axis: $\mathbf{r}_0(s) = s\hat{z}$. The unit vectors follow the usual Cartesian unit vectors with $\hat{s}$ replacing $\hat{z}$.

²⁰ The exception is hosing instability, which will be discussed in Sec. 3.5.1. The longitudinal field is also affected by the head erosion of the drive beam, but this is not particularly important from a beam dynamics perspective.

A particle's path in Cartesian coordinates is given by $\mathbf{x}(t) = x(t)\hat{x} + y(t)\hat{y} + z(t)\hat{z}$; however, we would like to use s as the independent variable instead of t . We introduce the reference particle with momentum $p_0(s)$ that travels along the reference trajectory. The reference particle has velocity in $\hat{s}$ with magnitude $v_0 = \sqrt{\frac{c^2 p_0^2}{m^2 c^2 + p_0^2}}$, and the co-moving coordinate can be defined in terms of the reference particle as $\xi = v_0 t - z$. A particle's path in the new coordinates is $\mathbf{x}(s) = x(s)\hat{x} + y(s)\hat{y} + \xi(s)\hat{s}$. In a linear accelerator, the evolution of each particle is an initial value problem that depends on the particle's initial position $\mathbf{x}_0$ and momentum $\mathbf{p}_0$. In our new coordinate system, a particle is then fully defined by its six dimensional phase space coordinates $(x, p_x, y, p_y, \xi, p_s)$. It is slightly more convenient to work with trace space coordinates in the transverse direction, defined as the angles $x' = \frac{dx}{ds} = \frac{p_x}{p_s}$ and $y' = \frac{dy}{ds} = \frac{p_y}{p_s}$, where prime denotes the derivative with respect to s . In the longitudinal direction, it is more convenient to use the fractional momentum deviation from the reference particle $\delta = (p - p_0)/p_0$. Trace space coordinates are $(x, x', y, y', \xi, \delta)$.

We want the transverse equation of motion in trace space. The force from the electromagnetic fields on a single particle in the beam is given by the Lorentz force law

$$\frac{d\mathbf{p}}{dt} = \mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad (2.115)$$

where the momentum is $\mathbf{p} = \gamma_b m \mathbf{v}$ for a relativistic particle. We start with the x component. Making the ultrarelativistic approximation for the reference particle $v_0 \approx c$ allows us to write the equation of motion as

$$\frac{dp_x}{ds} = \frac{q}{c} E_x + q \left(\frac{v_y}{c} B_z - \frac{v_z}{c} B_y \right). \quad (2.116)$$

Using the definition of relativistic momentum $p_x = \gamma m v_x$ gives

$$\begin{aligned} \gamma_b m \frac{dv_x}{ds} + m v_x \frac{d\gamma_b}{ds} &= \frac{q}{c} E_x + q \left(\frac{v_y}{c} B_z - \frac{v_z}{c} B_y \right), \\ \frac{dv_x}{ds} + \frac{\gamma'_b}{\gamma_b} v_x &= \frac{q}{\gamma_b m c} E_x + \frac{q}{\gamma_b m} \left(\frac{v_y}{c} B_z - \frac{v_z}{c} B_y \right). \end{aligned} \quad (2.117)$$

If we assume the deviations from the reference orbit are small $x' \ll 1$, and the particles are ultrarelativistic, then $x' \approx \frac{v_x}{c}$ and $\frac{v_z}{c} \approx 1$. We can divide by c and simplify to

$$x'' + \frac{\gamma'_b}{\gamma_b} x' = \frac{q}{\gamma_b m c^2} E_x + \frac{q}{\gamma_b m c} (y' B_z - B_y). \quad (2.118)$$

Similarly, in the y direction we have

$$y'' + \frac{\gamma'_b}{\gamma_b} y' = \frac{q}{\gamma_b m c^2} E_y - \frac{q}{\gamma_b m c} (x' B_z - B_x). \quad (2.119)$$

We can integrate these equations to find the transverse motion of a particle if we know $\gamma_b(s)$ of the particle and the fields it experiences: $\mathbf{E}_\perp$ and $\mathbf{B}$.

The fields contain contributions from the mean space charge field of the other particles in the beam and the external fields from the accelerator. We will show in Sec. 2.3.5 that the mean space charge fields are negligible for an ultrarelativistic beam; this leaves us with only the external fields from the accelerator. Using Eq. (2.95), we showed that $B_z = 0$ in a highly nonlinear PWFA; this removes the x - y coupling from the equations of motion. Introducing the focusing field from Eq. (2.105), we have

$$x'' + \frac{\gamma'_b}{\gamma_b} x' = \frac{q}{\gamma_b m c^2} W_x. \quad (2.120)$$

The y direction has the same form. Assuming the beam is made of electrons, we can use Eq. (2.108) for W_x

$$x'' + \frac{\gamma'_b}{\gamma_b} x' + \frac{\omega_p^2}{2\gamma_b c^2} x = 0. \quad (2.121)$$

The equation for the y direction has the same form. The coefficient on the x term is called $K(s)$. For our nonlinear PWFA, $K(s) = \frac{\omega_p(s)^2}{2\gamma_b c^2}$, for other accelerator elements it takes on different values.

In the special case of constant energy, we get Hill's equation

$$x'' + K(s)x = 0. \quad (2.122)$$

Strictly, this is only Hill's equation if $K(s)$ is periodic, otherwise it is a harmonic oscillator with time dependent frequency. Eq. (2.122) is a linear, homogeneous, second order differential equation.²¹

The solution can always be written as a linear function of the initial conditions

$$\mathbf{x}(s) = \begin{pmatrix} x(s) \\ x'(s) \end{pmatrix} = \mathbf{M}(s|s_0) \mathbf{x}(s_0) = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} x(s_0) \\ x'(s_0) \end{pmatrix}. \quad (2.123)$$

²¹ Eq. (2.121) is also a linear, homogeneous second order differential equation so the solution can be described by transfer matrices. The determinant of the transfer matrices, however, is no longer 1 due to the dampening term.

The matrix $\mathbf{M}$ is called the transport matrix, importantly it only depends on $K(s)$ and not on the particle's initial conditions. The matrix from s_0 to s_2 can be constructed by matrix multiplication if the matrices from s_0 to s_1 and s_1 to s_2 are known:

$$\mathbf{M}(s_2|s_0) = \mathbf{M}(s_2|s_1)\mathbf{M}(s_1|s_0) \quad (2.124)$$

Hill's equation is applicable if the particle energy is constant, and the force on the particle depends linearly on x . Two common accelerator elements have these features: drift spaces (M_d) and quadrupoles (M_q). For both, the transport matrix can be found analytically [50]:

$$\mathbf{M}_d = \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}, \quad \mathbf{M}_q = \begin{pmatrix} \cos(\sqrt{K}l) & \frac{1}{\sqrt{K}} \sin(\sqrt{K}l) \\ -\sqrt{K} \sin(\sqrt{K}l) & \cos(\sqrt{K}l) \end{pmatrix}. \quad (2.125)$$

Here, l is the length of the element. A drift space is a region without any electric or magnetic fields. The transport matrix is intuitive: particles travel in straight lines so their direction of propagation x' does not change; the straight lines have angle x'_0 and intercept of x_0 so the particle position must be $x = x_0 + x'_0 l$. The beam size within a conventional accelerator is controlled by electromagnetic quadrupoles. A quadrupole is a focusing magnet where B_x varies linearly in y and B_y varies linearly in x . Because the magnetic field is curl and divergence free within the accelerator, a quadrupole can only focus in one direction (positive K). In the other transverse direction, it is defocusing (negative K). Typically, quadrupoles are relatively short, $\sqrt{K}l \ll 1$, and the transport matrix can be expanded in the thin lens approximation as

$$\mathbf{M}_{q,focusing} = \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix}, \quad \mathbf{M}_{q,defocusing} = \begin{pmatrix} 1 & 0 \\ 1/f & 1 \end{pmatrix}, \quad (2.126)$$

where $f = \lim_{l \rightarrow 0} \frac{1}{\sqrt{K}l}$ is the focal length of the quadrupole. A quadrupole acts like a lens; a particle's position is unchanged as it passes through, but its angle is changed proportional to how far it is from the axis. Fig. 2.11 shows how particles behave as they pass through a drift space and a thin lens quadrupole.

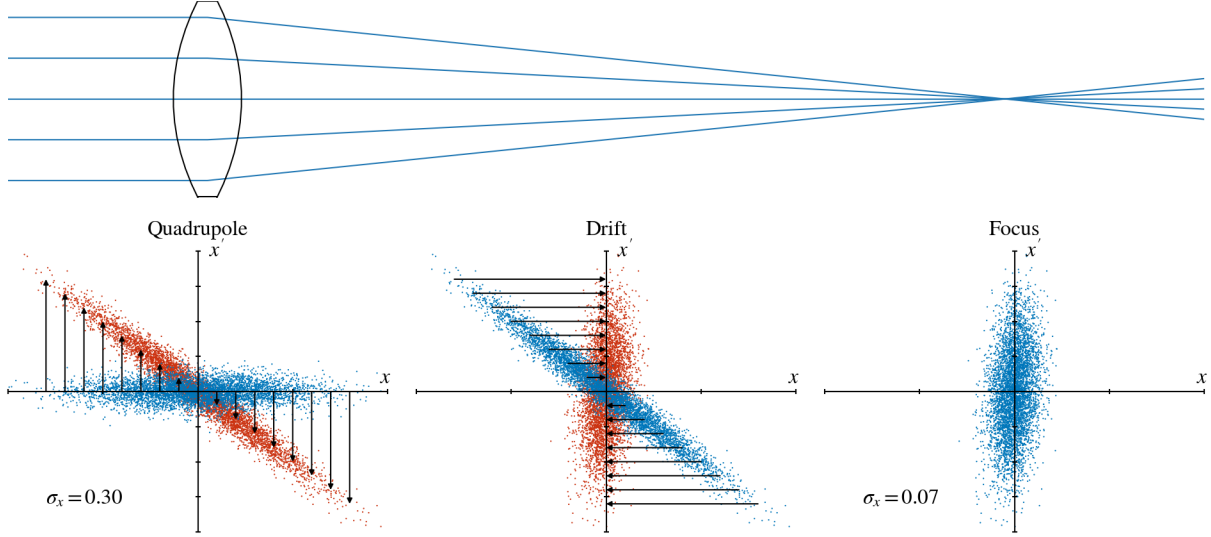


Figure 2.11: Evolution of an electron passing through a thin lens quadrupole and then a drift space. The top plot shows the trajectory of particles through the system. The bottom plots show the trajectory in phase space. In the quadrupole, the particle position is unchanged but its angle changes proportional to its position. In a drift space, the opposite happens; a particle's angle is unchanged but its position changes proportional to its angle.

2.3.2 Courant-Snyder Parameters

The general solution to Hill's equation for an arbitrary, periodic focusing force is given in terms of the Courant-Snyder (CS) parameters. Even though K is not periodic within a PWFA stage, CS parameters are still used, and it is important to understand how they are defined. We follow the work of Courant and Snyder [51] who first defined the CS parameters, with additional information from Lee [50]. We switch the order of the derivation, providing the more intuitive definition first, before introducing the original definition.

A solution to Hill's equation—even for non-periodic K —is a harmonic oscillator with s dependent amplitude and phase

$$x_1(s) = w(s)e^{i\Phi(s)}, \quad (2.127)$$

where $w(s)$ and $\Phi(s)$ are arbitrary amplitude and phase functions. Substituting back into Hill's equation, Eq. (2.122), we find that $x_1(s)$ is a solution if $w(s)$ and $\Phi(s)$ satisfy the differential

equations

$$w'' + K(s)w - \frac{1}{w^3} = 0, \quad \Phi' = \frac{1}{w^2}. \quad (2.128)$$

In finding the phase equation, we have chosen the integration constant to be 0. The solution $x_2(s) = w(s)e^{-i\Phi(s)}$ also satisfies Eq. (2.122) and is linearly independent from x_1 . We define the classic CS parameters α , β , and γ as

$$\beta = w^2, \quad (2.129)$$

$$\alpha = -\frac{\beta'}{2}, \quad (2.130)$$

$$\gamma = \frac{1 + \alpha^2}{\beta}. \quad (2.131)$$

Substituting β into Eq. (2.128) gives a differential equation for β :

$$\frac{1}{2}\beta'' + K\beta - \frac{1}{\beta} \left[1 + \left(\frac{1}{2}\beta' \right)^2 \right] = 0. \quad (2.132)$$

The general solution to Hill's equation is given by $x = ax_1 + bx_2$. A single particle starts with initial conditions $x_0 = x(s_0)$ and $x'_0 = x'(s_0)$, which can be used to determine the coefficients a and b once $x_{1,2}$ are known. The CS parameters, however, are ambiguous; we are free to choose boundary conditions for Eq. (2.132) and the resulting β still results in a valid solutions to Hill's equation. If $K(s)$ is assumed periodic, then β must satisfy periodic boundary conditions $\beta(s) = \beta(s + L)$ and $\beta'(s) = \beta'(s + L)$ with the same period, L , as K . The periodic boundary is sufficient to uniquely define β .

The definition of γ can be understood by considering Courant and Snyder's original definitions for α , β , and γ . We start by writing out the transport matrix from our solution to Hill's equation

$$\mathbf{M}(s_2|s_1) = \begin{pmatrix} \sqrt{\frac{\beta_2}{\beta_1}}(\cos \Phi + \alpha_1 \sin \Phi) & \sqrt{\beta_1 \beta_2} \sin \Phi \\ -\frac{1+\alpha_1 \alpha_2}{\sqrt{\beta_1 \beta_2}} \sin \Phi + \frac{\alpha_1 - \alpha_2}{\sqrt{\beta_1 \beta_2}} \cos \Phi & \sqrt{\frac{\beta_1}{\beta_2}}(\cos \Phi - \alpha_2 \sin \Phi) \end{pmatrix}. \quad (2.133)$$

Enforcing the periodic boundary condition simplifies the matrix when written over a single period

$$\mathbf{M}(s + L|s) = \begin{pmatrix} \cos \Phi + \alpha \sin \Phi & \beta \sin \Phi \\ -\gamma \sin \Phi & \cos \Phi - \alpha \sin \Phi \end{pmatrix} = \mathbf{I} \cos \Phi + \mathbf{J} \sin \Phi, \quad (2.134)$$

where $\mathbf{I}$ is the identity matrix and

$$\mathbf{J} = \begin{pmatrix} \alpha & \beta \\ -\gamma & -\alpha \end{pmatrix}. \quad (2.135)$$

Periodic boundary or not, the transport matrix must have a determinant of one because Hill's equation is linear without any first derivative terms. Eq. (2.134) is a general form for any matrix M with determinate of one when $\mathbf{J}$ has zero trace and unit determinant. Requiring the determinant to be one introduces the condition

$$\beta\gamma - \alpha^2 = 1 \quad (2.136)$$

which is the origin of Eq. (2.131). For any transport matrix $\mathbf{M}$ from s_0 to s_1 , $\mathbf{M}$ can be written in the form of Eq. (2.134) defining $\beta(s_0, s_1)$ regardless of periodicity. $\beta(s_0, s_1)$ is only related to the envelope function $\beta(s)$ when periodic boundary conditions are enforced.

The motion of a single particle is given by a linear combination of x_1 and x_2 , rewriting in terms of an initial phase and amplitude gives

$$x(s) = \sqrt{\epsilon\beta(s)} \cos(\Phi(s) + \Phi_0), \quad (2.137)$$

where $\Phi(s)$, called the *betatron phase advance*, is given by integrating Eq. (2.128):

$$\Phi(s) = \int_0^s \frac{ds'}{\beta(s')}. \quad (2.138)$$

The initial amplitude ϵ is called the *single particle emittance*. It is easy to verify by substitution that

$$C(x, x') = \gamma x^2 + 2\alpha x x' + \beta x'^2 = \epsilon \quad (2.139)$$

is an invariant of the motion called the Courant-Snyder invariant. The invariant traces out an ellipse in phase space, shown in Fig. 2.12. At any point in the accelerator, a particle will exist at a position on the ellipse determined by its betatron phase. The emittance of the particle defines the magnitude of the deviation from the reference orbit.

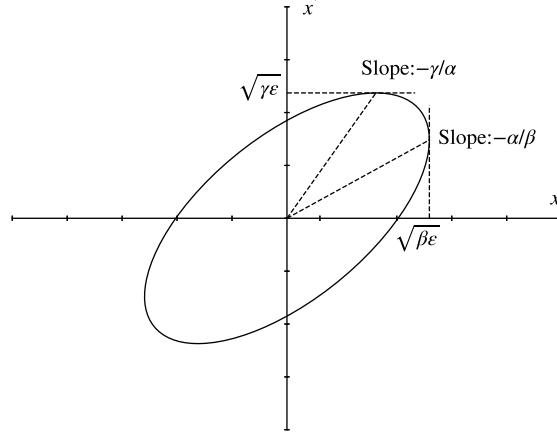


Figure 2.12: Phase space ellipse of the Courant-Snyder invariant. The single particle emittance is the area of the ellipse and is an invariant of the motion. The shape of the ellipse is set by the CS parameters, which are functions of the focusing force in the machine. A particle's phase space coordinates lie at a position on the ellipse defined by the particle's betatron phase. Adapted from [50, 52].

The CS parameters at two different points in the accelerator are related by [50]

$$\begin{pmatrix} \beta_2 \\ \alpha_2 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} M_{11}^2 & -2M_{11}M_{12} & M_{12}^2 \\ -M_{11}M_{21} & M_{11}M_{22} + M_{12}M_{21} & -M_{12}M_{22} \\ M_{21}^2 & -2M_{21}M_{22} & M_{22}^2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \alpha_1 \\ \gamma_1 \end{pmatrix}, \quad (2.140)$$

where M_{ij} represents an element of the transport matrix $\mathbf{M}(s_2|s_1)$. This matrix is an easy way to find the evolution of the CS parameters through linear accelerator elements such as drifts and quadrupoles.

The CS parameters are defined in terms of the transport matrix, which only depends on the focusing force $K(s)$; if all particles see the same focusing force, then the CS parameters are defined solely by $K(s)$, i.e. they are a property of the accelerator. Note that α and γ are fully defined by β ; as such, it is common to describe the transverse focusing design of an accelerator by specifying $\beta(s)$ through the lattice.²² As discussed earlier, the beta function is only well defined for an accelerator made of periodic cells where the periodic boundary condition can be imposed. A

²² From Eq. (2.118), we can see that the focusing force depends on the particle energy. $\beta(s)$ is typically specified at the energy of the reference particle.

circular accelerator is naturally periodic, and many linear accelerators are composed of identical, repeating cells, allowing both to be described by the beta function. In a single stage PWFA, we need to introduce an initial condition to solve for β , we will cover this in the next section.

2.3.3 The Beam Distribution and Emittance

Up to now we have only discussed the motion of a single particle. Typically, we are interested in the properties of a particle beam, a collection of particles each with its own position and velocity. With sufficient particles, the beam can be treated as a distribution in 6D phase space $g(x, p_x, y, p_y, \xi, \delta)$ (we will also use the trace space distribution $f(x, x', y, y', \xi, \delta)$). The full 6D phase space is difficult to visualize and to measure experimentally. Instead, 2D projections and slices are used. Typically, the x and y plane behave independently and we can limit ourselves to the $f(x, p_x, \xi, \delta)$ subspace, the 4D y subspace will behave in a similar way.

We start with the transverse direction. The projection onto the x - x' plane, Fig. 2.13(a), is given by

$$f_{\perp}(x, x') = \int d\xi d\delta f(x, x', \xi, \delta). \quad (2.141)$$

A longitudinal slice can also be projected onto the x - x' plane, Fig. 2.13(b),

$$f_{\xi}(x, x'; \xi) = \frac{1}{N(\xi)} \int d\xi' d\delta f(x, x', \xi, \delta) \delta_D(\xi' - \xi), \quad (2.142)$$

and similarly for an energy slice, Fig. 2.13(c),

$$f_{\delta}(x, x'; \delta) = \frac{1}{N(\delta)} \int d\xi d\delta' f(x, x', \xi, \delta) \delta_D(\delta' - \delta), \quad (2.143)$$

where $N(\xi) = \int dx dx' d\delta f(x, x', \xi, \delta)$, $N(\delta) = \int dx dx' d\xi f(x, x', \xi, \delta)$, and δ_D is the Dirac delta function. In all three cases, we end up with a distribution in x and x' , we can perform similar integrals in momentum space to get a distribution in x and p_x . The distribution is characterized

by its moments:

$$\begin{aligned}
\langle x \rangle &= \int dx dx' x f_{\perp}(x, x'), & \langle x' \rangle &= \int dx dx' x' f_{\perp}(x, x'), \\
\langle x^2 \rangle &= \int dx dx' x^2 f_{\perp}(x, x'), & \langle x'^2 \rangle &= \int dx dx' x'^2 f_{\perp}(x, x'), \\
\sigma_x^2 &= \langle x^2 \rangle - \langle x \rangle^2, & \sigma_{x'}^2 &= \langle x'^2 \rangle - \langle x' \rangle^2, & \sigma_{xx'} &= \langle xx' \rangle - \langle x \rangle \langle x' \rangle,
\end{aligned} \tag{2.144}$$

where σ_x and $\sigma_{x'}$ are the RMS spot size and RMS divergence, respectively.

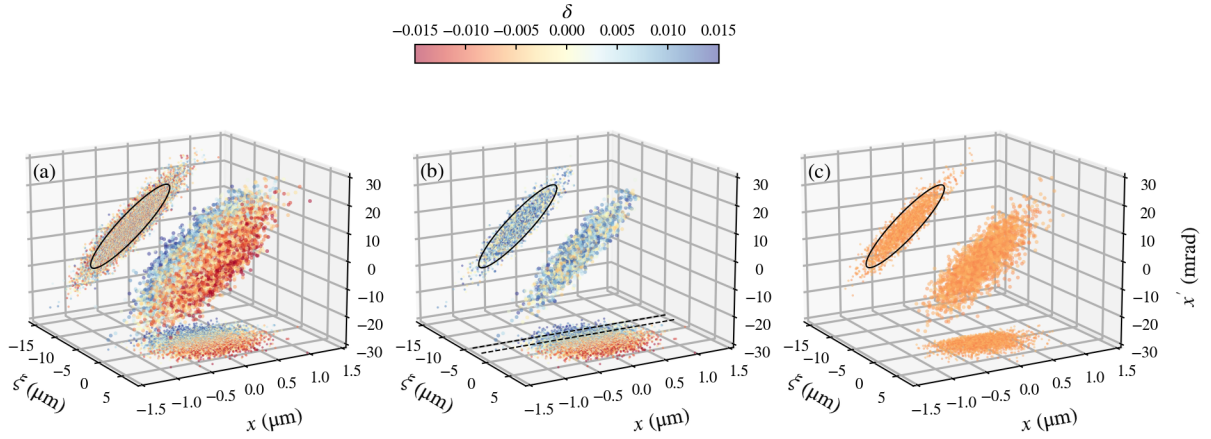


Figure 2.13: 4D phase space distribution of an electron beam. The projected x - x' distribution is shown in (a). The x - x' distribution of a longitudinal slice is shown in (b) and of an energy slice is shown in (c).

For a single particle with constant energy, the single particle emittance is an invariant of the motion. The shape of the ellipse is determined solely by β , which depends on the focusing force of the accelerator, while the size of the ellipse is determined by the emittance. Consider a beam where all the particles have the same, constant, energy (equivalently, a single energy slice of a beam); the evolution of every particle is described by the same transport matrix [31, 50]. We can take advantage of this by writing a matrix equation for the evolution of the entire beam,

$$\mathbf{X}(s) = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_N \end{pmatrix} = \mathbf{M}(s|s_0) \mathbf{X}(s_0), \tag{2.145}$$

where $\mathbf{X}$ is a $4 \times N$ matrix with each column being the transverse coordinates of a single particle.

The beam moments are given by taking the outer product of the matrix, for simplicity, we assume

$\langle x \rangle = \langle x' \rangle = 0$:

$$\Sigma = \begin{pmatrix} \langle x^2 \rangle & \langle xx' \rangle \\ \langle xx' \rangle & \langle x'^2 \rangle \end{pmatrix} = \frac{\mathbf{X} \mathbf{X}^T}{N} \quad (2.146)$$

As the particles propagate through the accelerator, the evolution of the sigma matrix is given by

$$\Sigma(s) = \frac{\mathbf{M} \mathbf{X}(s_0) [\mathbf{M} \mathbf{X}(s_0)]^T}{N} = \frac{\mathbf{M} \mathbf{X}(s_0) \mathbf{X}^T(s_0) \mathbf{M}^T}{N} = \mathbf{M} \Sigma(s_0) \mathbf{M}^T. \quad (2.147)$$

Because the determinant of the transport matrix is 1, the determinant of Σ remains constant as the beam propagates. We define the *geometric emittance* as

$$\epsilon = \sqrt{\det(\Sigma)} = \sqrt{\langle x^2 \rangle \langle x'^2 \rangle - \langle xx' \rangle^2}, \quad (2.148)$$

which is a constant of the motion for the entire beam. As for the single particle emittance, the geometric emittance is a measure of the deviation of the beam from the reference orbit.

In defining the geometric emittance as a constant of the motion, we have assumed all the particles have the same energy. For a beam that undergoes nonlinear focusing and has some energy spread, it is useful to define a statistical definition of the CS parameters. In the case of a monoenergetic beam in a linear lattice, all particles have the same CS parameters, and the statistical definition should reduce to the single particle values. We follow the derivation of Buon [53]. The beam distribution can always be rotated to a coordinate system w and w' such that $\sigma_{w'}$ is minimized. Rotating the coordinate system by θ we have the condition

$$\frac{\partial}{\partial \theta} \sigma_{w'}^2 = 0. \quad (2.149)$$

The position of a particle in the new coordinate system is

$$w' = x' \cos \theta - x \sin \theta. \quad (2.150)$$

Combining Eq. (2.149) and Eq. (2.150) and solving for θ gives

$$\tan 2\theta = \frac{2 \langle xx' \rangle}{\langle x^2 \rangle - \langle x'^2 \rangle}. \quad (2.151)$$

The two moments are given by

$$\sigma_{w,w'}^2 = \frac{1}{2} \left(\langle x^2 \rangle + \langle x'^2 \rangle \pm \frac{2 \langle xx' \rangle}{\sin 2\theta} \right). \quad (2.152)$$

In the new coordinates, the geometric emittance is given simply as $\epsilon = \sigma_w \sigma_{w'}$. We can define the envelope ellipse in the new coordinate system as

$$\frac{w^2}{\sigma_w^2} + \frac{w'^2}{\sigma_{w'}^2} = 1. \quad (2.153)$$

Transforming back to x and x' , the ellipse equation becomes

$$x^2 \langle x'^2 \rangle - 2xx' \langle xx' \rangle + x'^2 \langle x^2 \rangle = \epsilon. \quad (2.154)$$

Letting the CS parameters define the envelope ellipse the same way they define the CS invariant ellipse (see Eq. (2.139)) we get

$$\beta = \frac{\langle x^2 \rangle}{\epsilon}, \quad \gamma = \frac{\langle x'^2 \rangle}{\epsilon}, \quad \alpha = -\frac{\langle xx' \rangle}{\epsilon}. \quad (2.155)$$

We can see from Eq. (2.155) that the CS parameters are related to Σ according to

$$\Sigma = \epsilon \begin{pmatrix} \beta & -\alpha \\ -\alpha & \gamma \end{pmatrix}. \quad (2.156)$$

The statistical definition of the CS parameters describes the properties of the beam distribution. The single particle CS parameters describe the evolution through a linear lattice. In a linear lattice, these two definitions are equivalent (for the proof see Ref. [53]).

Remember that the evolution of β is given by Eq. (2.132), but the boundary conditions are ill defined for a non-periodic accelerator. If the properties of the incoming beam are known, we can use Eq. (2.155) to define initial conditions for β and α , giving us sufficient boundary conditions to fully define β throughout the accelerator.

The importance of the emittance can be understood by considering a focusing system. A final focus composed of quadrupoles will focus the beam down to a minimum β , denoted β^* , that depends on the strength and arrangement of the quadrupoles. The beam size at the focus is $\sigma_x^2 = \beta\epsilon$. Producing a tightly focused beam requires minimizing the emittance as much as possible, as β^* is limited by the strength of the magnets.

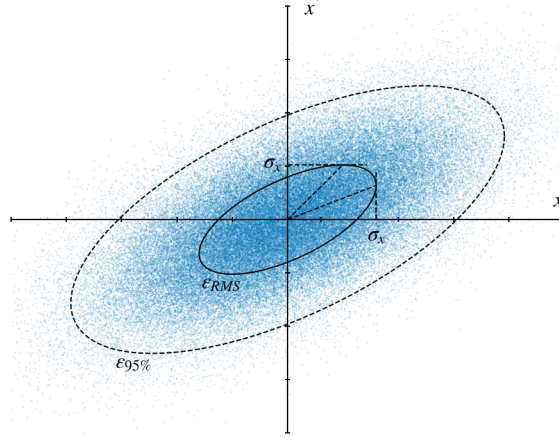


Figure 2.14: Phase space ellipse and geometric emittance for a beam of electrons. The RMS geometric emittance is a statistical measure of the beam. It can be used to define the CS parameters statistically. The CS ellipse for the beam is defined in terms of these parameters. If the particles are ordered by single particle emittance, then $\epsilon_{95\%}$ is the emittance that contains 95% of the particles.

Eq. (2.132) only describes the evolution of β if all the particles have the same energy and the energy remains constant.²³ In a plasma-based accelerator, the energy gain is significant, and the original definition of the CS parameters is not very useful. There are two approaches to handle both energy gain and energy spread in the beam. The first is to define the CS parameters using Eq. (2.155), which is always valid but removes most of the utility of introducing the parameters in the first place. A second approach is to define CS parameters for each energy slice, again using Eq. (2.155). This results in β , α , and γ being functions of δ and is useful for modelling the propagation between accelerator stages.

As we have seen, the emittance is preserved for a mono-energetic beam propagating through linear elements. Equivalently, the emittance of each energy slice in a beam with energy spread is preserved. Since the emittance is a measure of the area the beam occupies in transverse phase space, the total volume of the beam in phase space is preserved. This is a statement of Liouville's

²³ In a circular accelerator, the beta function at a given energy is fully defined by the focusing force and the periodic boundary conditions of the accelerator through Eq. (2.132). If the relative energy gain per revolution is small, the beta function evolves adiabatically through the solutions for different energies, and the formalism presented so far is sufficient to define the CS parameters.

Theorem, which formally says that the phase space density is constant around any trajectory in phase space, as long as the system is described by Hamiltonian mechanics. Ignoring space charge forces, the particles evolve under the influence of external electromagnetic fields according to the Hamiltonian

$$H(t) = c\sqrt{m^2c^2 + (\mathbf{p} - q\mathbf{A})^2} + q\phi. \quad (2.157)$$

In the absence of accelerating fields, we can write down a Hamiltonian for the transverse motion of each particle in terms of the trace space coordinates [50, 54]. All the particles see the same Hamiltonian so the initial beam distribution $f(x, x', \xi, \delta)$ is a distribution in the phase space of that Hamiltonian.²⁴ The number of particles within a volume of phase space is $f(x, x', \xi, \delta)dx dx' d\xi d\delta$. According to Liouville's Theorem, if we follow a particle trajectory, $x(s), x'(s)$, the distribution function remains constant $f(x(s), x'(s), \xi, \delta) = f(x, x', \xi, \delta)$. The number of particles within a phase space volume about the particle trajectory thus remains constant. The result is a phase space volume enclosed by a surface of constant phase space density. The projection of this volume onto the x - x' plane is the geometric emittance.

The particles in the beam evolve according to the Hamiltonian of Eq. (2.157) in the presence of accelerating fields. The canonical coordinates are then x , y , and z and the momenta p_x , p_y , and p_z . Importantly, x' and y' are no longer canonical momentum. Consequently, the geometric emittance is no longer the conserved phase space quantity. This can be understood intuitively by recalling that $x' = p_x/p_z$. When a particle is accelerated, p_z increases while p_x remains the same. Consequently x' shrinks for all particles and the geometric emittance decreases, this effect is called *adiabatic dampening*. The conserved quantity is the volume of the beam in 6D phase space surrounded by a 5D surface of constant phase space density.²⁵ The projection of this volume onto

²⁴ Adding in interactions between the beam particles means each particle experiences a slightly different Hamiltonian. As a result, the initial beam distribution no longer describes the phase space distribution of a single Hamiltonian and Liouville's theorem cannot be applied. For more details see [54]

²⁵ The emittance is defined as the phase space volume enclosed by a surface of constant phase space density. The definition is ambiguous in so far as the choice of surface, effectively a choice of scaling parameter. One approach is to take the surface that envelopes some percentage of the beam particles. The most common choice is the 95% emittance. However, only scaling the distribution by the RMS values, i.e. using Eq. (2.155), will result in CS parameters that can be propagated using Eq. (2.140). Unless otherwise specified, the RMS emittance is used throughout this thesis. For an in depth discussion, see Ref. [52]

the x - p_x plane is called the *normalized emittance*, defined as

$$\epsilon_n = \frac{1}{m_e c} \sqrt{\sigma_x^2 \sigma_{p_x}^2 - \sigma_{xp_x}^2}, \quad (2.158)$$

which is the same form as the expression for the geometric emittance with x' replaced by p_x . The use of the normalized emittance allows comparison between the quality of the beam before and after acceleration in the plasma stage. The geometric emittance is not suitable because it will always decrease as a result of acceleration. The normalized emittance is the most important figure of merit for the transverse quality of the beam produced by an accelerator.

The normalized and geometric emittance are projections of the 6D phase space volume onto the transverse coordinates. As such, they are not constants of the motion and can increase or decrease depending on the correlations in the beam. That being said, the normalized emittance can typically only grow as the beam traverses an accelerator.²⁶ The growth happens because the transverse dimension is often initially uncorrelated with the longitudinal dimension. Further, the normalized emittance for a slice of the beam with fixed energy and longitudinal position is preserved by linear elements. Each slice then has the same constant emittance, but the projection of the all the slices depends on how well they overlap. In this case, the normalized emittance can only grow from the initial value and cannot be reduced below the initial value. An example is shown in Fig. 2.15, where the normalized emittance of each energy slice is the same, but the total projected emittance is larger. Further, if the correlations between the longitudinal and transverse dimensions are complex, the beam cannot even be corrected back to the initial value. As we will show in this dissertation, plasma-based accelerators can introduce complex transverse-longitudinal correlations in the beam that cannot be easily corrected. As such, it is important to preserve the projected emittance within the plasma stage.

To end this section, we want to comment on one last effect that can cause the “effective” full 6D emittance to grow without violating Liouville’s theorem. Liouville’s theorem concerns the

²⁶ There are several techniques to reduce the beam emittance, called dampening. One is using the radiation reaction from synchrotron radiation to create a dissipative system. The particles are circulated in a dampening ring to produce sufficient synchrotron radiation to reduce the emittance.

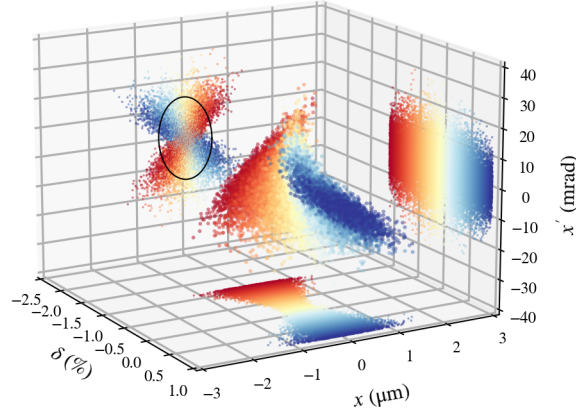


Figure 2.15: Growth of the normalized emittance. The projected emittance, the area of the ellipse on the x - x' plane, can typically only increase from its initial value. If the beam starts without any longitudinal correlations, then the longitudinal slice emittance is preserved. As the beam propagates it can develop transverse longitudinal correlations, in this case a twist along the beam, causing the projected emittance to grow. The projected emittance cannot be reduced below the individual slice emittance.

values of the continuous phase space density function. Particularly in the presence of nonlinear focusing terms, the density function can develop into thin filaments in phase space that technically satisfy Liouville's theorem. However, as the size of the filaments reduces down to the scale of the discrete particle spacing, the continuous approximation breaks down and the effective phase space volume of the discrete beam is much larger than its initial volume. In many practical cases, such as particle colliders and synchrotron radiation light sources, the length scales of interest are even larger than the inter-particle spacing. In these cases, the effective emittance grows when the filaments are smaller than the coarse length scales important for such beams.

2.3.4 Gaussian Beam Distribution

The beam distribution can take on any form, but the Gaussian distribution is commonly found in accelerators and typically used for theoretical work. In this section, we present the functional

form of the distribution and derive some useful relations. The beam density is given by

$$f(x, x', \xi) = \frac{1}{(2\pi)^{3/2} \epsilon \sigma_\xi} \exp \left[-\frac{1}{2\epsilon_{RMS}^2} (\sigma_{x'}^2 x^2 - 2\sigma_{xx'} x x' + \sigma_x^2 x'^2) \right] \exp \left(-\frac{\xi^2}{2\sigma_\xi^2} \right). \quad (2.159)$$

The distribution can be transformed into action-angle coordinates (J, ψ) :²⁷

$$x = \sqrt{2\beta J} \cos \psi, \quad x' = -\sqrt{\frac{2J}{\beta}} \sin \psi. \quad (2.160)$$

The Gaussian distribution in action angle coordinates is

$$f(J, \psi, \xi) = \frac{1}{(2\pi)^{3/2} \epsilon \sigma_\xi} \exp \left(-\frac{J}{\epsilon_{RMS}} \right) \text{rect} \left(\frac{\psi}{2\pi} \right) \exp \left(-\frac{\xi^2}{2\sigma_\xi^2} \right). \quad (2.161)$$

Generating the beam in action-angle coordinates is convenient for numerical work, particularly particle tracking.

The Gaussian distribution can be expanded to include the y direction by adding in an additional transverse exponential term. The charge density is given by the spacial distribution

$$\rho(x, y, \xi) = \frac{Q}{(2\pi)^{3/2} \sigma_x \sigma_y \sigma_\xi} \exp \left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} - \frac{\xi^2}{2\sigma_\xi^2} \right), \quad (2.162)$$

where Q is the total beam charge. The beam number density is $n_b = -\rho(x, y, \xi)/e$. The current profile of an ultrarelativistic beam is

$$\lambda(\xi) = c \int dx dy \rho(x, y, \xi) = \frac{cQ}{(2\pi)^{1/2} \sigma_\xi} \exp \left(-\frac{\xi^2}{2\sigma_\xi^2} \right) \quad (2.163)$$

2.3.5 Beam Self Fields

An electron beam generates its own electric and magnetic fields that act back on the beam. The fields are given by the inhomogeneous wave equations, Eqs. (2.25) and (2.26), with the beam charge and current as the source:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{E} = -\frac{e}{\epsilon_0} \nabla n_b - \frac{e}{\epsilon_0 c^2} c \hat{z} \frac{\partial n_b}{\partial t}, \quad (2.164)$$

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{B} = \mu_0 e c \nabla \times (n_b \hat{z}). \quad (2.165)$$

²⁷ For the derivation of action angle variables see Appendix B.

Following the same procedure as the derivation of linear plasma waves, we take the ultrarelativistic limit and transform into the co-moving coordinate:

$$\nabla_{\perp}^2 \mathbf{E} = -\frac{e}{\epsilon_0} \nabla_{\perp} n_b \quad (2.166)$$

$$\nabla_{\perp}^2 \mathbf{B} = \mu_0 e c \left(\frac{\partial n_b}{\partial y} \hat{x} - \frac{\partial n_b}{\partial x} \hat{y} \right) \quad (2.167)$$

If we now take the cross product of $\frac{1}{c} \mathbf{v}$ with Eq. (2.166), we get

$$\nabla_{\perp}^2 \left(\frac{1}{c^2} \mathbf{v} \times \mathbf{E} \right) = \mu_0 e c \left(\frac{\partial n_b}{\partial y} \hat{x} - \frac{\partial n_b}{\partial x} \hat{y} \right), \quad (2.168)$$

and thus

$$\mathbf{B} = \frac{1}{c^2} \mathbf{v} \times \mathbf{E}, \quad (2.169)$$

which is a general result for any cold beam²⁸ (see Jackson Sec. 11.10 [32]). Now that we know how to find $\mathbf{B}$ from $\mathbf{E}$, we just need to solve Eq. (2.166). Immediately we can see that $E_z = 0$. If the beam is cylindrically symmetric, we can then use Gauss law to find the electric field

$$E_r(r, \xi) = -\frac{e}{\epsilon_0} \frac{1}{r} \int_0^r r' dr' n_b(r', \xi). \quad (2.170)$$

We should be slightly careful about the ultrarelativistic approximation in this case [55]. The E_z field is only zero if $v = c$, it is straightforward to find the field for beams with finite energy. Again assuming cylindrical symmetry, we can write Faraday's law as

$$\begin{aligned} \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \left(\frac{\partial E_r}{\partial z} - \frac{\partial E_z}{\partial r} \right) &= -\frac{\partial B_{\phi}}{\partial t}. \end{aligned} \quad (2.171)$$

Using Eq. (2.169) and transforming to the co-moving coordinate with velocity v_b , we get

$$\left(\frac{\partial E_r}{\partial \xi} + \frac{\partial E_z}{\partial r} \right) = -v_b \frac{\partial B_{\phi}}{\partial \xi} = -\frac{v_b^2}{c^2} \frac{\partial E_r}{\partial \xi}. \quad (2.172)$$

Using the definition of γ_b , we get the relation

$$\frac{\partial E_z}{\partial r} = -\frac{1}{\gamma_b^2} \frac{\partial E_r}{\partial \xi}. \quad (2.173)$$

²⁸ A cold beam has no momentum spread; all particles have $x' = y' = \delta = 0$. Because x' and y' are very small for a ultrarelativistic beam, the cold beam assumption is valid for this case.

For a beam with a Gaussian profile the fields can be calculated analytically. Inserting Eq. (2.162) into Eq. (2.170), we find

$$E_r(r, \xi) = \frac{Q}{(2\pi)^{3/2} \sigma_\xi \epsilon_0} \exp\left(-\frac{\xi^2}{2\sigma_\xi^2}\right) \frac{1}{r} \left[1 - \exp\left(-\frac{r^2}{2\sigma_r^2}\right)\right]. \quad (2.174)$$

The longitudinal field is found using Eq. (2.173). The electric fields from a Gaussian beam are shown in Fig. 2.16.

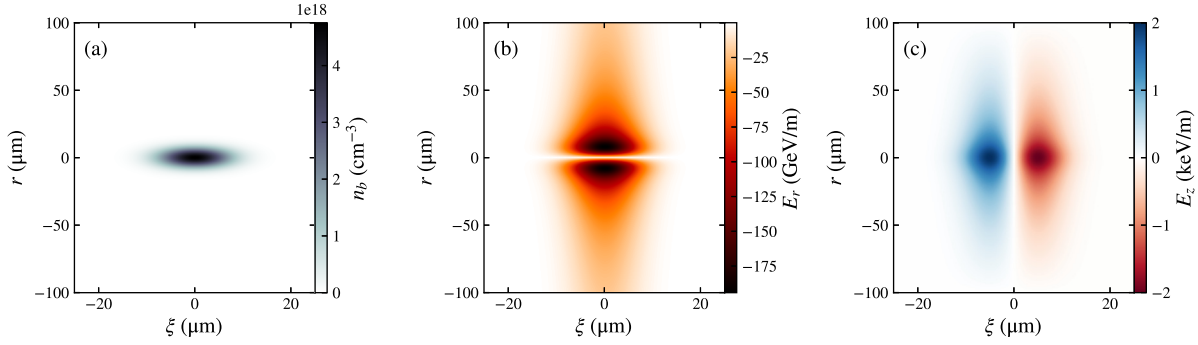


Figure 2.16: Electric fields from a Gaussian electron beam. (a) the electron number density. (b) the radial electric field. (c) the longitudinal electric field. The longitudinal electric field is found by integrating Eq. (2.173) from the simulation boundaries, the actual field will be larger.

The force on a particle in the beam from the self fields is given by the Lorentz force law.

Replacing the magnetic field using Eq. (2.169) we have

$$\mathbf{W}_\perp = \frac{E_\perp}{\gamma_b^2} \quad (2.175)$$

because the focusing force from the magnetic field cancels out the defocusing force from the electric field up to order $1/\gamma_b^2$. Ultrarelativistic beams propagate ballistically in free space; as a result, we can ignore the self fields when calculating the beam dynamics.

2.3.6 Longitudinal Effects

The evolution of the beam particles in a plasma wakefield accelerator is not limited to the transverse direction. The particles evolve in longitudinal phase space (ξ - δ space) as they are accelerated. The most important feature of the dynamics is the ξ dependence of the accelerating

field present in both plasma-based accelerators and RF accelerating cavities. The equation of motion for a single particle is found by starting with the z component of the Lorentz force law (Eq. (2.115)):

$$\frac{dp_s}{ds} = \frac{q}{c}E_z + q\left(\frac{v_x}{c}B_y - \frac{v_y}{c}B_x\right). \quad (2.176)$$

Inserting the definition of p_s , making the ultrarelativistic approximation, and assuming x' and y' are small gives

$$\gamma'_b = \frac{q}{mc^2}E_z + \frac{q}{mc}(x'B_y - y'B_x). \quad (2.177)$$

Our longitudinal momentum coordinate is $\delta = (p - p_0)/p_0$, but we have a differential equation for the particle energy. In the ultrarelativistic limit $p \approx \gamma_b mc$, allowing us to rewrite δ in terms of γ_b : $\delta = (\gamma_b - \bar{\gamma}_b)/\bar{\gamma}_b$, where the bar represents the reference particle. δ is therefore not only the relative momentum spread, but also the relative energy spread. The reference particle has transverse velocities $x' = y' = 0$ by definition, so the momentum and energy of the reference particle is given by

$$p'_0 = mc\gamma'_b = \frac{q}{c}E_z, \quad \bar{\gamma}_b = \frac{q}{mc^2}E_z. \quad (2.178)$$

Changes in the δ distribution primarily result from the ξ dependence of E_z . The magnetic field terms $x'B_y$ and $y'B_x$ are typically negligible and can be ignored.²⁹

The longitudinal evolution is coupled to the transverse evolution through the $(\gamma'_b/\gamma_b)x'$ term in Eq. (2.118). The form of Eq. (2.121) makes it clear that this term is a time dependent dampening that depends on the rate of energy gain. This dampening process is the adiabatic dampening discussed earlier and is shown schematically in Fig. 2.17. A force along $\hat{s}$ increases p_s ; as a result, the particle's angle, given by $x' = p_x/p_s$, is reduced. A second source of transverse-longitudinal coupling enters through the γ_b in the denominator of $K(s)$. We will find that this second source of coupling is important for understanding the emittance evolution within a PWFA.

Later on we will discuss the FACET-II accelerator. Here, we introduce the 6D transport

²⁹ From Eqs. (2.106) and (2.107), we see that the magnetic field is a factor of c smaller than the electric field (typical for electromagnetic waves). The relative strength of the magnetic force to the electric force is then given by $x' \ll 1$, and the magnetic force can be ignored.

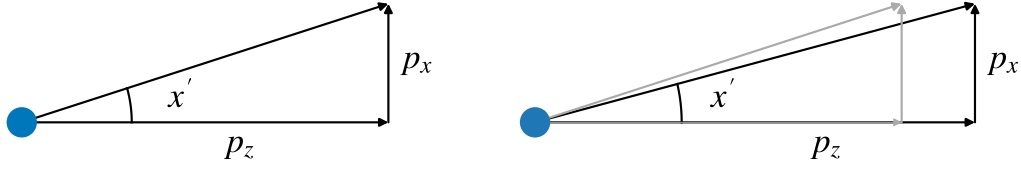


Figure 2.17: Adiabatic dampening. As a particle is accelerated, p_z increases while p_x stays the same. As a result, the trajectory angle x' decreases.

matrix required to understand the discussion. The evolution of a single particle through a section of beamline is typically dominated by linear effects with small higher order corrections. The evolution of a particle can then be expanded as

$$x_2^i = R_{ij}x_1^j + T_{ijk}x_1^jx_1^k + U_{ijkl}x_1^jx_1^kx_1^l + \dots, \quad (2.179)$$

where x^i is now the full 6D vector of phase space coordinates, R_{ij} is the linear transport matrix, and T_{ijk} and U_{ijkl} are the second and third order transport matrices. If only the linear elements are important, then the first order transport matrix behaves the same as the transport matrices for Hill's equation. We can write

$$\mathbf{x}(s) = \begin{pmatrix} x(s) \\ x'(s) \\ y(s) \\ y'(s) \\ \xi(s) \\ \delta(s) \end{pmatrix} = \mathbf{R}(s|s_0)\mathbf{x}(s_0) = \begin{pmatrix} R_{11} & R_{12} & R_{13} & R_{14} & R_{15} & R_{16} \\ R_{21} & R_{22} & R_{23} & R_{24} & R_{25} & R_{26} \\ R_{31} & R_{32} & R_{33} & R_{34} & R_{35} & R_{36} \\ R_{41} & R_{42} & R_{43} & R_{44} & R_{45} & R_{46} \\ R_{51} & R_{52} & R_{53} & R_{54} & R_{55} & R_{56} \\ R_{61} & R_{62} & R_{63} & R_{64} & R_{65} & R_{66} \end{pmatrix} \begin{pmatrix} x(s_0) \\ x'(s_0) \\ y(s_0) \\ y'(s_0) \\ \xi(s_0) \\ \delta(s_0) \end{pmatrix}. \quad (2.180)$$

As before, the total transport matrix for multiple elements can be created by multiplying the individual matrices together per Eq. (2.124). It is worth noting that, not only is the linear approach invalid within a plasma-based accelerator, the series expansion in Eq. (2.179) does not converge because $x(s)$ depends sinusoidally on $\delta(s_0)$ well beyond the radius of convergence of the Taylor series. However, the linear approach is still useful within the conventional part of the accelerator.

2.4 Particle in Cell Simulation

The full electromagnetic interaction between the beam and plasma can be very complex and difficult to tackle theoretically. It is common practice to turn to particle in cell (PIC) simulations to help understand the behavior of the accelerator and design or interpret experiments. Particle in cell codes treat the beam and plasma kinetically as a collection of macro particles each representing many real particles. Macroparticles are initialized such that they sample the full 6D phase-space distribution of the particles.³⁰ The motion of the macro particles is solved for in a continuous space while the fields from the macroparticles are solved for on a grid using a finite difference time domain solution to Maxwell's equations. The simulation loop proceeds as follows:

- (1) Deposit the charges and currents from the macro particles into a charge and current density on the grid.
- (2) Calculate the fields on the grid from the charge and current density using a numerical solution to Maxwell's equations.
- (3) Interpolate the fields from the grid to the individual macroparticle positions.
- (4) Push the particles to their new positions/momenta by integrating the particle's equations of motion.

Typically, the simulation of a plasma-based accelerator is performed on a co-moving grid that travels at constant speed along with the particle beam. The grid then only needs to be large enough to contain the beam and the wake, rather than the entire length of the accelerator.

PIC simulations are computationally difficult and time consuming. A full simulations of a single plasma stage typically takes days even when running on thousands of CPU cores. For example, the plasma source for FACET-II has a density of $3.5 \times 10^{16} \text{ cm}^{-3}$ which results in a

³⁰ Typically, the macroparticles are initialized on a grid in the position degrees of freedom with variable weights representing the phase space density. Each macro particle is then assigned momentum randomly from a distribution that can depend on the particle's position. The result is a uniform sampling of the initial position distribution of the beam and a random sampling of the initial momentum distribution and can support correlations between the two.

plasma wavelength of $\lambda_p \approx 175 \mu\text{m}$. From Sec. 2.2.3, we know that the size of the blowout is approximately equal to λ_p . The simulation box needs to be approximately twice as large as the blowout. For our example, this is a cubic simulation box $350 \mu\text{m}$ to a side. The cell size is set by the scale of the electron beam. For FACET-II, this is on the order of $\sigma_{RMS} = 4 \mu\text{m}$. The beam size needs to be resolved in order to capture the dynamics of the wake formation; assuming 5 cells per σ gives a cell size of $0.8 \mu\text{m}$. The grid is then about 440 cells in each direction for a total of 85 million cells. For stability and accuracy reasons,³¹ the time step is tied to the cell size by the speed of light. For a 0.5 m long plasma source this is 625,000 time steps. Using the code VSim [56] with 1600 CPU cores on NERSC's Cori supercomputer, a time step takes about 1 second with 8 beam and 8 plasma macro particles per grid cell. With these parameters, simulating a full accelerator stage takes almost a week. It is not possible to optimize an accelerator using such lengthy simulations. Analytic approaches provide a way to reduce the parameter space before turning to simulations.

³¹ For the electromagnetic solver to be stable, the time step needs to satisfy the CFL condition which sets the upper limit. The dispersion relation for EM waves is most accurate if the time step is as close to the CFL limit as possible.

Chapter 3

Evolution of the Witness Beam

Plasma-based accelerators cannot take the place of conventional accelerators unless they can produce beams of comparable quality. In this chapter, we consider how the properties of the accelerated witness beam evolve as the beam traverses a plasma stage. The goal of the chapter is to determine what requirements the plasma source must satisfy in order to produce high quality beams. We focus on the transverse evolution of the witness beam and the ways in which this evolution depends on the density profile of the plasma source.

The witness beam can be fully characterized by its 6D phase space distribution. The evolution of this distribution is complex; multiple physical processes compete in driving the beam's dynamics. Of primary importance are chromatic effects. The beam always starts with some initial energy spread, and the energy distribution will evolve during acceleration if the wake is not perfectly loaded. Because the particles have different energies, they have different betatron oscillation frequencies. If we slice the phase space distribution in energy, we will find each slice evolves differently. To complicate matters, the change in the beam's energy distribution during acceleration will cause the betatron frequency to vary in time. As a result, particles will mix between energy slices leading to complex evolution and correlations in 6D phase space.

The full 6D phase space is typically too difficult to work with analytically. In any case, it is rarely important for accelerator applications. Instead, we focus on the projections of the full phase space onto the various 2D planes as well as projections of phase space slices. Of primary importance to applications is the transverse beam moments and emittance, both of the projected

beam, and longitudinal or energy slices.

The projected emittance will grow due to chromatic effects as it traverses the plasma if the divergence of the beam is not matched to the focusing force of the plasma. The beam sizes required to match to the plasma are very small, typically beyond the focusing capability of conventional magnetic optics. Placing a ramped density profile at the entrance of the plasma source focuses the beam, reducing the beam size. If the ramp profile is chosen correctly for a given incoming beam, it will focus the beam down to the required size, and the emittance growth will be minimized. We calculate the beam evolution in a plasma ramp in order to specify requirements on the ramp length and shape.

We start with a description of the transverse evolution in a uniform plasma, without acceleration, to introduce the mechanism of chromatic phase spreading. We then extend this to a longitudinally non-uniform plasma to show how the profile of the plasma can impact the final properties of the beam. We demonstrate that a wide range of ramp shapes can be used to preserve the emittance of the beam, provided the ramp length is chosen carefully. Next, we introduce acceleration into the picture and present an approach, using FACET-II as an example, to find the bulk plasma density and minimum plasma width. Following this, we combine the longitudinal and transverse dynamics to develop a general theory for the beam evolution. We show that longitudinal field results in growth of the slice emittances. We demonstrate that minimizing the total plasma density in the ramps minimizes the negative impacts on the beam. Finally, with FACET-II as our target application, we put this all together to show how to design a plasma profile that preserves the beam emittance while maximizing the energy gain and minimizing the energy spread.

Sec. 3.2 as well as Secs. 3.6.1 and 3.6.2 are largely a direct reproduction of Ariniello *et al.* [57] © 2019 American Physical Society. Sections have been broken up and their order changed to make a more coherent narrative in the thesis. Sec. 3.4 is composed of direct reproductions of several sections from Ariniello *et al.* [58] © 2022 American Physical Society.

3.1 Transverse Dynamics in a Uniform Plasma

The electrons in the witness beam experience strong focusing fields within the wake. The important features of the transverse beam evolution can be understood without taking into account the longitudinal fields. In this case, the particle energy is constant, and the beam evolves in the transverse direction according to Hill's equation

$$\frac{d^2x}{ds^2} + K(s)x = 0. \quad (3.1)$$

The focusing strength of the plasma is constant in s but depends on the particle energy (see Sec. 2.3.1) according to

$$K = \frac{\omega_p^2}{2\gamma_b c^2} = \frac{\omega_p^2}{2\bar{\gamma}_b c^2} \frac{1}{1 + \delta}, \quad (3.2)$$

where a bar denotes the value of the reference particle. Typically, we write the focusing force in terms of the betatron wavenumber $k_\beta^2 = K$.

Each particle undergoes simple harmonic motion with wavenumber k_β in the transverse direction, Fig. 3.1(a). The oscillation wavenumber for each particle depends on the particle's energy because $k_\beta = \bar{k}_\beta / \sqrt{1 + \delta}$. For a given energy slice, the particles will oscillate coherently and each individual particle will trace an ellipse in transverse phase space, Fig. 3.1(b). The slice, as a whole, will thus rotate in phase space. The different slices, however, will rotate at different rates.

3.1.1 Evolution of the Beta Function

For each energy slice, we can use the CS parameter formalism to describe the beam evolution. The initial values of β and α are given by Eq. (2.155) for each slice. It is then a simple matter of solving Eq. (2.132) using these initial conditions to get the evolution of the CS parameters. We use the solution given by Eq. (2.140). The transport matrix for constant K is simply the thick quadrupole transport matrix of Eq. (2.125). Combining, we get

$$\beta = \frac{1}{2} \left(\beta_0 + \frac{\gamma_0}{k_\beta^2} \right) + \frac{1}{2} \left(\beta_0 - \frac{\gamma_0}{k_\beta^2} \right) \cos(2k_\beta s) - \frac{\alpha_0}{k_\beta} \sin(2k_\beta s), \quad (3.3)$$

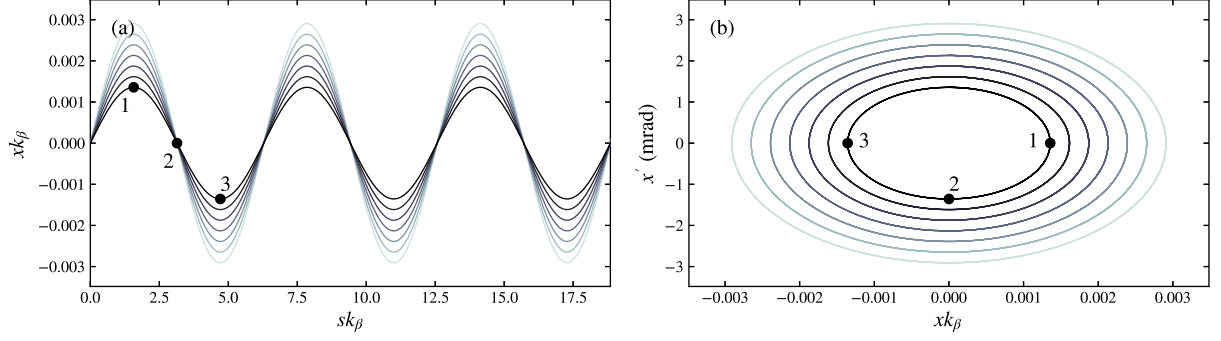


Figure 3.1: Motion of a single particle in a uniform ion column. (a) shows the particle trajectory along the accelerator, (b) shows the trajectory in transverse phase space. The curves, from dark to light, correspond to particles with increasing single particle emittance. The particles in each curve have the same energy.

where β_0 , α_0 and γ_0 are the CS parameters at $s = 0$. Remember that β is proportional to the RMS size of the beam, so the beam will undergo envelope oscillations, shown in Fig. 3.2, unless $\beta_0 = \beta_m = \frac{1}{k_\beta}$ and $\alpha_0 = 0$. A beam that satisfies this condition is called a *matched beam* and propagates without any changes in transverse RMS size. The matched condition represents the special case where the ellipse of the CS invariant is concentric with the ellipse formed by the phase space trajectories of the individual particles. The different energy slices of the beam primarily differ through their phase (the different slices rotate at different rates). Notice, however, that the matched beta function depends on the beam energy as well. If all the energy slices have the same initial size, only one of them can be matched to the plasma and the others will undergo small amplitude envelope oscillations.

3.1.2 Emittance Growth

In a uniform plasma without accelerating fields, the energy slice emittance is constant. Since the different energy slices rotate at different rates in phase space, as they rotate, they no longer overlap. The result is that the projected emittance will tend to grow as the beam propagates through the plasma. The evolution of the 3D phase space is shown in Fig. 3.3. We derive a general expression that describes emittance growth in Sec. 3.4.2.

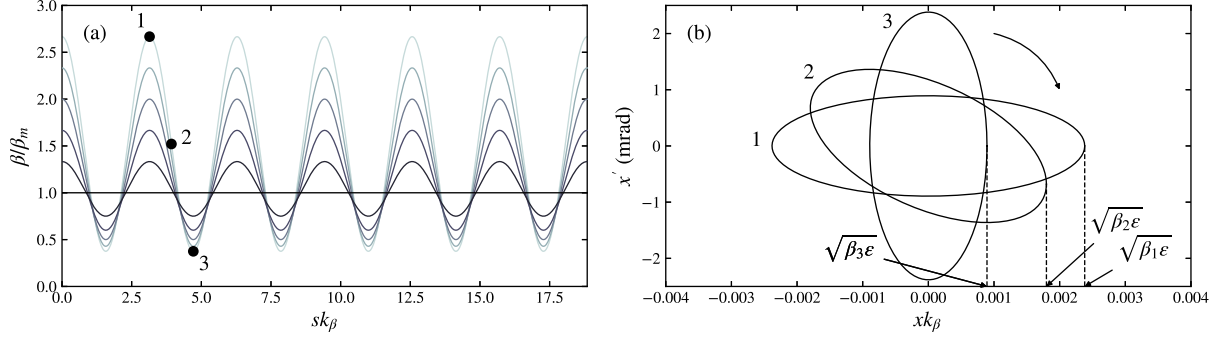


Figure 3.2: Evolution of the beta function in a uniform ion column. (a) evolution of β for different initial values β_0 , lighter is larger β_0 . If the beam starts matched, $\beta = \beta_m$, then no envelope oscillations occur. (b) evolution of the CS invariant ellipse in transverse phase space.

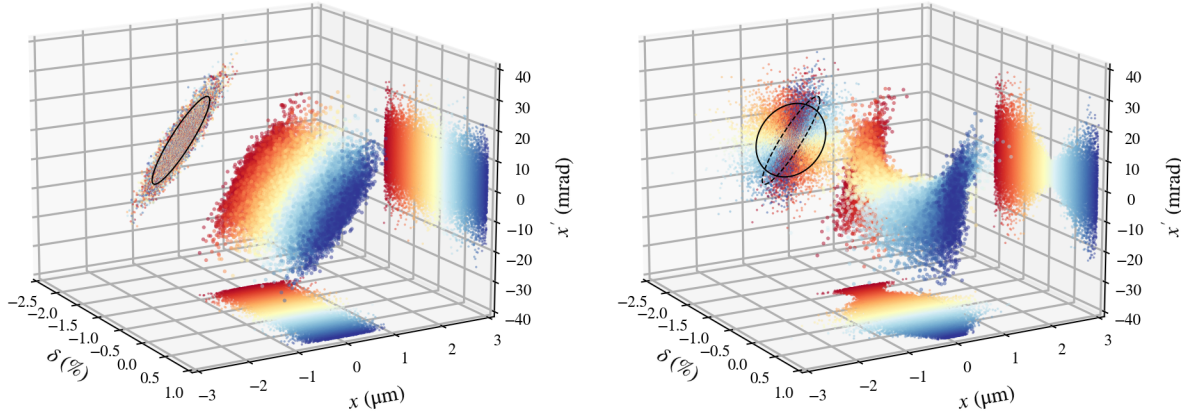


Figure 3.3: Projected emittance growth in 3D phase space. Color represents particle energy. All energy slices start with the same emittance and CS parameters (black ellipse, left). As the beam propagates, each energy slice rotates in phase space with an energy dependent frequency. After sufficient rotations the different energy slices no longer overlap and the projected emittance (black ellipse, right) is larger than the initial emittance (dashed ellipse).

3.2 Transverse Dynamics in a Nonuniform Plasma

Any real plasma source has a longitudinally non-uniform plasma density. Even if the bulk of the plasma is uniform, the density must taper to zero at the ends. The tapered region at the start of the plasma source is called the entrance ramp; the region at the end is called the exit ramp.

Conventional magnetic focusing optics typically are not strong enough to focus the electron beam down to the size needed to match to the bulk plasma. The ramps provides a way to focus the beam down further and control the divergence when exiting the plasma source. Even if matching is not required, the ramps will still be there and will change the beam's parameters. In this section, we work out solutions for the evolution of the beta function in nonuniform plasmas at constant energy. General, analytic solutions for the beam dynamics in an arbitrary shaped ramp are not known. We derive solutions in the adiabatic (long ramp) and perturbative (short ramp) limits and discuss the behavior of the beam in the non-adiabatic regime between the two limits.

This section is written in normalized units where $1/k_{\beta 0}$ is the unit of length and $\eta = n/n_0$ is used for plasma density. The length scale comes from the matched CS parameters, defined as: $\beta_m = 1/k_\beta$, $\alpha_m = -\beta'_m/2$, and $\gamma_m = (1 + \alpha_m^2)/\beta_m$. β_m sets the natural length scale of the transverse evolution. In a uniform plasma with no energy gain, β_m is constant and $\alpha_m = 0$. The matched CS parameters are functions of s due to the variation in plasma density and particle energy along the accelerator. Often, we are concerned with matching a beam that has been focused by a set of conventional optics. The focus of the conventional optics is called the vacuum waist. The vacuum waist is located at a position denoted by s^* with a minimum beam size denoted by β^* .

3.2.1 Adiabatic Evolution

In an adiabatic ramp, the plasma density changes slowly compared to the length scale of the beam's betatron oscillations. Focusing of a beam using a slowly changing plasma density has been considered before [59, 60, 12, 15, 16, 17]. However, the full solution for the beam dynamics, including the evolution of a mismatched beam, has not been explicitly derived. The adiabatic solution is an expansion in α_m ; accordingly, we define the adiabaticity parameter as $\mathcal{A} = |\alpha_m|$. The adiabatic limit is defined as

$$\mathcal{A} = |\alpha_m| = \frac{1}{4\eta k_\beta} \left| \frac{d\eta}{ds} \right| = \frac{1}{4\eta^{3/2}} \left| \frac{d\eta}{ds} \right| \ll 1, \quad (3.4)$$

which is equivalent to saying that the relative change in η over a distance k_β^{-1} is small. This makes no reference to the rate of change of the beam's beta function, only the rate of change of the plasma density.

The transverse dynamics of a single particle are described by the Hamiltonian

$$H = \frac{1}{2}x'^2 + \frac{1}{2}\eta x^2. \quad (3.5)$$

If the plasma density changes adiabatically, the evolution of the beam can be found using the theory of adiabatic invariants. The action-angle variables J_0 and θ_0 in a uniform density plasma are generated by the function

$$F_1(x, \theta_0) = -\frac{x^2}{2\beta_m} \tan \theta_0, \quad (3.6)$$

where the subscript 0 denotes that these variables are only action-angle variables for the Hamiltonian with constant η .

For a Hamiltonian with longitudinally varying plasma density, J_0 and θ_0 are still valid canonical variables, but they have a dependence on s given by

$$\begin{aligned} J'_0 &= -\beta'_m \frac{\partial}{\partial \theta_0} \left(\frac{\partial F_1}{\partial \beta_m} \right) \\ \theta'_0 &= \frac{\partial H}{\partial J_0} + \beta'_m \frac{\partial}{\partial J_0} \left(\frac{\partial F_1}{\partial \beta_m} \right). \end{aligned} \quad (3.7)$$

Using the definition of α_m and taking the derivatives of the generating function gives expressions for J'_0 and θ'_0 where the s dependence of η is a clear perturbation:

$$J'_0 = 2\frac{\alpha_m J_0}{\beta_m} \cos 2\theta_0, \quad \theta'_0 = \frac{1}{\beta_m} - \frac{\alpha_m}{\beta_m} \sin 2\theta_0. \quad (3.8)$$

These expressions can be expanded in powers of α_m , the lowest order term for each coordinate is

$$\theta_0^{(0)} = \int_0^s \frac{d\xi}{\beta_m(\xi)}, \quad J_0^{(0)} = J_0(0). \quad (3.9)$$

Appendix B shows how to use the coordinates to find the transport matrix for single particle motion and calculate the adiabatic solution for the β function. The beta function evolves according to

$$\beta = \beta_m(c_0 + c_1 \cos 2\theta + c_2 \sin 2\theta), \quad (3.10)$$

where $\theta = \theta_0^{(0)}$. The constants are functions of the initial conditions:

$$\begin{aligned} c_0 &= \frac{1}{2} \left(\frac{\beta_0}{\beta_{m0}} + \beta_{m0}\gamma_0 \right) \\ c_1 &= \frac{1}{2} \left(\frac{\beta_0}{\beta_{m0}} - \beta_{m0}\gamma_0 \right) \\ c_2 &= -\alpha_0, \end{aligned} \tag{3.11}$$

where $\beta_{m0} = \beta_m(0)$. In the adiabatic limit the β function oscillates with an amplitude proportional to the matched beam size about a value that is offset from, and proportional to, the matched beam size.

The first order solution is inconsistent in that $\beta'(0) \neq -2\alpha_0$. As a result, the solution does not include oscillations in the beta function that occur when $\alpha_0 = 0$ and $\alpha_{m0} \neq 0$. It was pointed out by Xu *et al.* [18] that these oscillations would never appear in the lowest order adiabatic theory. A ramp with $\eta = 1/(1 - 2\alpha_m s)^2$, which has constant $\mathcal{A}$, can be solved exactly. This solution can be used to derive a higher order correction to c_1 , c_2 , and c_3 .

To find the beam evolution in the $\eta = 1/(1 - 2\alpha_m s)^2$ ramp, we start by taking the derivative of Eq. (2.132), which gives a linear third order differential equation describing the evolution of β :

$$\beta''' + 4\eta\beta' + 2\eta'\beta = 0. \tag{3.12}$$

Using the definition of β_m the equation can be written as

$$\beta_m^3 \beta''' + 4\beta_m \beta' + 8\alpha_m \beta = 0. \tag{3.13}$$

If $\beta_m = \beta_{m0} - 2\alpha_{m0}s$, corresponding to a plasma density of $\eta = 1/(\beta_{m0} - 2\alpha_m s)^2$, then the equation is a Cauchy-Euler equation. Changing variables to β_m makes this obvious

$$\beta_m^3 \alpha_{m0}^3 \frac{d^3 \beta}{d\beta_m^3} + \beta_m \alpha_{m0} \frac{d\beta}{d\beta_m} - \alpha_{m0} \beta = 0. \tag{3.14}$$

The solution has the form

$$\beta = C_n \beta_m^n, \tag{3.15}$$

where C_n is an arbitrary constant and n is a solution of the equation

$$(n-1)(\alpha_{m0}^2 n(n-2) + 1) = 0. \tag{3.16}$$

The beta function evolves according to

$$\beta = \beta_m (c_0 + c_+ e^{i2\Phi} + c_- e^{-i2\Phi}), \quad (3.17)$$

where

$$\Phi = \frac{1}{2\alpha_m} \sqrt{1 - \alpha_m^2} \ln \beta_m. \quad (3.18)$$

Insisting that the solution is real constrains the coefficients such that c_0 is real and $c_+^* = c_-$.

We can define a new pair of coefficients as $c_{\pm} = (c_1 \pm ic_2)/2$. The evolution is now written in the familiar form

$$\beta = \beta_m (c_0 + c_1 \cos 2\Phi + c_2 \sin 2\Phi). \quad (3.19)$$

The third order differential equation introduces an additional constant that can be removed by requiring that the solution satisfy Eq. (2.132). After some algebra, it is found that c_0 is a function of c_1 and c_2 :

$$c_0 = \sqrt{c_1^2 + c_2^2 + \frac{1}{1 - \alpha_m^2}}. \quad (3.20)$$

Letting $\beta_{m0} = 1$, the constants c_1 and c_2 are given in terms of the initial CS parameters at $s = 0$ as

$$c_1 = \frac{1}{2} \left(\beta_0 - \frac{\gamma_0 - 2\alpha_0\alpha_{m0} + \alpha_{m0}^2\beta_0}{1 - \alpha_{m0}^2} \right), \quad (3.21)$$

$$c_2 = \frac{\alpha_0 - \alpha_{m0}\beta_0}{\sqrt{1 - \alpha_{m0}^2}}. \quad (3.22)$$

The constants reduce to the adiabatic solution when $\alpha_{m0} \rightarrow 0$. The solution has the same form as Eq. (3.10), but with different constants. Expanding these constants to first order in α_m and inserting them into the adiabatic solution, captures the oscillations in β . This solution for the beta function corresponds to a transfer matrix where the second row is the derivative of the first. Working backwards it can be shown that the action variable must include some, but not all, of the

second order terms. The higher order constants are given by:

$$\begin{aligned}
c_0 &= \frac{1}{2} \left(\frac{\beta_0}{\beta_{m0}} + \beta_{m0}\gamma_0 - 2\alpha_0\alpha_{m0} \right) \\
c_1 &= \frac{1}{2} \left(\frac{\beta_0}{\beta_{m0}} - \beta_{m0}\gamma_0 + 2\alpha_0\alpha_{m0} \right) \\
c_2 &= -\alpha_0 + \frac{\alpha_{m0}\beta_0}{\beta_{m0}}.
\end{aligned} \tag{3.23}$$

The oscillations around the matched size disappear when $\beta_0 = \frac{\beta_{m0}}{\sqrt{1+\alpha_{m0}^2}}$ and $\alpha_0 = \frac{\alpha_{m0}}{\sqrt{1+\alpha_{m0}^2}}$. In the adiabatic limit this simplifies to $\beta_0 = \beta_{m0}$ and $\alpha_0 = \alpha_{m0}$. For a nonuniform plasma, we define this as the matched condition (this definition of matched was previously noted by Williams and Katsouleas [60]). If a beam is initially matched to an adiabatic plasma profile, the beam will remain matched to the plasma.

If it is assumed that $\beta_0 = \beta_{m0}$ and $\alpha_0 = \alpha_{m0}$, then very small oscillations of β occur. Oscillations can also develop due to rapid changes in α_m , even if the ramp remains adiabatic. Oscillations induced by discontinuous changes in α_m can be explained using the adiabatic solution with higher order constants. For example, if the gradient of the plasma density at the beginning of the exit ramp is discontinuous, a beam that is initially matched to the plasma in the accelerating region, where $\beta_0 = 1$ and $\alpha_0 = 0$, will not be matched to the exit ramp. Subsequently, it will undergo oscillations. Examples of these oscillations for two exit ramps where the beta function can be found exactly are shown in Fig. 3.4. The ramp is identical to the primary ramp considered in [18], which has an exact solution for the beam evolution. As shown in the figure, the oscillations are not captured by the lowest order adiabatic solution, but they are captured by using the higher order correction to the constants.

3.2.2 Emittance Growth of the Matched and Unmatched Beam

Even if the reference energy slice is matched to the plasma, different energy slices of the beam will experience different focusing forces leading to projected emittance growth. Each energy component has a different matched beta function and will undergo adiabatic evolution according to Eq. (3.10). To lowest order in δ , the projected emittance growth is a function of the beam's

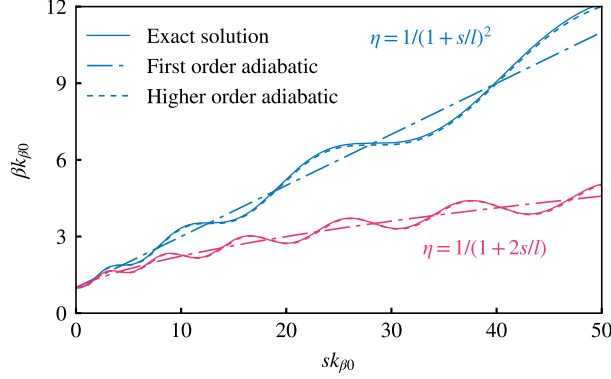


Figure 3.4: Comparison of the lowest order adiabatic solution and the higher order adiabatic solution. Two different ramp profiles are portrayed, both with $l = 5/k_{\beta 0}$. The exact solution for the β function evolution is shown as a thin solid line, the lowest order adiabatic solution as a dot-dashed line and the higher order solution as a dotted line. The higher order solution correctly predicts the oscillations of the beam around the matched β function, which is equivalent to the first order adiabatic solution.

chromatic amplitude [61]:

$$\frac{\Delta\epsilon}{\epsilon_0} = \frac{1}{2}W^2\sigma_\delta^2, \quad (3.24)$$

where σ_δ is the RMS fractional momentum deviation, and the chromatic amplitude W is defined as

$$W^2 = \left(\frac{\partial\alpha}{\partial\delta} - \frac{\alpha}{\beta} \frac{\partial\beta}{\partial\delta} \right)^2 + \left(\frac{1}{\beta} \frac{\partial\beta}{\partial\delta} \right)^2. \quad (3.25)$$

Here, all energy slices of the beam are assumed to start with the same geometric emittance ϵ_0 and all the terms are evaluated at $\delta = 0$. The expansion is only valid if $\beta(\delta)$ is approximately linear (even for small δ the quadratic term can sometimes be important).

It can be seen from Eq. (2.132) that $\beta(\delta)$ is given by making the replacement $\beta_m \rightarrow \beta_m\sqrt{1+\delta}$. Additionally, to account for any chromatic effects of the incoming beam, the initial conditions are functions of δ : $\beta_0(\delta)$, $\alpha_0(\delta)$ and $\gamma_0(\delta)$. Making the above substitution for β_m , β_{m0} , and α_{m0} gives

$$\beta(\delta) = \beta_m \left[c_3 \frac{\delta}{2} (1 - \cos 2\psi) + c_2 \sqrt{1+\delta} \sin 2\psi + c_0 + c_1 \cos 2\psi \right], \quad (3.26)$$

where $c_3 = \beta_{m0}\gamma_0 - 2\alpha_{m0}\alpha_0$, and

$$\psi = \frac{\theta}{\sqrt{1+\delta}}. \quad (3.27)$$

All the c_i constants are functions of δ through β_0 , α_0 , and γ_0 . β_m is the matched beta function for the average beam energy $\langle\gamma_b\rangle$. It is convenient to expand ψ in δ : $\psi \approx (1 - \delta/2)\theta$. Evaluating the chromatic amplitude will give the lowest order emittance growth for a beam with arbitrary chromatic initial conditions.

Consider the case where all energy components start with the same initial conditions, and the reference energy is matched to the plasma. In this case the beam starts with its projected emittance equal to the emittance of each energy slice. $\beta(\delta)$ is given by

$$\beta(\delta) = \beta_m \left[\frac{\delta}{2}(1 - \cos 2\psi) + 1 \right]. \quad (3.28)$$

Evaluating the chromatic amplitude shows that the emittance evolves according to

$$\frac{\Delta\epsilon}{\epsilon_0} = \frac{1}{4}(1 - \cos 2\theta)\sigma_\delta^2 \quad (3.29)$$

to lowest order in δ . This expression also holds true in a uniform plasma. The beam evolution and emittance evolution in an adiabatic plasma are shown in Fig. 3.5 for two different initial conditions. The model matches the numerical solution quite well in both cases. An oscillation of the projected emittance can occur so long as the emittance of each energy slice is smaller than the projected emittance of the entire beam. In the presence of non-uniform longitudinal fields, particle exchange between energy slices will lead to slice emittance growth and saturation of the projected emittance of the beam. The emittance will saturate to the maximum of the oscillation of the projected emittance. Figure 3.5(b) shows the emittance evolution for initial conditions where each energy slice is matched to the plasma. For most matched beams, the emittance growth in an adiabatic section would be negligible, on the order of σ_δ^2 .

If the reference energy is mismatched, $\beta(\delta)$ cannot be approximated as a linear function and the chromatic amplitude no longer predicts the emittance growth. The emittance growth in a uniform plasma from chromatic phase mixing has been found by [18] to be

$$\epsilon = \epsilon_{sat} \sqrt{1 - \frac{(\gamma_0\beta_m + \beta_0/\beta_m)^2 - 4}{(\gamma_0\beta_m + \beta_0/\beta_m)^2} \left(\frac{\sin \Delta\psi}{\Delta\psi} \right)^2}, \quad (3.30)$$

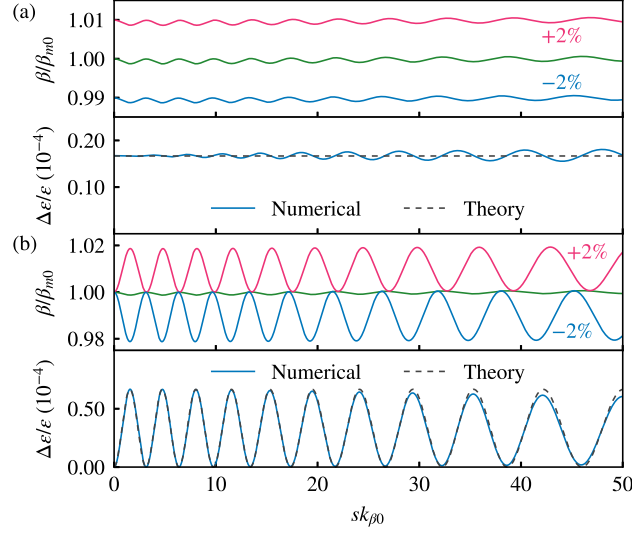


Figure 3.5: Emittance evolution for a beam in an adiabatic plasma where (a) all energy components are matched, and (b) start with the same Courant-Snyder parameters. The beam has a uniform energy distribution from -2% to 2% with the extremes of the distribution plotted. The plasma profile is $\eta = 1/(1 + (as)^2)$ with $a = 0.05$.

where $\Delta\psi = \psi(\sqrt{3}\sigma_\delta) - \psi(-\sqrt{3}\sigma_\delta)$. The saturated emittance is given by $\epsilon_{sat} = \epsilon_0(\gamma_0\beta_m + \beta_0/\beta_m)/2$, and represents the asymptotic upper limit of the emittance after many betatron oscillations. The phase mixing process works the same in an adiabatic plasma, so Eq. (3.30) can be used to describe the emittance growth of a mismatched beam in an adiabatic ramp.

It is useful to quantify the adiabatic ramp by the total phase advance in the ramp, θ_f . The chromatic phase spread can then be written as $\Delta\psi \approx -\theta_f\delta$. Most ramps have θ_f on the order of 10s of radians and $\Delta\psi$ is small. Even for long adiabatic ramps, the emittance growth in the ramp is negligible. Adiabatic ramps are a robust option for matching a beam into a plasma. There is almost no emittance growth in the ramp if the beam is initially matched, and the exact shape of the ramp is inconsequential, so long as the plasma density changes slowly and $\mathcal{A}$ remains small.

3.2.3 Non-Adiabatic Evolution

Ramps that are non-adiabatic, or have sections that are non-adiabatic, can focus beams over a much shorter distance than an adiabatic ramp with a matched beam. In general, the

beam evolution in a non-adiabatic ramp cannot be solved for exactly; however, some ramp profiles have exact solutions that demonstrate the general character of the evolution. One such ramp is $\eta = 1/(1 - 2\alpha_m s)^2$. The exact solution for the ramp is given in the Sec. 3.2.1. It is similar to the adiabatic solution except for a difference in constants and a correction to the phase. In the non-adiabatic regime it is impossible to choose a set of initial conditions such that the oscillating terms vanish for all s .

The fastest way to focus the beam in a non-adiabatic ramp is to harness the phase of the oscillation where the beam size is most rapidly decreasing. The non-adiabatic regime is unique in that the beam can undergo oscillations and still end up matched to the plasma. This is impossible in a smooth adiabatic ramp. In a well chosen non-adiabatic ramp, the beam undergoes almost exactly a quarter betatron oscillation and ends matched to the plasma. This is similar to the scheme presented in Sec. 3.6.2. The focusing process is shown in Fig. 3.6 where a non-adiabatic ramp is compared to an adiabatic ramp. The non-adiabatic ramp decreases the beta function by a factor of 30 in only $29/k_{\beta 0}$ with the beam undergoing approximately a quarter of a betatron oscillation. An adiabatic ramp at the upper limit of the adiabatic regime ($\alpha_m = 0.1$ for the majority of the ramp) requires a distance of $152/k_{\beta 0}$ to achieve the same degree of focusing.

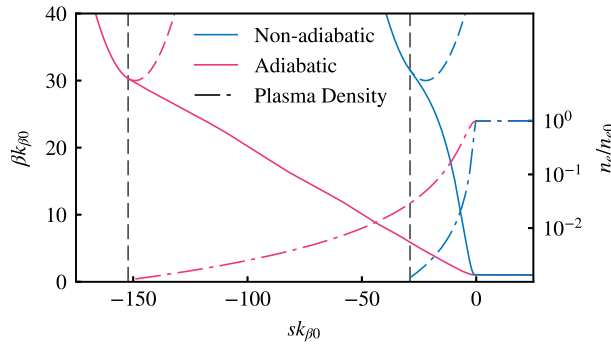


Figure 3.6: Evolution of two identical beams with shifted vacuum waist locations through an adiabatic ramp (magenta) and a non-adiabatic ramp (blue). In the adiabatic ramp, the beam evolution follows the adiabatic solution, reducing β by a factor of 30 in a distance of $152 k_{\beta 0}^{-1}$. In the non-adiabatic ramp, the beam undergoes a quarter of a betatron oscillation, reducing β by a factor of 30 in $29 k_{\beta 0}^{-1}$. In both cases the beam ends matched to the uniform density acceleration section, with $\beta = 1/k_{\beta 0}$ and $\alpha = 0$.

Generally, numerical methods must be implemented to design a smooth, matching ramp that provides rapid focusing. The ramp must transition from non-adiabatic to adiabatic exactly at the point where the beam oscillation passes through the matched solution. In practice, the matched Courant-Snyder parameters at the top of the non-adiabatic section can be numerically back propagated through a ramp of a given shape while scanning the characteristic ramp length until the ideal matching ramp length is identified.

Analytic solutions exist for ramps where $\beta_m(s)$ is linear or quadratic (see [16] for the quadratic case) which can be used to approximate the non-adiabatic part of an arbitrary ramp. Non-adiabatic ramps are sensitive to the exact ramp shape due to the interplay between oscillation and matching. These difficulties make non-adiabatic ramps potentially difficult to design and implement in situations where the plasma density cannot be precisely controlled.

3.2.4 Perturbative Plasma Ramp

If the plasma has a sufficiently low density, the motion arising from the focusing force of the ion column can be treated as a perturbation to the vacuum propagation of the beam. The unperturbed Hamiltonian is the free space Hamiltonian and it is well known that the canonical transformation obtained using Hamilton's principle function results in the coordinates given by

$$A = x'_0 \quad B = x - As. \quad (3.31)$$

According to classical, time dependent perturbation theory, the time evolution of these coordinates is given by

$$\dot{A} = -\eta(s)(As + B) \quad \dot{B} = \eta(s)s(As + B). \quad (3.32)$$

The zeroth order solution is the constant solution of the unperturbed Hamiltonian: $A_0 = x'_0$ and $B_0 = x_0$. The first order solution is found by inserting the zeroth order solution into the right hand side of the differential equation and integrating. The first order solution can be written in the form

of a transport matrix for x and x' from $s = 0$ to s :

$$M = \begin{pmatrix} 1 - I_0 s + I_1 & s - I_1 s + I_2 \\ -I_0 & 1 - I_1 \end{pmatrix} \quad (3.33)$$

where I_n refers to integrals given by

$$I_n = \int_0^s \zeta^n \eta(\zeta) d\zeta. \quad (3.34)$$

The approximation is valid if $\frac{1}{2}\eta x^2 \ll \frac{1}{2}x'^2$, requiring that $\beta k_\beta^2/\gamma \ll 1$. A more detailed derivation is given in Appendix C. The β function evolution can be derived in the normal way from the components of the transport matrix. Figure 3.7 shows that the approximation well represents the beam evolution in the low density tail of an exit ramp with a Gaussian density profile.

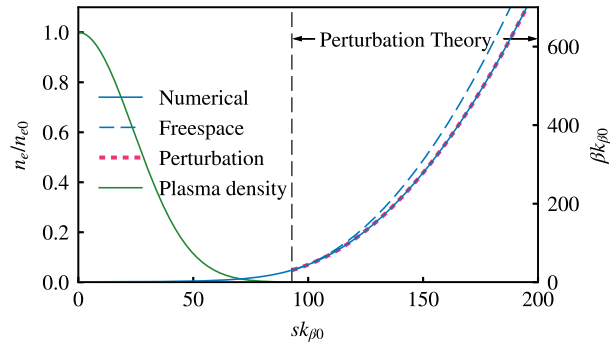


Figure 3.7: Evolution of an initially matched beam in an exit ramp with a long perturbative tail. The perturbative theory effectively models the small deviation in β from free space evolution caused by the long, low density tail of a Gaussian plasma density ramp (green).

As s becomes very large, the integrals converge for any practical ramp. Beyond this point, the plasma has no effect on the beam and the beam evolves as it would in vacuum. The beam will appear to have come from a vacuum waist of $\beta^* = 1/\gamma$:

$$\beta^* = \frac{1}{\gamma_0 + 2(\alpha_0 I_0 - \gamma_0 I_1) + \beta_0 I_0^2 - 2\alpha_0 I_0 I_1 + \gamma_0 I_1^2}, \quad (3.35)$$

where the I_n integrals are assumed to be convergent and evaluated to infinity. After the beam has left the perturbative region of an exit ramp, β^* will be the effective waist seen by the downstream magnetic focusing optics. For an entrance ramp, this formula can be used to calculate the focusing

of the beam that occurs at the foot of the ramp. It can be useful to drop terms of order I^2 to get a compact version of the formula.

It is straightforward to account for chromatics in the perturbative ramp, a factor of $1/(1+\delta)$ appears in front of each I_n integral. The transport matrix for off energy particles can easily be found and $\beta(\delta)$ solved for. The chromatic amplitude gives emittance growth induced in the perturbative part of the ramp, which will be negligible in most cases. The net effect of the perturbative ramp on the motion does not depend on the exact shape of the ramp, only the values of the integrals I_n evaluated over the entire ramp.

3.2.5 Summary of the Effects of the Plasma Ramps

A tailored longitudinal plasma density profile can be used to match a beam into or out of a plasma-based accelerator in order to minimize chromatic emittance growth in the beam. The length of the ramp, location of the vacuum waist, and emittance growth in the ramp can be calculated analytically for a large range of ramp shapes by separating arbitrary plasma profiles - or portions thereof - into three qualitatively distinct types. The first is an adiabatic plasma ramp where the plasma density changes slowly. In this type of ramp, a beam that starts out matched to the plasma remains matched to the plasma and undergoes minimal emittance growth. Adiabatic ramps are long but insensitive to the exact form of the density profile, so long as the gradient of the plasma density remains small. The second is a non-adiabatic plasma ramp where the beam is not matched to the plasma and undergoes a fraction of an oscillation. This type of plasma ramp can focus the beam quickly, but the beam dynamics are sensitive to the exact shape of the plasma density profile. The third is a perturbative plasma ramp where the plasma density is low, and the beam evolution is only slightly perturbed from vacuum-like propagation. This type of ramp is somewhat insensitive to the form of the plasma density profile but can only cause minimal focusing of the beam. The beam dynamics can be calculated exactly for the adiabatic and perturbative ramps. The beam dynamics in a non-adiabatic section can be calculated numerically or approximated by a ramp where the dynamics are analytically solvable.

A wide range of ramp shapes can match the same beam into a plasma-based accelerator, but different density profiles can affect the focusing process. The fastest focusing can be achieved with a subset of the adiabatic plasma ramps, wherein a highly mismatched beam undergoes a quarter of a betatron oscillation in a low density uniform plasma (Sec. 3.6.2). In contrast, adiabatic ramps are long, but are insensitive to errors in the plasma density profile. Multiple ramp types can be combined to produce a ramp that suits a particular set of experimental needs. The theoretical work presented here can be used to compare different ramp shapes by studying the sensitivity of the emittance growth to the initial beam conditions and variations in the plasma density profile. This can be extended to include the effects of non-uniformity in the accelerating section of the plasma.

3.3 Acceleration and Wake Loading

In the previous sections, we worked out the emittance growth assuming the energy of each particle in the beam remained constant. This is instructive to understand the chromatic dephasing mechanism and the impact of the ramps, within which the energy change is minimal. The entire point of a plasma wakefield accelerator is to accelerate the particles and increase their energy. The relative energy gain within a single stage can be significant, particularly in the early stages of an accelerator. Significant changes in the energy distribution of the beam during acceleration will impact the chromatic dynamics, especially when wake loading is included.

Before discussing the general solution for the transverse dynamics with energy gain, we want to touch on the energy gain and energy spread expected from PWFA experiments at FACET-II. Sec. 3.2 on transverse dynamics was largely focused on how to design the entrance and exit ramps to preserve the beam emittance. This section focuses on how to choose the bulk plasma density and set a requirement for the plasma width. Although the simulations are specific to FACET-II, the approach is general. The FACET-II accelerator can deliver a wide range of beams, but there are still important limitations on the beam parameters. Choosing the plasma density to match the beam parameters has a large impact on the beam loading and thus the accelerator's final energy

spread and transformer ratio.

3.3.1 Acceleration and Loading for the FACET-II Beam

The drive beam for the two bunch configuration at FACET-II is designed to be as high current as possible while maintaining an approximately 3 to 1 ratio between the charge of the drive and witness beam. The ratio sets the trade-off between transformer ratio T and efficiency η ; 3 to 1 corresponds to $T \approx 1$ and $\eta \approx 30\%$. Early simulations of FACET-II were run to find a machine configuration with a nominal drive-witness separation of $150 \mu\text{m}$ which resulted in bunch lengths of $5.2 \mu\text{m}$ and $3.6 \mu\text{m}$ for the drive and witness beam, respectively [62, 63]. The accelerator is capable of delivering beams with a range of values around the nominal parameters while maintaining the drive-witness charge ratio. Achieving the maximum performance from the accelerator requires tuning the plasma density in concert with the witness bunch length and drive-witness spacing to maximize the energy and minimize the energy spread of the witness beam.

Table 3.1: FACET-II beam parameters. A * superscript refers to the value of the parameter at the witness beam waist.

	Drive Beam	Witness Beam
Charge (nC)	1.5	0.5
Energy (GeV)	10	10
$\epsilon_{nx,ny}$ ($\mu\text{m rad}$)	3.4, 3.0	3.2, 3.1
$\beta_{x,y}^*$ (cm)	70, 70	5, 5
$\alpha_{x,y}^*$	4.2, 1.6	0, 0
σ_z (μm)	5.2	3.6
σ_δ	1%	1%

We ran PIC simulations at three different plasma densities for witness beams with varying σ_z and delays Δz behind the drive beam. The charge, energy, bunch length, emittance, and energy spread used in the simulations are summarized in Table 3.1. All beams were initialized with Gaussian spatial distributions (see Sec. 2.3.4) and uniform energy distributions. Both beams were initialized matched to the plasma. The simulations were propagated until the wakefields stabilized. The final energy and estimated energy spread of the witness beam at the end of the accelerator

stage was compared between the different cases. The energy spread of the beam can be estimated for the full plasma using the fields from the short simulation with the following procedure: The accelerating field has a dependence on ξ given by some function $E_z(\xi, s)$, which in general depends on s . Assume the shape of the accelerating field is constant as the beam travels down the accelerator, $E_z(\xi, s) = E_z(\xi)$. Assuming the beam is monochromatic and starts with current density $\lambda(\xi)$, then the final energy distribution after a distance L is given by

$$\mathcal{E}(\xi) = \mathcal{E}_0 + E_z(\xi)L. \quad (3.36)$$

Here, $\mathcal{E}_0$ is the beam's initial energy and we have ignored any longitudinal motion of beam particles within the beam. The final average energy of the beam is given by

$$\bar{\mathcal{E}} = \frac{\int_{-\infty}^{\infty} \mathcal{E}(\xi) \lambda(\xi) d\xi}{\int_{-\infty}^{\infty} \lambda(\xi) d\xi}. \quad (3.37)$$

The final RMS energy spread in the beam is given by

$$\sigma_{\delta}^2 = \frac{\int_{-\infty}^{\infty} (\mathcal{E}(\xi) - \bar{\mathcal{E}})^2 \lambda(\xi) d\xi}{\bar{\mathcal{E}}^2 \int_{-\infty}^{\infty} \lambda(\xi) d\xi}. \quad (3.38)$$

The acceleration length is calculated by solving for L such that $\bar{\mathcal{E}} - \mathcal{E}_0 = 10 \text{ GeV}$. This is only an estimation because the final energy, and thus L , will depend on the transformer ratio.

Fig. 3.8 shows the set of simulations for a plasma density of $5 \times 10^{16} \text{ cm}^{-3}$ which demonstrate the typical behavior of the gradient and energy spread as σ_z and Δz are varied. As the witness beam moves further back in the wake, the average accelerating gradient increases until the rear of the beam starts to exceed the bounds of the wake. The maximum accelerating gradient is larger for shorter witness beams because they can be pushed further back in the wake while still being fully contained in the bubble. The gradient is relatively insensitive to the bunch length of the witness beam, while the energy spread is highly sensitive as the length determines how well the beam loads the wake. As shown from the simulations, a relatively short beam is required to sufficiently load the wake. Although not shown at this density, if the beam length is reduced even further, it will eventually overload the wake and the energy spread will increase. Finally, notice that there is an

optimum drive-witness separation for minimum energy spread that is typically shorter than the separation that results in the maximum accelerating gradient.

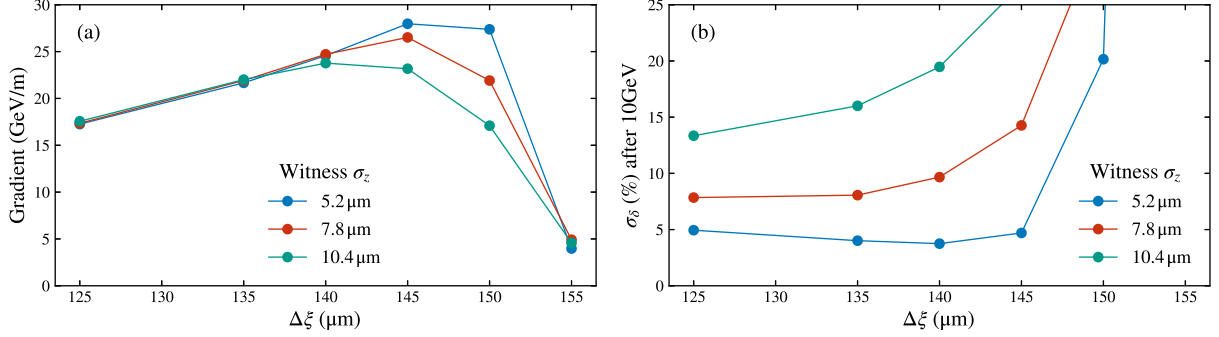


Figure 3.8: Average accelerating gradient and estimated final energy spread as a function of the drive-witness spacing and the witness beam bunch length. (a) the accelerating gradient of the witness beam increases as it is moved towards the rear of the wake. At some point, the beam no longer fits within the wake and the gradient drops off. (b) estimated energy spread of the witness beam from Eq. (3.38) after 10 GeV of acceleration. The energy spread depends on how well the beam loads the wake.

The accelerating gradient is not a suitable quantity to compare across plasma densities because the decelerating gradient felt by the drive beam varies. The more appropriate quantity is the transformer ratio which is approximately proportional to the final energy of the witness beam.¹

The transformer ratio and energy spread are compared across all the simulations in Fig. 3.9. The points near the lower right corner represent the best combination of large transformer ratio and low energy spread. For FACET-II, the optimal plasma density for these beam parameters is $2.6 \times 10^{16} \text{ cm}^{-3}$ with a $\sigma_z = 5.2 \mu\text{m}$ witness beam. This combination can be tuned for low energy spread or a balance between energy spread and transformer ratio by changing the spacing between the two beams.

¹ The final energy of the witness beam is the average accelerating gradient multiplied by the length of the accelerator. The length of the accelerator is set by how long the drive beam can sustain the wake. Due to drive beam effects, such as head erosion, this distance is not simply the energy of the drive beam divided by the average decelerating gradient. Effects such as head erosion are density dependent, so the transformer ratio is an approximate, but useful, comparison between densities.

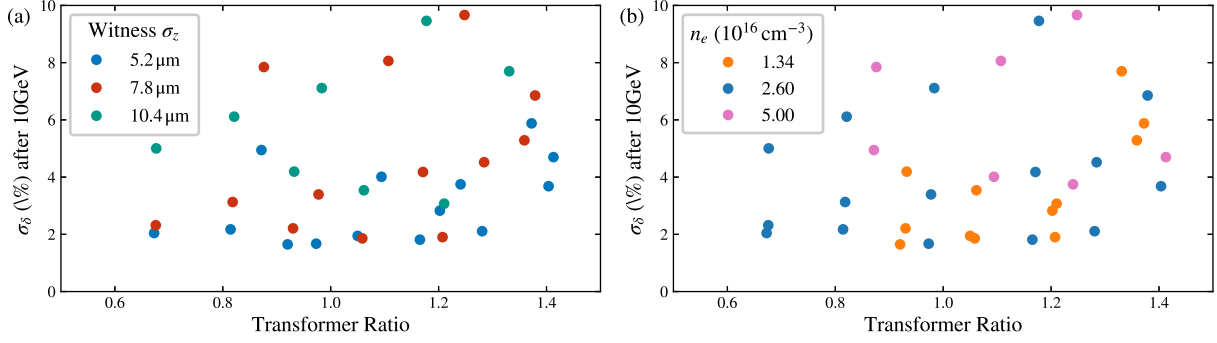


Figure 3.9: Transformer ratio and estimated final energy spread as a function of (a) plasma density and (b) witness beam bunch length. For the beam parameters considered here, a plasma density of $2.6 \times 10^{16} \text{ cm}^{-3}$ with a $\sigma_z = 5.2 \mu\text{m}$ witness beam produces both the largest transformer ratio and lowest energy spread.

3.3.2 Plasma Width

The plasma must be sufficiently wide to fully contain the wake. We know from nonlinear theory (Sec. 2.2.3) that the wake is approximately spherical with a diameter of λ_p . The linear perturbation around the wake is described by linear fluid theory and extends beyond the wake by about one skin depth c/ω_p [42]. We would expect the width of the plasma would need to be at least

$$W = \lambda_p + 2 \frac{c}{\omega_p} \approx 8.3 \frac{c}{\omega_p}, \quad (3.39)$$

in order to fully enclose the wake. To test this approximation, we ran a series of PIC simulations using approximate FACET-II beam parameters with different plasma widths. The simulations used a plasma density of $2.6 \times 10^{16} \text{ cm}^{-3}$, and the witness beam is located $180 \mu\text{m}$ behind the drive beam. The rest of the beam parameters are given in Table 3.1. The beams were propagated for 2 mm in the plasma to let the wakefields stabilize. The minimum plasma width according to Eq. (3.39) is $273 \mu\text{m}$, simulations were run with plasma widths from $140 \mu\text{m}$ to $280 \mu\text{m}$ in $35 \mu\text{m}$ increments. The resulting electron density is shown in Fig. 3.10 where the distortion of the bubble is clearly visible.

Reducing the plasma width primarily changes the longitudinal field. The final energy and energy spread of the witness beam, shown in Fig. 3.11, can then be used to quantify the impact

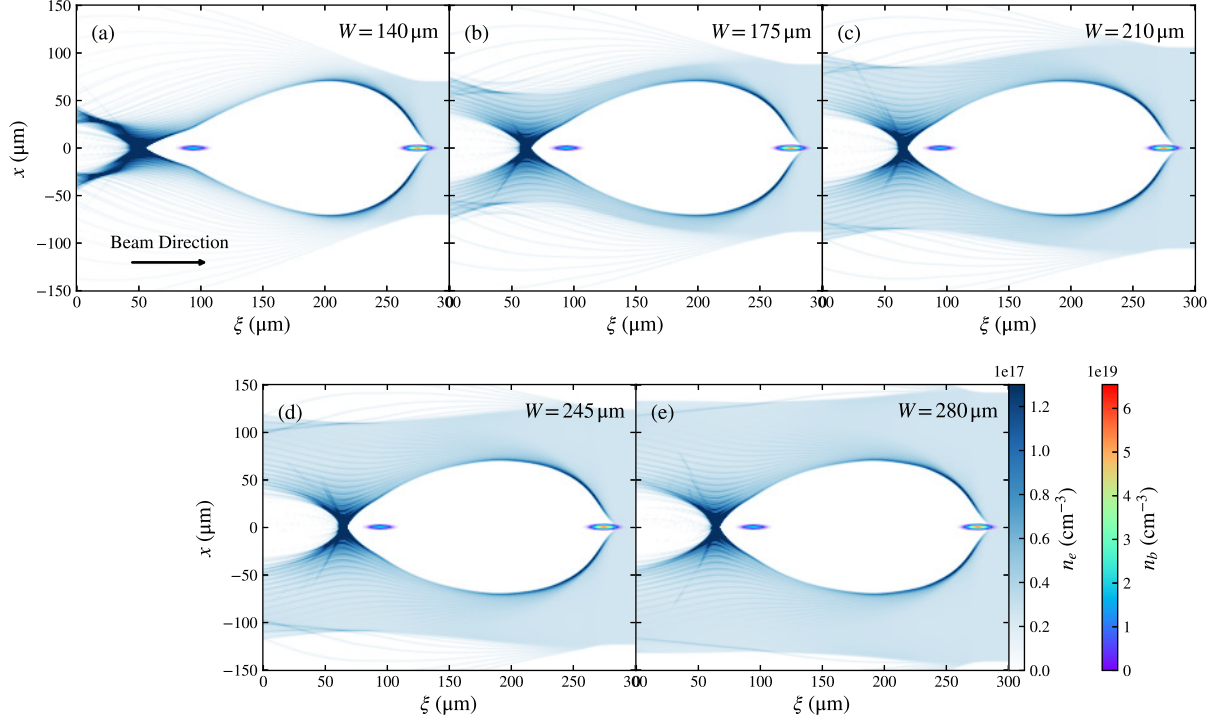


Figure 3.10: Plasma and beam density for plasmas with initial widths of (a) $140\mu\text{m}$, (b) $175\mu\text{m}$, (c) $210\mu\text{m}$, (d) $245\mu\text{m}$, and (e) $280\mu\text{m}$.

of the plasma width on the accelerator performance. Both parameters converge at a plasma width of $210\mu\text{m}$, well below the predicted value of $273\mu\text{m}$ from Eq. (3.39). The disagreement is not surprising; in the widest plasma simulated, the wake is found to be $213\mu\text{m}$ long and $136\mu\text{m}$ wide. The difference from $\lambda_p = 207\mu\text{m}$ results because the drive beam does not have sufficient current to drive a highly nonlinear wake. If we take the measured bubble size and add two skin depths, we get $202\mu\text{m}$ which agrees well with the simulation. The final bunch length of the FACET-II drive beam is not yet known, but some simulations predict a much larger peak current; therefore, Eq. (3.39) is a good target for the minimum width of the FACET-II plasma source.

The plasma width becomes important in the ramps where λ_p is much larger than in the bulk plasma. As can be seen in Fig. 3.11, the accelerating gradient is reduced as the plasma width narrows. In the ramps, this can be advantageous because it limits the energy change of the beam in the ramp. From the simulations of different plasma densities presented in the previous section,

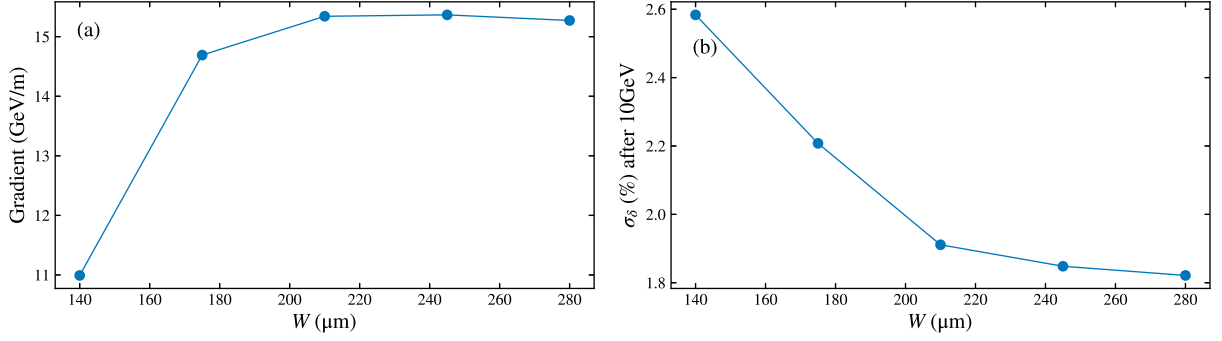


Figure 3.11: Convergence of the accelerating gradient and estimated final energy spread with plasma width. Convergence occurs around $W = 210\mu\text{m}$.

we know the witness beam only optimally loads the wake at one plasma density. As it traverses the ramp, the energy spread will grow and, at low plasma density, the witness beam can be in the decelerating phase of the wake. A narrow plasma in the ramps mitigates both of these effects.

3.4 A General Approach Including Acceleration

Chromatically driven emittance growth results from the variation in particle energy within the beam. Wake loading and acceleration will affect this energy variation and thus the beam evolution. In this section, we develop a model for the transverse dynamics of the beam that takes these features into account. We find that the presence of wake loading results in mixing between energy slices and thus growth in the energy slice emittance. Using our model, we are able to derive analytic saturation lengths for the projected and slice emittances. We finish with an analytic calculation of the emittance growth within a full plasma source with energy gain and density ramps.

3.4.1 Single Particle Motion

The equations of motion in the transverse directions for an electron in the witness beam are decoupled and given by Eq. (2.121). In general, K is a function of s through the local plasma density and the particle energy. An approximate solution to the equation of motion can be derived for a plasma density that varies adiabatically in s , i.e., $|\alpha_m(s)| \ll 1$. We start with the solution

derived in Appendix B for the motion of a single particle. Eq. (B.12) gives the transport matrix for a particle in an adiabatically varying plasma density, in the absence of energy gain. From this the transverse motion of a single particle is found to be [57, 20]:

$$\begin{aligned}
 x &= x_0 \sqrt{\frac{\beta_m}{\beta_{m0}}} (\cos \phi + \alpha_{m0} \sin \phi) + x'_0 \sqrt{\beta_m \beta_{m0}} \sin \phi \\
 x' &= x_0 \frac{(\alpha_{m0} - \alpha_m) \cos \phi - (1 + \alpha_{m0} \alpha_m) \sin \phi}{\sqrt{\beta_m \beta_{m0}}} \\
 &\quad + x'_0 \sqrt{\frac{\beta_{m0}}{\beta_m}} (\cos \phi - \alpha_{m0} \sin \phi),
 \end{aligned} \tag{3.40}$$

where the subscript 0 indicates the initial value of a variable at $s = s_0$ and $x' = dx/ds = p_x/p_s$, p_x and p_s are the longitudinal and transverse momentum of the particle, respectively. The betatron phase advance is defined as

$$\phi = \int_{s_0}^s k_\beta(s') ds'. \tag{3.41}$$

We are interested in the general case where the particles can gain and lose energy. Energy gain (loss) results in a reduction (increase) in the amplitude of the oscillations due to adiabatic dampening. We use the ansatz that the single particle motion has an additional s dependent amplitude $A(s)$:

$$x = A(s)x_c = x_0 A(s) \sqrt{\frac{\beta_m}{\beta_{m0}}} (\cos \phi + \alpha_{m0} \sin \phi) + x'_0 A(s) \sqrt{\beta_m \beta_{m0}} \sin \phi, \tag{3.42}$$

where x_c is the position of a single particle with constant energy given by Eq. (3.40). x_c satisfies the differential equation $x_c'' + k_\beta^2 x_c = 0$. Inserting the ansatz into Eq. (2.121) and requiring that A can take on any value gives a differential equation for A :

$$2A' + \frac{\gamma'_b}{\gamma_b} A = 0. \tag{3.43}$$

The solution of which is $A(s) = \sqrt{\gamma_{b0}/\gamma_b}$. For Ax_c to approximately satisfy the equation of motion, Eq. (2.121), the relative change in energy over one betatron period must be small $\gamma'_b/\gamma_b \ll k_\beta$ and $\gamma''_b/\gamma_b \ll k_\beta^2$.

The transport matrix M defines the motion of the particle based on its initial conditions

$$\begin{pmatrix} x \\ x' \end{pmatrix} = M \begin{pmatrix} x_0 \\ x'_0 \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} x_0 \\ x'_0 \end{pmatrix}. \quad (3.44)$$

Combining the adiabatic dampening term $A(s)$ with x_c gives the transport matrix for a particle in an adiabatic plasma with energy change:

$$\begin{aligned} M_{11} &= \sqrt{\frac{\gamma_{b0}}{\gamma_b} \frac{\beta_m}{\beta_{m0}}} (\cos \phi + \alpha_{m0} \sin \phi) \\ M_{12} &= \sqrt{\frac{\gamma_{b0}}{\gamma_b} \beta_m \beta_{m0}} \sin \phi \\ M_{21} &= \sqrt{\frac{\gamma_{b0}}{\gamma_b} \frac{(\alpha_{m0} - \alpha_m) \cos \phi - (1 + \alpha_{m0} \alpha_m) \sin \phi}{\sqrt{\beta_m \beta_{m0}}}} \\ M_{22} &= \sqrt{\frac{\gamma_{b0}}{\gamma_b} \frac{\beta_{m0}}{\beta_m}} (\cos \phi - \alpha_{m0} \sin \phi). \end{aligned} \quad (3.45)$$

In a uniform plasma, $\beta_m = \beta_{m0} \sqrt{\gamma_b / \gamma_{b0}}$. Inserting β_m into Eq. (3.45) and assuming the energy gain over a single betatron period is small, $\alpha_m \approx 0$, we recover the expression for single particle motion in a uniform plasma from [13, 19].

Fig. 3.12 shows a comparison between Eq. (3.44) and a numerical solution to Eq. (2.121) for a single electron travelling through a plasma-based accelerator. The plasma stage is designed to double the particle's energy and has identical adiabatic density ramps on each end. The entrance ramp reduces the incoming beta function by a factor of 10, and the beam undergoes approximately 12 betatron periods within the uniform plasma section. $k_{\beta u}$ is the betatron wavenumber at the start of the uniform plasma section, located at $sk_{\beta u} = 200$. The parameters used here are similar to those found in current beam driven plasma wakefield accelerator experiments such as those at FACET-II [62, 63] and FLASHForward [64, 65]. The analytic solution shows excellent agreement with the numerical solution; it accurately captures both the adiabatic dampening and focusing in the ramp. In the next section, we use this solution to derive general expressions for the evolution of the beam's moments and emittance.

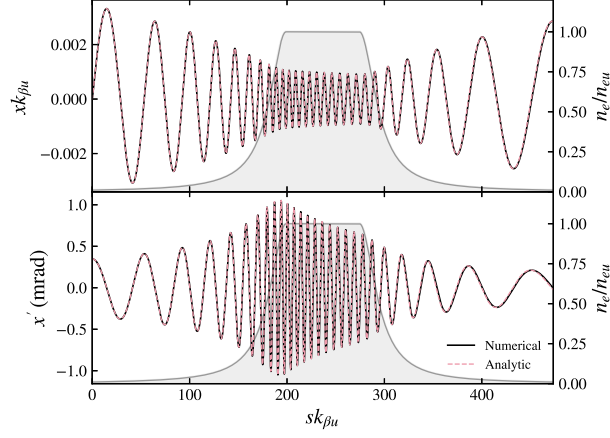


Figure 3.12: Single particle motion in a plasma accelerator with adiabatic ramps. The particles energy doubles within the accelerator. The top plot shows the evolution of the particle's transverse position x . The bottom plot shows the evolution of the particle's trajectory angle x' . The solid line shows the numerical solution to the transverse equation of motion, the dashed line shows the analytic expression from Eq. (3.45). The longitudinal plasma density is shown by the shaded region.

3.4.2 Evolution of the Beam Moments and Emittance

The transverse beam quality can be quantified using the normalized emittance [52]:

$$\epsilon_n = (1/m_e c) \sqrt{\sigma_x^2 \sigma_{p_x}^2 - \sigma_{xp_x}^2}, \quad (3.46)$$

where the beam sizes are $\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2$, $\sigma_{p_x}^2 = \langle p_x^2 \rangle - \langle p_x \rangle^2$ and the correlation is $\sigma_{xp_x} = \langle xp_x \rangle - \langle x \rangle \langle p_x \rangle$. In the ultrarelativistic limit the normalized emittance simplifies to

$$\epsilon_n^2 = \langle \gamma_b \rangle^2 (\sigma_\delta^2 \sigma_x^2 \langle x'^2 \rangle + \epsilon^2), \quad (3.47)$$

where $\sigma_\delta^2 = (\langle \gamma_b^2 \rangle - \langle \gamma_b \rangle^2) / \langle \gamma_b \rangle^2$ and ϵ is the geometric emittance defined as

$$\epsilon^2 = \sigma_x^2 \sigma_{x'}^2 - \sigma_{xx'}^2 \quad (3.48)$$

with $\sigma_{x'}^2 = \langle x'^2 \rangle - \langle x' \rangle^2$ and $\sigma_{xx'} = \langle xx' \rangle - \langle x \rangle \langle x' \rangle$. In most cases, the first term in Eq. (3.47) is very small and the geometric emittance dominates, giving the more familiar formula $\epsilon_n \approx \langle \gamma_b \rangle \epsilon$.

Emittance growth results from the γ_b dependence of k_β . Different energy slices of the beam oscillate with different frequencies in the ion channel. Over time the different slices dephase and

no longer overlap in transverse phase space as shown in Fig. 3.13. Even if the beam has no initial energy spread, this chromatic dephasing will occur if imperfect wake loading induces energy spread in the beam during acceleration. This same effect dampens out the centroid oscillations of an offset beam. To calculate the growth explicitly requires the evaluation of the moments of the beam distribution.

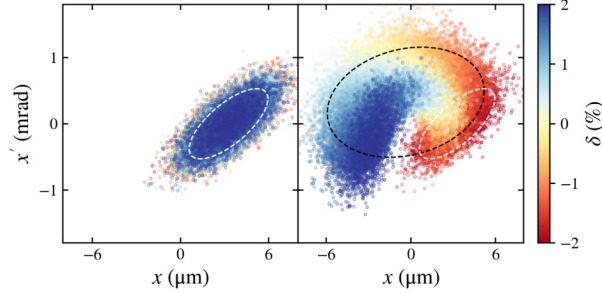


Figure 3.13: When the witness beam is mismatched or offset transversely with respect to the ion column, it will undergo emittance growth through chromatic dephasing. Different energy slices of the beam will rotate at different frequencies in transverse phase space, leading to growth in the projected emittance (the area enclosed by the white dashed line grows to the area enclosed by the black dashed line).

The beam moments can be evaluated under the assumption that there is no initial correlation between the longitudinal distribution and the transverse distribution. In this case the phase space distribution of the witness beam at s_0 can be written as $f_0(x, x', \xi, \delta) = f_{\perp}(x, x')f(\xi, \delta)$, where $\delta = (p - p_0)/p_0$ parameterizes the energy spread in the beam and p_0 is the momentum of the reference particle. Here, we are ignoring the vertical transverse dimension because it is decoupled and evolves independently in an analogous fashion. We let $f_{\perp}$ describe the witness beam without offset; i.e., the beam centroid is located at the origin in transverse phase space: $\int dx dx' x f_{\perp}(x, x') = 0$ and $\int dx dx' x' f_{\perp}(x, x') = 0$. Further, we assume that the distribution is normalized to 1.

If the witness beam is then given an initial offset of Δx in x and $\Delta x'$ in x' , the beam distribution at s_0 can be written as $f_{\Delta}(x, x', \xi, \delta) = f_{\perp}(x - \Delta x, x' - \Delta x')f(\xi, \delta)$. Single particle evolution is described by Eq. (3.44) and is a function of $s, x_0, x'_0, \xi_0, \delta_0$: $x(s, x_0, x'_0, \xi_0, \delta_0)$, where the ξ_0 and δ_0 dependence enters through the relativistic factor $\gamma_b(s, \xi_0, \delta_0)$. The functional form of

γ_b depends on the acceleration model chosen, we show several examples in Secs. 3.4.5 and 3.4.6. Making the change of variables to $u = x - \Delta x$ and $u' = x' - \Delta x'$, we can write the single particle evolution in terms of u and u' : $x(s, u_0 + \Delta x, u'_0 + \Delta x', \xi_0, \delta_0)$. The first beam moment is given by:

$$\langle x \rangle = \int d\xi_0 d\delta_0 f(\xi_0, \delta_0) \int du_0 du'_0 x(u_0, u'_0) f_\perp(u_0, u'_0). \quad (3.49)$$

The other beam moments, $\langle x^2 \rangle$, $\langle x' \rangle$, $\langle x'^2 \rangle$, and $\langle xx' \rangle$ take the same form.

To make the integral tractable, we make the assumption that the energy spread is small: $\delta \ll 1$. This allows us to treat the matched CS parameters and the adiabatic dampening term, $A = \sqrt{\gamma_{b0}/\gamma_b}$, in Eq. (3.45) as constant with respect to ξ_0 and δ_0 . We use bars to denote values for the reference particle: $\bar{\gamma}_b = \gamma_b(s, \xi_0 = 0, \delta_0 = 0)$. Using $\gamma_b = \bar{\gamma}_b$ in A , β_m , α_m , and γ_m for all particles defines a matched set of CS parameters for the beam; this is in contrast to the approach taken in Sec. 3.2 (Ref. [57]), where different matched CS parameters are defined for each energy slice. The only dependence on ξ_0 and δ_0 remaining in Eq. (3.45) is in the energy dependence of the betatron phase ϕ .

We have to integrate over u and u' in Eq. (3.49) first because x depends on the longitudinal coordinates. The u and u' integrals can be written in terms of moments of the witness beam's initial (non-offset) phase space distribution. The second central moments of the phase space distribution, σ_x , $\sigma_{x'}$, and $\sigma_{xx'}$, are defined as: $\sigma_x^2 = \int dx dx' x^2 f_\perp(x, x')$, $\sigma_{x'}^2 = \int dx dx' x'^2 f_\perp(x, x')$, and $\sigma_{xx'} = \int dx dx' xx' f_\perp(x, x')$. These moments can be expressed in terms of the CS parameters: $\beta = \sigma_x^2/\epsilon$, $\gamma = \sigma_{x'}^2/\epsilon$, and $\alpha = -\sigma_{xx'}/\epsilon$.

Evaluating the u , u' integrals, and writing the beam moments in terms of the beam's initial CS parameters gives the following expressions for the moments of the offset beam:

$$\langle x \rangle = \sqrt{\frac{\bar{\gamma}_{b0}}{\bar{\gamma}_b}} \beta_m \epsilon_0 (d_1 C_1 + d_2 S_1) \quad (3.50)$$

$$\langle x^2 \rangle = \frac{\bar{\gamma}_{b0}}{\bar{\gamma}_b} \beta_m \epsilon_0 (a + b_1 C_2 + b_2 S_2) \quad (3.51)$$

$$\langle x' \rangle = \sqrt{\frac{\bar{\gamma}_{b0}}{\bar{\gamma}_b} \frac{\epsilon_0}{\beta_m}} [-d_1 S_1 + d_2 C_1 - \alpha_m (d_1 C_1 + d_2 S_1)] \quad (3.52)$$

$$\begin{aligned} \langle x'^2 \rangle = & \frac{\bar{\gamma}_{b0}}{\bar{\gamma}_b} \frac{\epsilon_0}{\beta_m} [(1 + \alpha_m^2)a - (1 - \alpha_m^2)(b_1 C_2 + b_2 S_2) \\ & + 2\alpha_m(b_1 S_2 - b_2 C_2)] \end{aligned} \quad (3.53)$$

$$\langle xx' \rangle = \frac{\bar{\gamma}_{b0}}{\bar{\gamma}_b} \epsilon_0 [b_2 C_2 - b_1 S_2 - \alpha_m(a + b_1 C_2 + b_2 S_2)], \quad (3.54)$$

where we have kept terms of order α_m^2 . The a , b , and d terms are all unit-less constants that depend on the beam's initial conditions:

$$\begin{aligned} d_1 &= \frac{\Delta x}{\sqrt{\beta_{m0}\epsilon_0}} \\ d_2 &= \alpha_{m0} \frac{\Delta x}{\sqrt{\beta_{m0}\epsilon_0}} + \sqrt{\frac{\beta_{m0}}{\epsilon_0}} \Delta x' \\ a &= \frac{1}{2} (\beta_0 \gamma_{m0} + \gamma_0 \beta_{m0} - 2\alpha_0 \alpha_{m0} + d_1^2 + d_2^2) \\ b_1 &= \frac{\beta_0}{\beta_{m0}} + d_1^2 - a \\ b_2 &= d_1 d_2 - \alpha_0 + \alpha_{m0} \frac{\beta_0}{\beta_{m0}}. \end{aligned}$$

The first three constants have physical meaning; d_1 and d_2 are the normalized phase space coordinates for the motion of the beam centroid and, as we will show later, a is the relative emittance growth at saturation.

The s dependence of the moments enters through $\bar{\gamma}_b$ directly (adiabatic dampening) and through the matched CS parameters (plasma focusing), which are functions of $\bar{\gamma}_b$; however, the primary dependence on s is given by the C and S terms which capture both the chromatic phase spreading and the betatron oscillations. These terms are integrals over the longitudinal phase space:

$$\begin{aligned} C_1 &= \int d\xi_0 d\delta_0 f(\xi_0, \delta_0) \cos[\phi(\xi_0, \delta_0)] \\ S_1 &= \int d\xi_0 d\delta_0 f(\xi_0, \delta_0) \sin[\phi(\xi_0, \delta_0)] \\ C_2 &= \int d\xi_0 d\delta_0 f(\xi_0, \delta_0) \cos[2\phi(\xi_0, \delta_0)] \\ S_2 &= \int d\xi_0 d\delta_0 f(\xi_0, \delta_0) \sin[2\phi(\xi_0, \delta_0)]. \end{aligned}$$

The two dimensional longitudinal phase space distribution can always be reduced to a one dimensional distribution of betatron phase advance $f_\phi(\phi)$. The integrals can then be written in the form:

$$C_1 = \int d\phi f_\phi(\phi) \cos \phi, \quad (3.55)$$

where we have only written the C_1 integral for brevity. The s dependence enters through the variation of the f_ϕ distribution as the beam propagates through the plasma. At this point, the saturated emittance can be determined. As the different energy components of the beam dephase, the width of the f_ϕ distribution will grow. At saturation, the width is much larger than 2π and the C and S integrals tend towards zero. Setting C and S to zero in the moments and calculating the emittance gives the saturated emittance

$$\epsilon_{nsat} = \epsilon_{n0} \sqrt{1 + \sigma_\delta^2} \frac{1}{2} \left(\beta_0 \gamma_{m0} + \gamma_0 \beta_{m0} - 2\alpha_0 \alpha_{m0} + \frac{\Delta x^2}{\beta_{m0} \epsilon_0} + \Delta x'^2 \frac{\beta_{m0}}{\epsilon_0} + \Delta x \Delta x' \frac{\alpha_{m0}}{\epsilon_0} \right), \quad (3.56)$$

where terms higher than α_m have been dropped. The pre-factor $\sqrt{1 + \sigma_\delta^2}$ can be set equal to one if the final energy spread is sufficiently small.

The detailed emittance evolution can be found by rewriting the C and S integrals. We start by writing the betatron phase as $\phi = \bar{\phi} + \Delta\phi(\xi_0, \delta_0)$. Further, we note that the four C and S integrals can be written as the real and imaginary parts of two complex integrals. After removing the $e^{i\bar{\phi}}$ term from the them, the integrals generically evaluate to a pair of complex numbers with real amplitudes H_1, H_2 , and arguments ψ_1, ψ_2 :

$$I_1 = \int d\xi_0 d\delta_0 f(\xi_0, \delta_0) e^{i\Delta\phi(\xi_0, \delta_0)} = H_1 e^{i\psi_1}, \quad (3.57)$$

$$I_2 = \int d\xi_0 d\delta_0 f(\xi_0, \delta_0) e^{i2\Delta\phi(\xi_0, \delta_0)} = H_2 e^{i\psi_2}. \quad (3.58)$$

The C and S integrals are then given by

$$C_1 = H_1 \cos(\bar{\phi} + \psi_1)$$

$$S_1 = H_1 \sin(\bar{\phi} + \psi_1)$$

$$C_2 = H_2 \cos(2\bar{\phi} + \psi_2)$$

$$S_2 = H_2 \sin(2\bar{\phi} + \psi_2).$$

If one of three conditions is met or approximately met, the expression for the emittance simplifies significantly. The first condition is if $\psi_1 = \psi_2 = 0$. The second is if $\psi_1 = \psi_2 = \pi/2$. The third is if $\psi_2 = 2\psi_1$. In all the examples we consider later, either the first or third condition is satisfied. The emittance then evolves according to

$$\epsilon_n = \epsilon_{nsat} \sqrt{1 - \frac{b_1^2 + b_2^2}{a^2} H_2^2 - \frac{1}{a^2} \{a(d_1^2 + d_2^2) - [b_1(d_1^2 - d_2^2) + 2b_2d_1d_2] H_2\} H_1^2}, \quad (3.59)$$

where H_2 terms describe emittance growth due to mismatch and H_1 terms describe emittance growth due to transverse offsets in position and angle. In addition, H_1 , when combined with the adiabatic dampening, describes the dampening of the beams centroid oscillations. The $H(s)$ functions describe the amount of chromatic coherence in the beam, they get smaller as the beam propagates and dephases. ψ_1 is the betatron phase difference between the projected beam and the reference particle.

3.4.3 Solution in a Uniform Plasma

In Fig. 3.14 we compare the results of numerical particle tracking with Eq. (3.59) for a uniform plasma without energy gain. The witness beam has an initial energy spread and is either mismatch to the plasma (red), offset transversely from the drive beam (blue), or both mismatched and offset (black). In this simple case, Eq. (3.41) gives the betatron phase advance as

$$\phi = \bar{\phi} \frac{1}{\sqrt{1 + \delta_0}}.$$

Here, ϕ is independent of ξ_0 because the accelerating field is constant (in this case zero) in ξ_0 . The betatron phase advance of the reference particle is $\bar{\phi} = \bar{k}_\beta s$, where $\bar{k}_\beta = \omega_p / (c\sqrt{2\gamma_b})$ is the betatron wavenumber of the reference particle. Expanding to first order in δ_0 simplifies the expression to $\phi = \bar{\phi} - \bar{\phi}\delta_0/2$. Since ϕ depends linearly on δ_0 , the I_1 and I_2 integrals are Fourier transforms of the energy distribution with frequencies $\omega_\delta = \bar{\phi}/2$ and $2\omega_\delta = \bar{\phi}$, respectively. If the distribution of energy is even (symmetric about γ_b), then $H_1 = \hat{f}_\delta(\omega_\delta) = \hat{f}_\delta(\bar{\phi}/2)$, $H_2 = \hat{f}_\delta(\bar{\phi})$, and $\psi_1 = \psi_2 = 0$. Here, $\hat{f}_\delta$ is the Fourier transform of the energy distribution f_δ . Eq. (3.59) then fully describes the evolution of the projected emittance.

In the case of a uniform energy spread between $-\Delta\delta/2$ and $\Delta\delta/2$, $\hat{f}_\delta(\bar{\phi}) = \text{sinc}\Delta\Phi$ and $\hat{f}_\delta(\bar{\phi}/2) = \text{sinc}(\Delta\Phi/2)$, where $\Delta\Phi$ is the range of ϕ spanned by the energy spread $\Delta\Phi = \Delta\delta\bar{\phi}$. In the absence of transverse offset, this formula reduces to that given in Ref. [18]. The emittance growth of such a beam is shown in Fig. 3.14(a) for various initial mismatches and transverse offsets. The analytic expression shows excellent agreement with the numerical particle tracking code. The small oscillations of the numerical result about the theoretical emittance from Eq. (3.59) are due to treating β_m and γ_b as constants and are of order σ_δ^2 (for more details see Ref. [57]).

There are two length scales for the emittance growth. The growth due to mismatch saturates when $\Delta\Phi = \pi$ giving a saturation length of $s_2 = \pi/(\Delta\delta\bar{k}_\beta)$, while the emittance growth due to transverse offset requires twice the length to saturate: $s_1 = 2\pi/(\Delta\delta\bar{k}_\beta)$.

In the case of a Gaussian distribution of energy, $f_\delta(\delta_0) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\delta_0^2}{2\sigma_\delta^2}\right)$ and $\hat{f}_\delta(\bar{\phi}) = \exp\left(-\frac{\bar{\phi}^2\sigma_\delta^2}{2}\right)$. Without any transverse offset, $\Delta x = \Delta x' = 0$, the solution reduces to that given in Ref. [19]. The theoretical solution is compared to numerical particle tracking in Fig. 3.14(b). As can be seen in the figure, the saturated emittance is the sum of the emittance growth due to an offset and mismatch.

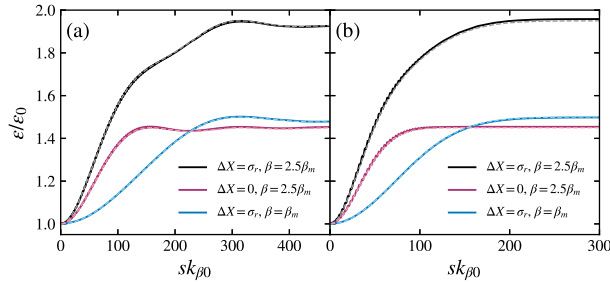


Figure 3.14: Emittance growth of a beam in a uniform plasma without energy gain. The beam is transversely offset in the blue and black curves. The beam is mismatched in the red and black curves. (a) shows the emittance evolution for a beam with uniform energy spread, (b) shows the evolution for a beam with Gaussian energy spread. The dashed lines represent the analytic theory and solid lines the numerical particle tracking results. In all cases, the energy spread is 2%. The mismatch driven emittance growth saturates in half the distance of the offset driven emittance growth. The final saturated emittance growth is the sum of the mismatch and offset.

3.4.4 Evolution of the Slice Moments and Slice Emittance

Although the projected emittance is normally taken as the figure of merit, both the longitudinal slice emittance and the energy slice emittance are of practical interest. We start with the longitudinal slice emittance, which is of importance for both light sources and colliders. For an individual slice at ξ , the I_1 integral takes the form:

$$\begin{aligned} I_1(\xi) &= \frac{1}{N(\xi)} \int d\xi_0 d\delta_0 f(\xi, \delta_0) \delta_D(\xi_0 - \xi) e^{i\Delta\phi(\xi_0, \delta_0)} \\ I_1(\xi) &= \frac{1}{N(\xi)} \int d\delta_0 f(\xi, \delta_0) e^{i\Delta\phi(\xi, \delta_0)} = H_1(\xi) e^{i\psi_1(\xi)}, \end{aligned} \quad (3.60)$$

where $N(\xi) = \int d\delta_0 f(\xi, \delta_0)$ is a normalization factor and δ_D is the Dirac delta function. I_2 has the same form with $\Delta\phi$ replaced by $2\Delta\phi$. The expressions for the moments and emittance of each slice are given by Eqs. (3.50)-(3.54) and Eq. (3.59), but with H and ψ replaced by the ξ dependent expressions from Eq. (3.60). Further accuracy can be achieved by using $\bar{\gamma}_b$ and β_m calculated for each slice rather than the beam as a whole.

Similarly, the energy slice emittance is found by replacing the I_1 and I_2 integrals with

$$I_1(\delta) = H_1(\delta) e^{i\psi_1(\delta)} = \frac{1}{N(\delta)} \int d\xi_0 d\delta_0 f(\xi_0, \delta_0) \delta_D(\delta - \frac{\gamma_b}{\bar{\gamma}_b} + 1) e^{i\Delta\phi(\xi_0, \delta_0)}, \quad (3.61)$$

where γ_b is a function of ξ_0 and δ_0 , $N(\delta) = \int d\xi_0 d\delta_0 f(\xi_0, \delta_0) \delta_D(\delta - \frac{\gamma_b}{\bar{\gamma}_b} + 1)$. Again, I_2 has the same form as I_1 with $\Delta\phi$ replaced by $2\Delta\phi$. Growth in the energy slice emittance is due to particle exchange between energy slices as a consequence of imperfect loading of the wake.

Energy slice emittance is important for two reasons: First, projected emittance growth can be reversed down to the energy slice emittance using a suitable achromatic beam line. An example of this technique is presented in Ref. [61]. Second, the energy slice emittance influences the results of some emittance measurements, such as those taken using the butterfly technique. Third, the performance of a transverse gradient undulator depends on the energy slice emittance [66, 67, 68].

3.4.5 Emittance Growth of a Beam with Imperfect Wake Loading

In the previous sections, we develop general expressions for the projected and slice emittance growth. In this section, we work out analytic solutions for the projected and slice emittance for

two situations where the witness beam does not perfectly load the wake.

Perfect loading requires the witness beam to have a trapezoidal current profile to produce a uniform accelerating field [49]. Existing accelerators and PWFA experiments tend to use beams that have Gaussian or other non-trapezoidal current profiles. Depending on the length and current of the witness beam, the accelerating field can be approximated around the beam centroid as a linear or quadratic function of ξ . A linear function is a good approximation when the witness beam does not have sufficient current to flatten the wake and a quadratic function is a good approximation if the witness beam has too much current. To demonstrate this, we ran a series of particle in cell (PIC) simulations using the code VSim [56]. A 0.5 nC witness beam and 2 nC drive beam were propagated through a $3.5 \times 10^{16} \text{ cm}^{-3}$ plasma. Both beams had Gaussian current profiles. Simulations were run for witness beams of different lengths, and thus, different peak currents. Fig. 3.15 shows the longitudinal electric field in the region of the witness beam for two beams with the same charge, but bunch lengths of $\sigma_z = 4.0 \mu\text{m}$ and $\sigma_z = 8.0 \mu\text{m}$. The electric field around the shorter beam is well represented by a quadratic function centered on the beam centroid while the field around the longer beam can be approximated as linear. We proceed to work out the projected emittance, longitudinal slice emittance, and energy slice emittance growth for both cases.

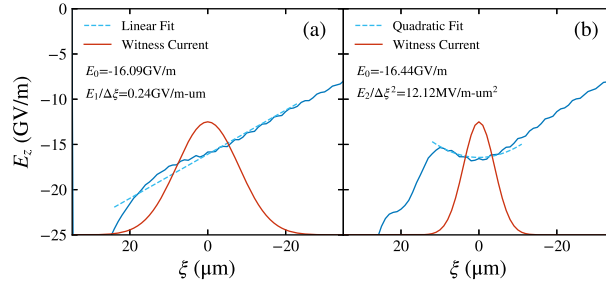


Figure 3.15: Longitudinal variation in the accelerating field for two witness beams of different lengths. Both beams have the same charge. The beam in (a) is too long and does not sufficiently load the wake, the accelerating field varies approximately linearly along the beam as shown by the dashed blue line. (b) shows a beam that is slightly too short, the variation in the accelerating field is well approximated by a quadratic function. In both cases the field from $-3\sigma_{\xi 0}$ to $3\sigma_{\xi 0}$ was used for fitting.

In the linear case, the accelerating field is given by $E_z = -E_0 - E_1\xi/\Delta\xi$. Here, E_0 is the accelerating field at $\xi = 0$ and $E_1/\Delta\xi$ describes the slope of the electric field; $\Delta\xi$ is an arbitrary length scale, we typically set it to the bunch length. The energy of a particle in the beam as a function of s depends on ξ_0 and δ_0 according to

$$\gamma_b = \bar{\gamma}_b + \delta_0\gamma_{b0} + E_1s\frac{\xi_0}{\Delta\xi}\frac{e}{m_ec^2}, \quad (3.62)$$

where $\bar{\gamma}_b = \gamma_{b0} + E_0se/(m_ec^2)$. It is straightforward to evaluate Eq. (3.41) to get the betatron phase advance of each particle

$$\phi = \frac{\omega_p}{c\sqrt{2}} \frac{2m_ec^2}{E_0e} \frac{1}{1 + E_1\xi_0/(E_0\Delta\xi)} \left(\sqrt{\gamma_b} - \sqrt{\gamma_{b0}(1 + \delta_0)} \right).$$

If we make the reasonable assumption that $E_1\xi_0/(E_0\Delta\xi) \ll 1$, then ϕ can be expanded in the aforementioned quantity giving

$$\phi = \frac{\omega_p}{c\sqrt{2}} \frac{2m_ec^2}{E_0e} \left(1 - \frac{E_1\xi_0}{E_0\Delta\xi} \right) \left(\sqrt{\gamma_b} - \sqrt{\gamma_{b0}(1 + \delta_0)} \right).$$

Even for simple distributions, the C and S integrals are not closed form without further expansion of ϕ . We use the fact that $\gamma_b \approx \bar{\gamma}_b$ to expand the square root in the above expression in $E_1\xi_0/(E_0\Delta\xi)$ and δ_0 :

$$\phi = \bar{\phi} - \delta_0 \frac{\bar{\phi}}{2} \sqrt{\frac{\gamma_{b0}}{\bar{\gamma}_b}} + \frac{E_1\xi_0}{E_0\Delta\xi} (\bar{k}_\beta s - \bar{\phi}), \quad (3.63)$$

where we have dropped terms higher than first order. The betatron phase advance of the reference particle is

$$\bar{\phi} = \frac{\omega_p}{c\sqrt{2}} \frac{2m_ec^2}{E_0e} (\sqrt{\bar{\gamma}_b} - \sqrt{\gamma_{b0}}). \quad (3.64)$$

Because ϕ depends linearly on ξ_0 and δ_0 as $\Delta\phi = -\delta_0\omega_\delta - \xi_0\omega_\xi$, the I integrals are related to the 2D Fourier transform of the distribution:

$$I_1 = \hat{f}(\omega_\xi, \omega_\delta), \quad I_2 = \hat{f}(2\omega_\xi, 2\omega_\delta), \quad (3.65)$$

where

$$\begin{aligned} \omega_\xi &= -\frac{E_1}{E_0\Delta\xi} (\bar{k}_\beta s - \bar{\phi}), \\ \omega_\delta &= \frac{\bar{\phi}}{2} \sqrt{\frac{\gamma_{b0}}{\bar{\gamma}_b}}. \end{aligned} \quad (3.66)$$

Consider a beam with a Gaussian longitudinal distribution with length $\sigma_{\xi 0}$ and energy spread $\sigma_{\delta 0}$,

$$f(\xi_0, \delta_0) = \frac{1}{2\pi\sigma_{\xi 0}\sigma_{\delta 0}} \exp\left(-\frac{\xi_0^2}{2\sigma_{\xi 0}^2} - \frac{\delta_0^2}{2\sigma_{\delta 0}^2}\right). \quad (3.67)$$

The distribution is symmetric so $\hat{f}$ is real and $\psi_1 = \psi_2 = 0$. H_1 and H_2 are given by

$$\begin{aligned} H_1 &= e^{-\sigma_{\xi 0}^2 \omega_{\xi}^2 / 2 - \sigma_{\delta 0}^2 \omega_{\delta}^2 / 2}, \\ H_2 &= e^{-2\sigma_{\xi 0}^2 \omega_{\xi}^2 - 2\sigma_{\delta 0}^2 \omega_{\delta}^2}, \end{aligned} \quad (3.68)$$

and the projected emittance and moment evolution is given by Eq. (3.59) and Eqs. (3.50)-(3.54), respectively.

The longitudinal slice emittance of a Gaussian beam is straightforward to calculate. Evaluating Eq. (3.60) gives $H_1 = e^{-\sigma_{\delta 0}^2 \omega_{\delta}^2 / 2}$ and $\psi_1 = -\omega_{\xi} \xi$. H_2 and ψ_2 are given by replacing ω_{δ} with $2\omega_{\delta}$ and ω_{ξ} with $2\omega_{\xi}$ and the emittance growth is given by Eq. (3.59). The longitudinal slice emittance has no dependence on ξ because all the slices have the same initial energy spread. The phase term ψ_1 shows that the beam offset will vary sinusoidally along the bunch as shown in the top of Fig. 3.16 because $\psi_1 \propto \xi_0$. The oscillation frequency is given by ω_{ξ} .

When finding the energy slice emittance, it will be convenient to use the RMS energy spread:

$$\sigma_{\delta}^2 = \sigma_{\delta 0}^2 \frac{\gamma_{b0}^2}{\bar{\gamma}_b^2} + \sigma_{\xi 0}^2 \frac{s^2 E_1^2 e^2}{\bar{\gamma}_b^2 \Delta \xi^2 m_e^2 c^4}. \quad (3.69)$$

Using Eq. (3.61), we calculate H_1 and ψ_1 for each energy slice:

$$\begin{aligned} H_1 &= \exp\left[-\frac{\sigma_{\xi 0}^2 \sigma_{\delta 0}^2}{2\sigma_{\delta}^2} \left(\frac{sE_1 e}{\bar{\gamma}_b \Delta \xi m_e c^2} \omega_{\delta} - \frac{\gamma_{b0}}{\bar{\gamma}_b} \omega_{\xi}\right)^2\right] \\ \psi_1 &= \delta \frac{1}{\sigma_{\delta}^2} \left(\frac{\gamma_{b0}}{\bar{\gamma}_b} \sigma_{\delta 0}^2 \omega_{\delta} + \frac{sE_1 e}{\bar{\gamma}_b \Delta \xi m_e c^2} \sigma_{\xi 0}^2 \omega_{\xi}\right). \end{aligned} \quad (3.70)$$

H_2 and ψ_2 are given by replacing ω_{δ} with $2\omega_{\delta}$ and ω_{ξ} with $2\omega_{\xi}$. The emittance growth is given by Eq. (3.59) and the moments are given by Eqs. (3.50)-(3.54).

As with the longitudinal slice emittance, the energy slice emittance is the same for all slices because H_1 is independent of δ and the betatron phase offset is linearly proportional to the slice energy because ψ_1 is proportional to δ . Chromatic dephasing cannot occur within an energy slice

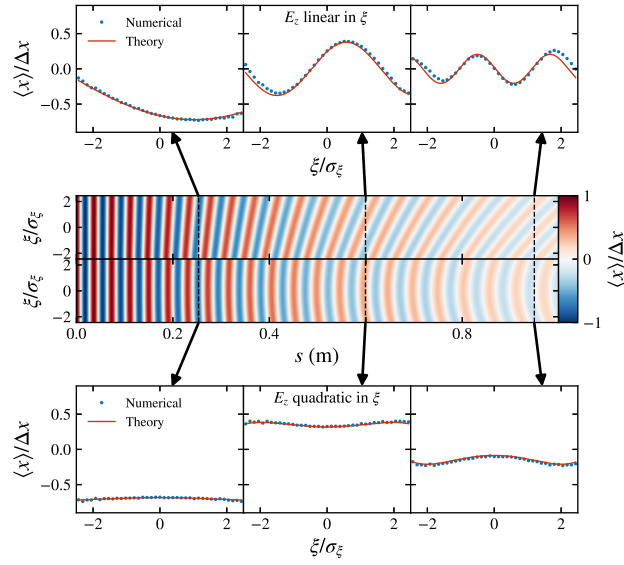


Figure 3.16: Centroid oscillations for each longitudinal slice in the beam. The top of the figure shows the evolution for the longer beam shown in Fig. 3.15(a), and the bottom shows the shorter beam from Fig. 3.15(b). In both cases the initial energy spread is $\sigma_{\delta 0} = 4\%$. The magnitude of the oscillations is dampened by chromatic phase spreading due to the initial energy spread within each slice. The oscillations decohere due to the energy differences the slices pick up as a result of imperfect wake loading. In (a) the accelerating field varies significantly across the slices and the oscillations rapidly decohere. In (b) the accelerating field along the beam is more uniform and the beam oscillates nearly coherently.

unless particles are able to mix between slices. The variation in accelerating field along the wake causes this mixing to occur.

There are several emittance saturation length scales. The transverse offset and mismatch driven emittance growth are determined by H_1 and H_2 respectively. We define the saturation length as the distance s when the argument of the exponential in H equals -1. The saturation length for the longitudinal slice emittance growth due to mismatch is given by $2\sigma_{\delta 0}^2 \omega_\delta^2 = 1$; this length also quantifies the contribution of the initial energy spread to the emittance growth. The length scale for the emittance growth due to imperfect wake loading is $2\sigma_{\xi 0}^2 \omega_\xi^2 = 1$. Finally, the

saturation length for the energy slice emittance growth is given by

$$\frac{2\sigma_{\xi 0}^2\sigma_{\delta 0}^2}{\sigma_{\delta}^2}\left(\frac{sE_1e}{\bar{\gamma}_b\Delta\xi m_e c^2}\omega_{\delta}-\frac{\gamma_{b0}}{\bar{\gamma}_b}\omega_{\xi}\right)^2=1.$$

The solution for s for each of these expressions is summarized in Table 3.2. To calculate the saturation length for the energy slice emittance, we assumed $\sigma_{\delta 0}^2\gamma_{b0}^2 \ll \sigma_{\xi 0}^2 s^2 E_1^2 e^2 / (\Delta\xi^2 m_e^2 c^4)$ (i.e. the initial energy spread is smaller than the energy spread induced by the wake), to simplify the expression for σ_{δ} . This assumption is typically well satisfied for a plasma-based accelerator with a single stage of reasonable length.

The saturation lengths reveal several properties about the emittance growth. First, the energy slice emittance always has a saturation length longer than that of the longitudinal slice emittance. Consequently, the energy slice emittance is smaller than the longitudinal slice emittance at every s . Second, both slice emittances are approximately independent of $\sigma_{\xi 0}$ and E_1 ; they depend only on the initial energy spread $\sigma_{\delta 0}$. Third, if the initial energy spread is too low, the slice emittances will never fully saturate because acceleration reduces the relative energy spread faster than dephasing can occur. Mathematically, this appears as a divergence in the saturation lengths of the slice emittances. Saturation occurs because the spread in betatron phase becomes larger than 2π ; the phase spread in a longitudinal slice is $\Delta\phi = \sigma_{\delta 0}\omega_{\delta}$, which does not grow without bound. The maximum phase spread is found by taking the large s limit:

$$\lim_{s \rightarrow \infty} \sigma_{\delta 0}\omega_{\delta}(s) = \sigma_{\delta 0}\gamma_{b0}k_{\beta 0}\frac{m_e c^2}{E_0 e}. \quad (3.71)$$

The maximum emittance is found by inserting this limit into the expression for H_1 and H_2 . This also means the magnitude of the transverse oscillations of the longitudinal slices are not dampened to zero. The above conclusions apply to the energy slice emittance as long as E_1 is sufficiently large ($\sigma_{\delta 0}^2\gamma_{b0}^2 \ll \sigma_{\xi 0}^2 s^2 E_1^2 e^2 / (\Delta\xi^2 m_e^2 c^4)$).

Figure 3.17 compares the growth of the longitudinal slice, energy slice, and projected emittance. The accelerating field is the same as that shown in Fig. 3.15(a), but the beam length has been doubled to $\sigma_{\xi 0} = 16.0 \mu\text{m}$ to exaggerate the difference between the projected and slice emittance saturation lengths. The initial energy spread is $\sigma_{\delta 0} = 0.05$. The saturation length due to the

Table 3.2: Saturation lengths for the projected and slice emittances. The projected emittance growth is driven by contributions from both the initial energy spread and the energy spread induced by imperfect wake loading. The shorter of the two saturation lengths dominates the projected emittance growth. In contrast, both energy and longitudinal slice emittance growth depend only on the initial energy spread, assuming the variation in accelerating field is sufficiently strong.

	Saturation Length - Offset Beam	Saturation Length - Mismatched Beam
Imperfect Loading Contribution ^a	$s_{1\xi} = \frac{E_0 e}{\gamma_{b0} m_e c^2} L + \sqrt{\left(\frac{E_0 e}{\gamma_{b0} m_e c^2} L\right)^2} + 2L$	$s_{2\xi} = \frac{E_0 e}{2\gamma_{b0} m_e c^2} L + \sqrt{\left(\frac{E_0 e}{2\gamma_{b0} m_e c^2} L\right)^2} + L$
Longitudinal Slice Emittance (Initial Energy Spread Contribution)	$s_{1\delta 0} = 4m_e c^2 \gamma_{b0} \frac{m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - e E_0}{(m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - 2e E_0)^2}$	$s_{2\delta 0} = 4m_e c^2 \gamma_{b0} \frac{2m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - e E_0}{(2m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - 2e E_0)^2}$
Energy Slice Emittance ^b	$s_{1\delta} \approx 8m_e c^2 \gamma_{b0} \frac{m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}}}{(m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - 2e E_0)^2}$	$s_{2\delta} \approx 8m_e c^2 \gamma_{b0} \frac{2m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}}}{(2m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - 2e E_0)^2}$
Accelerating field quadratic in ξ		
Imperfect Loading Contribution ^c	$s_{1\xi} = \frac{E_0 e}{\gamma_{b0} m_e c^2} Q + \sqrt{\left(\frac{E_0 e}{\gamma_{b0} m_e c^2} Q\right)^2} + 2Q$	$s_{2\xi} = \frac{E_0 e}{2\gamma_{b0} m_e c^2} Q + \sqrt{\left(\frac{E_0 e}{2\gamma_{b0} m_e c^2} Q\right)^2} + Q$
Longitudinal Slice Emittance (Initial Energy Spread Contribution)	$s_{1\delta 0} = 4m_e c^2 \gamma_{b0} \frac{m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - e E_0}{(m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - 2e E_0)^2}$	$s_{2\delta 0} = 4m_e c^2 \gamma_{b0} \frac{2m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - e E_0}{(2m_e c \sigma_{\delta 0} \omega_p \sqrt{\gamma_{b0}} - 2e E_0)^2}$

^a $L = \frac{E_0^2 \Delta \xi^2}{E_1^2 \sigma_{\xi 0}^2 k_{\beta 0}^2} + 2\sqrt{2} \frac{m_e c^2 \gamma_{b0} \Delta \xi}{k_{\beta 0} E_1 e \sigma_{\xi 0}}$. ^b Assuming E_1 is sufficiently large $\sigma_{\delta 0}^2 \gamma_{b0}^2 \ll \sigma_{\xi 0}^2 s^2 E_1^2 c^2 / (\Delta \xi^2 m_e c^4)$.

^c $Q = 4 \frac{E_0^2 \Delta \xi^4}{E_1^2 \sigma_{\xi 0}^4 k_{\beta 0}^2} + 8 \frac{m_e c^2 \gamma_{b0} \Delta \xi^2}{k_{\beta 0} E_1 e \sigma_{\xi 0}^2}$

variation in the accelerating field is shorter than that due to the beam's initial energy spread. As a result, the saturation length due to imperfect wake loading adequately describes the projected emittance growth. The difference between the numerical and analytic solution for the projected emittance arises due to the large final energy spread in this example. This growth can be handled analytically using the approach presented in Ref. [57]; however, for typical experimental parameters the energy spread is small enough that the correction is negligible.

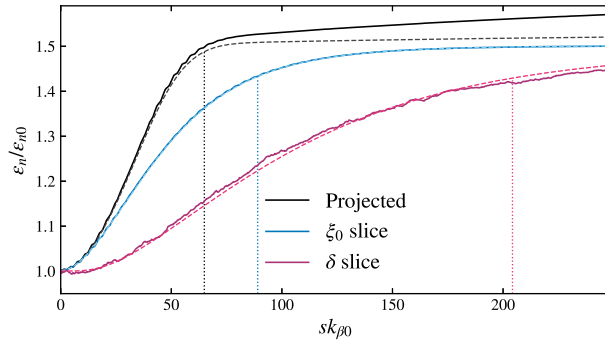


Figure 3.17: Growth of the projected, longitudinal slice, and energy slice emittance of the witness beam in a wake where the accelerating field varies linearly in ξ . The solid lines are the emittance from numerical particle tracking, the dashed lines are the analytic theory. The dotted lines show the saturation lengths from Table 3.2: black is the contribution from the non-uniform accelerating field to the projected emittance, blue is the initial energy spread contribution as well as the longitudinal slice, and magenta is the energy slice. The projected emittance is dominated by contribution from the non-uniform accelerating field.

For the quadratic case, the accelerating field is given by $E_z = -E_0 - E_2\xi^2/\Delta\xi^2$, where $E_2/\Delta\xi^2$ is the quadratic fitting parameter. As before, $\Delta\xi$ is an arbitrary length scale typically set to the bunch length. The energy of a particle in the beam is

$$\gamma_b = \bar{\gamma}_b + \delta_0\gamma_{b0} + E_2s\frac{\xi_0^2}{\Delta\xi^2}\frac{e}{m_e c^2}, \quad (3.72)$$

where $\bar{\gamma}_b\gamma_{b0} + E_0se/(m_e c^2)$, the same as before. Following the same procedure as before, the betatron phase advance of the particle is given by

$$\phi = \bar{\phi} - \delta_0\frac{\bar{\phi}}{2}\sqrt{\frac{\gamma_{b0}}{\bar{\gamma}_b}} + \frac{E_2\xi_0^2}{E_0\Delta\xi^2}(\bar{k}_\beta s - \bar{\phi}). \quad (3.73)$$

The reference phase advance $\bar{\phi}$ is given by Eq. (3.64). Unlike before, ϕ does not depend linearly on

ξ_0 and thus the C and S integrals are no longer Fourier transforms. To keep the math simple, we define the w_ξ as

$$w_\xi = -\frac{E_2}{E_0 \Delta \xi^2} (\bar{k}_\beta s - \bar{\phi}) \quad (3.74)$$

while ω_δ is given by Eq. (3.66).

We again assume the Gaussian longitudinal phase space distribution of Eq. (3.67). In this case, the I integrals are straightforward and the evolution is described by

$$\begin{aligned} H_1 &= (1 + 4\sigma_{\xi 0}^4 w_\xi^2)^{-1/4} e^{-\sigma_{\delta 0}^2 \omega_\delta^2 / 2} \\ \psi_1(w_\xi, \omega_\delta) &= \frac{1}{2} \arctan(2\sigma_{\xi 0}^2 w_\xi). \end{aligned} \quad (3.75)$$

H_2 and ψ_2 are given by making the replacement $\omega_\delta \rightarrow 2\omega_\delta$ and $w_\xi \rightarrow 2w_\xi$. In this case ψ does not strictly satisfy the requirements for Eq. (3.59) to be valid; however, $2\sigma_{\xi 0}^2 w_\xi$ must be small when the emittance is still growing (the emittance saturation length for mismatch is given by $\sigma_{\xi 0}^2 w_\xi = 2$), allowing us to expand the arctan and satisfy the condition that $\psi_2 = 2\psi_1$. To calculate the saturation lengths, we let $(1 + 16\sigma_{\xi 0}^4 w_\xi^2)^{-1/4} \approx e^{-1}$ for mismatch (H_2) which corresponds to $\sigma_{\xi 0}^2 w_\xi = 2$. For transverse offset we assume $\sigma_{\xi 0}^2 w_\xi = 4$ in H_1 in order to calculate the saturation length. The saturation lengths are shown in Table 3.2.

As before, all longitudinal slices have the same slice emittance. The evolution of each longitudinal slice is given by $H(\omega_\delta) = e^{-\sigma_{\delta 0}^2 \omega_\delta^2 / 2}$ and $\psi(w_\xi) = -w_\xi \xi^2$. The saturation length is the same as in the linear loading case. If the bunch starts offset, the ξ^2 dependence of the phase leads to a large region in the center of the bunch that undergoes transverse oscillations in phase. This is evident in the bottom of Fig. 3.16.

The energy slice emittance is analytically tractable, but the solution is cumbersome and the H and ψ functions are not easily extracted. To compare to the linear case, we use numerical particle tracking to propagate a witness beam in the accelerating field from Fig. 3.15(b) and then numerically evaluate Eq. (3.61) to get the theoretical prediction. For this example, the beam is mismatched but not transversely offset. Fig. 3.18 shows a comparison of the energy slice spot size evolution between the linear and quadratic cases. In both cases the spot size varies across

the energy slices at the exit of the plasma. Unlike the linear loading case, the emittance in the quadratic loading case varies across the energy slices. This variation is described by the H_2 function which is visible in the figure as a δ dependent dampening of spot size oscillation. The low energy slices of the beam disappear because particles can only move to slices with larger δ (if the initial energy spread is uncorrelated).

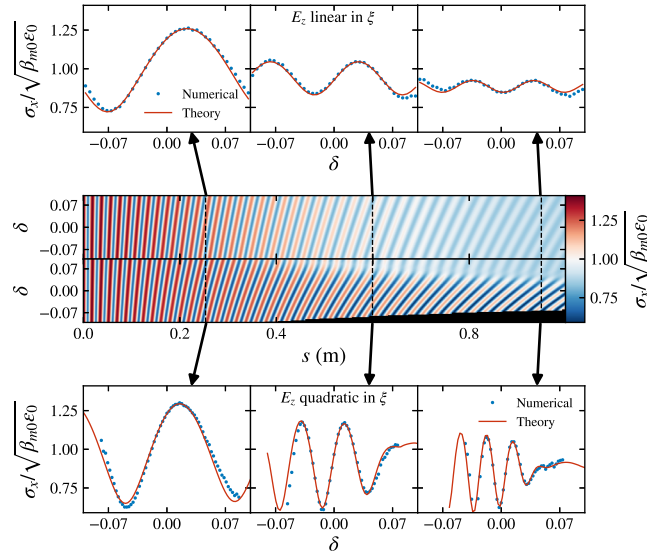


Figure 3.18: Spot size for each energy slice in the beam. The top of the figure shows the evolution for the longer beam shown in Fig. 3.15(a) and the bottom shows the shorter beam from Fig. 3.15(b). In both cases the initial energy spread is $\sigma_{\delta 0} = 4\%$. Chromatic dephasing cannot occur within a slice; the dampening of the β function oscillations is due to particle exchange between slices driven by wake loading. If E_z is linear in ξ (top), the emittance grows uniformly for all slices and the beam size only varies with the β function of each slice. If E_z is quadratic in ξ (bottom), the emittance varies across the slices leading to more complex dynamics.

The variation in beam parameters with energy could be used to measure the wake loading. If the beam is intentionally mismatched into the plasma, the C-S parameters and emittance of each energy slice will have a dependence on the wake loading. An imaging spectrometer can then indirectly measure the loading by looking at the variation in σ_x with energy.

3.4.6 Emittance Evolution in a Plasma Source with Density Ramps

In the previous examples, we have only considered plasma sources with uniform density. Here, we analytically calculate the quantities necessary to find the emittance growth in a plasma source with adiabatic density ramps at the entrance and exit. We assume the accelerating field varies with plasma density according to the simple model

$$E = E_0 (\sqrt{\eta} - 2\eta) - E_1 \eta \frac{\xi_0}{\Delta\xi}, \quad (3.76)$$

where $\eta = n_e/n_{eu}$, n_{eu} is the uniform density in the bulk, $E_1\eta/\Delta\xi$ is the slope of the wakefield in the blowout regime [41, 42] and $E_0 (\sqrt{\eta} - 2\eta)$ describes the variation in longitudinal phase and maximum wake amplitude with plasma density [69]. The fields can be approximated from first principles as

$$E_0 = \frac{\kappa}{2} \pi c \sqrt{\frac{n_{eu} m_e}{\epsilon_0}}, \quad \frac{E_1}{\Delta\xi} = \frac{n_{eu} e}{2\epsilon_0}, \quad (3.77)$$

where κ is a multiplier accounting for the current of the drive beam. κ typically varies between 1 and 2. Beam loading is ignored in this example to keep the math tractable.

The energy of a single particle is then given by

$$\begin{aligned} \gamma_b = & \gamma_{b0} + \delta_0 \gamma_{b0} - \frac{E_0 e}{m_e c^2} G_1 \\ & + \frac{e}{m_e c^2} \left(2E_0 + E_1 \frac{\xi_0}{\Delta\xi} \right) G_2, \end{aligned} \quad (3.78)$$

where

$$G_1(s) = \int_0^s ds' \sqrt{\eta(s')}, \quad G_2(s) = \int_0^s ds' \eta(s'). \quad (3.79)$$

To find the betatron phase advance in the ramp we assume the energy spread and energy gained in the ramp is small compared to the beam's centroid energy throughout the ramp. We can then expand k_β as

$$k_\beta = \frac{\omega_{pu}}{c\sqrt{2\gamma_{b0}}} \eta \left(1 - \frac{1}{2} \frac{\Delta\gamma_b}{\gamma_{b0}} \right),$$

where $\Delta\gamma_b = \gamma_b - \gamma_{b0}$ and $\omega_{pu}^2 = n_{eu} e^2 / m_e \epsilon_0$ is the plasma frequency of the uniform density region. As before, the betatron phase advance is linear in δ_0 and ξ_0 : $\Delta\phi = -\delta_0 \omega_\delta - \xi_0 \omega_\xi$. The I integrals

are given by Eq. (3.65) with

$$\begin{aligned}\omega_\xi &= \frac{\omega_{pu}}{c\sqrt{2\gamma_{b0}}} \frac{E_1 e}{2m_e c^2 \gamma_{b0} \Delta\xi} D_2 \\ \omega_\delta &= \frac{1}{2} \frac{\omega_{pu}}{c\sqrt{2\gamma_{b0}}} G_1.\end{aligned}\tag{3.80}$$

The betatron phase advance of the reference particle is

$$\bar{\phi} = \frac{\omega_{pu}}{c\sqrt{2\gamma_{b0}}} \left[G_1 - \frac{E_0 e}{2m_e c^2 \gamma_{b0}} (2D_2 - D_1) \right],\tag{3.81}$$

where

$$D_1 = \int_0^s ds' \sqrt{\eta(s')} G_1(s'), \quad D_2 = \int_0^s ds' \sqrt{\eta(s')} G_2(s').\tag{3.82}$$

The emittance growth and beam moments are then found by getting H_1 , H_2 , ψ_1 , and ψ_2 from the I integrals and then using Eq. (3.59) to find the emittance and Eqs. (3.50)-(3.54) to find the moments. Without carrying out the full calculation, it is apparent that minimizing the G and D integrals will minimize ω_ξ , ω_δ , and $\Delta\gamma_b$, i.e., minimize the undesirable impacts the ramp has on the beam.

As the beam enters the uniform density region, it is accelerated significantly and we can no longer assume $\Delta\gamma_b \ll \gamma_{b0}$. Instead, we use the solution presented in Sec. 3.4.5 with the addition of an initial phase and an initial energy of $\gamma_0 \rightarrow \gamma_{bl} = \bar{\gamma}_b(s = l)$, where l is the length of the ramp. The exit ramp is handled the same way as the entrance ramp except the initial energy is $\gamma_{bL} = \bar{\gamma}_b(s = l + L)$, where L is the length of the uniform plasma. The resulting piece-wise functions for $\bar{\phi}$, ω_ξ , and ω_δ are fairly cumbersome and are given in Appendix E.

Qualitatively, the solution is similar to that of a uniform plasma presented in Sec. 3.4.5. The primary difference is the conversion of beam size into divergence by the entrance ramp (focusing) and divergence into beam size by the exit ramp (defocusing). This is shown in Fig. 3.19 where σ_x and $\sigma_{x'}$ are plotted for a mismatched beam propagating through a plasma source with ramps. In this example, the ramps are short and only a small amount of chromatic dephasing occurs in them. The majority of the emittance growth occurs in the bulk plasma.

If the ramps are long or poorly shaped, they can have significant impacts on the witness beam. The integrals G_1 , G_2 , D_1 , and D_2 determine the amount of phase spread, betatron oscillation, and,

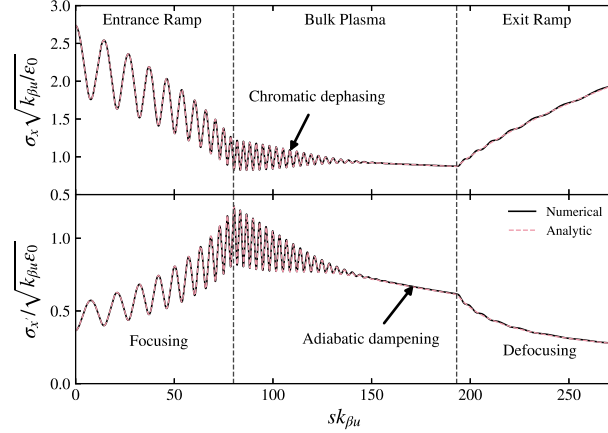


Figure 3.19: Evolution of the spot size (top) and divergence (bottom) of a mismatched beam entering a plasma source with entrance and exit ramps. The beam is initially focused by the entrance ramp into the bulk of the plasma source where it undergoes chromatic dephasing and becomes matched to the plasma. As the beam is accelerated, the divergence continues to decrease due to adiabatic dampening before the beam is defocused by the exit ramp.

in the case of a beam driven wake, drive beam energy loss in the ramp. All the integrals are reduced if the integrated plasma density— G_2 —is minimized. The adiabatic ramp that minimizes the impact on the beam therefore has the highest density gradient possible while remaining adiabatic. Solving Eq. (3.4) gives the optimal ramp as $\eta = 1/(1-2\alpha_m s)^2$. From experience, the ramp needs $|\alpha_m| < 0.1$ for the ramp to be well described by adiabatic theory. In Fig. 3.20, we compare several different adiabatic ramp shapes that all focus the beam by the same amount. The ramp shapes that minimize the density integral induce smaller impacts on the witness beam. The full ramp shape is important and the impact of the ramp on the beam cannot be described using only the half width and the adiabatic parameter.

3.5 Additional Effects

We have limited our discussion on witness beam dynamics to chromatic effects within the wake. This description is complete if the wakefields are given by the expressions derived in

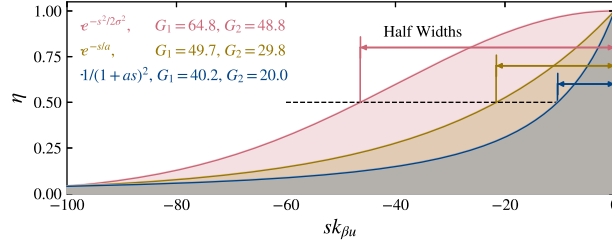


Figure 3.20: Three plasma ramps with different profiles that all reduce the beta function by a factor of 5. The chromatic phase spread, betatron oscillations and drive beam energy loss in the ramp all depend on the integral over the plasma density (G_2 shaded area). The first ramp (blue, bottom) is the most efficient adiabatic ramp possible while the Gaussian (red, top) is barely adiabatic at the tail. The ramp shape can have a significant impact on the beam, and the ramp half width is not a good indicator of the effectiveness of the ramp.

Sec. 2.2.3.² Two main mechanisms can modify the field structure and introduce dynamics beyond those described here. The first is feedback between the witness beam and the wake other than beam loading of the longitudinal field. The second is the motion of the ions within the wake.

3.5.1 Feedback on the Wake

The wake structure depends on the beam distribution of both the drive and witness beam. As the beams propagate through the accelerator and experience betatron oscillations and acceleration, their phase space distributions will change. These changes can feedback onto the wake, changing the wake structure, which changes the wakefields and then impacts the beam evolution. The only place the beam distribution entered into our description of a highly nonlinear wake was through the current distribution of Eq. (2.74). For most of the wake, r_b is larger than the beam radius and Eq. (2.74) depends only on the longitudinal beam distribution. At the front of the drive beam, however, r_b is less than the beam radius and the transverse distribution matters.

The blowout starts somewhere within the front of the drive beam and expands radially towards the rear of the drive beam. The particles within the blowout experience the transverse focusing force of the ion column given by the transverse wakefield of Eq. (2.108). Many of the

² By complete, we mean a full description within the validity of the assumptions we have made, i.e., the ultrarelativistic approximation.

particles, particularly at the front of the beam, are outside of the bubble and do not see the focusing force. These particles will diverge away from the reference orbit following a vacuum-like trajectory, and the density at the front of the beam will gradually reduce. The start of the blowout will then move further back along the drive beam in a process called *head erosion* [70, 71]. The erosion is even worse if the plasma is beam ionized rather than pre-ionized. An example of head erosion in a beam ionized lithium plasma is shown in Fig. 3.21. Particles not captured by the focusing forces of the wake diverge away from the bubble well before their energy is depleted.

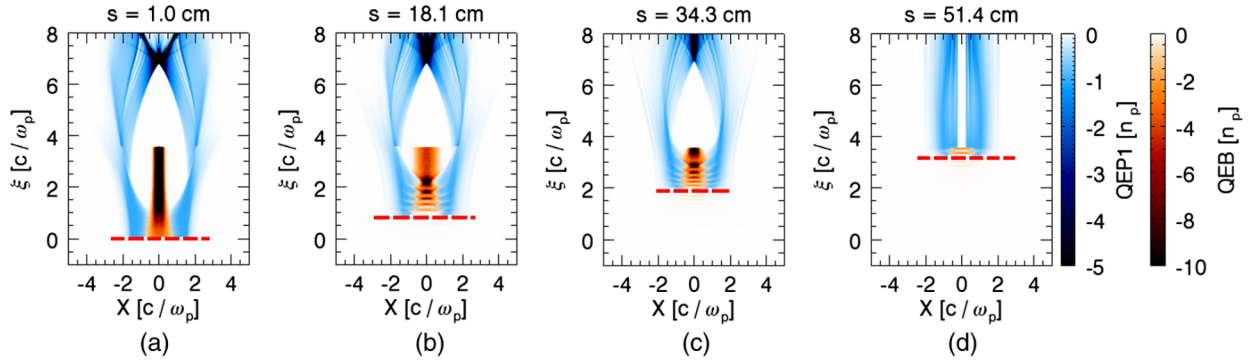


Figure 3.21: Simulation of head erosion in a PWFA where the lithium plasma is ionized by the drive beam. The blue represents the plasma density; the red is the drive beam density. The electrons at the front of the beam do not see the focusing force from the plasma and diverge as the beam propagates. Eventually, the majority of the drive beam charge is lost. Reproduced from Ref. [71].

Eq. (2.74) was derived assuming the drive beam was cylindrically symmetric. This symmetry is partly responsible for the irrelevance of the beam's transverse distribution on the evolution of the wake. We have already discussed one way to break the symmetry—transverse offsets of the witness beam. Each electron beam will try and drive a wake centered on the beam. If the witness beam is offset, it will displace the center of the wake along the length of the beam. This will feedback on the transverse dynamics of the rear of the beam. It ends up this feedback is unstable and the beam offset will grow exponentially in a process called the *hosing instability*. The same effect can happen in the drive beam if the beam is tilted or asymmetric. A simulation showing the growth of the transverse offset along the drive beam is shown in Fig. 3.22(a), notice the asymmetry in the wake as it oscillates resonantly with the beam. Hosing was first predicted

in the early days of plasma wakefield acceleration and many approaches for modelling it have been developed [72, 73, 74, 75, 76].

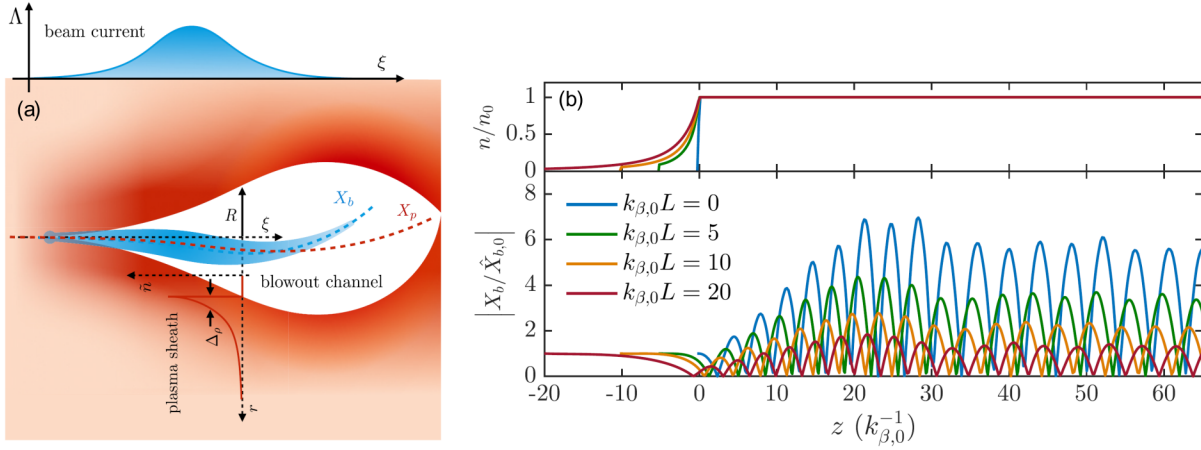


Figure 3.22: Hosing instability in a PWFA. (a) the centroid offset of the beam X_b grows along its length, driven resonantly with centroid offsets of the wake X_p . The beam travels right to left (b) growth of the centroid offset of the rear of the beam relative to the beam's initial offset $X_{b,0}$. The plasma ramps (top) result in a dampening and eventual saturation of the instability. Reproduced from Refs. [77, 78]

Hosing can be seeded either by initial transverse offsets between the drive and witness beam or small asymmetries in the transverse distributions of the two beams. Because hosing is an instability, any small seed will eventually result in large amplitude oscillations. As a result, theoretical models of hosing, particularly earlier ones, tended to predict catastrophic beam loss in a PWFA. Luckily, experimentalists were undeterred and found that hosing in actual experiments was not observed [79]. Several different mechanisms for suppression of the instability have been proposed. Mehrling *et al.* [78] suggested that the energy loss of the drive beam, energy spread in the witness beam, and ramps at the entrance and exit of the plasma could all act to suppress hosing. In a second paper, Mehrling *et al.* [80] suggested that ion motion would also act to suppress the instability. The most likely reason hosing has not yet been observed is that most experiments have a large energy spread in the witness beam, suppressing the resonance that drives the instability.

3.5.2 Ion Motion

When we derived the wake structure and wakefields in both the linear and nonlinear regime, we assumed the ions were stationary because the timescale of their dynamics is much longer than the timescale of the electron dynamics. As far as wake formation goes, this is a good assumption. The ion motion is localized to the transverse center of the wake, where the electron beams are, and therefore, has no effect on the fields near the electron sheath. Since the fields near the sheath are unchanged, the evolution of the innermost electron, and thus the bubble radius r_b , is unchanged by the ion motion except at the very rear of the wake, well behind the witness beam. The wakefields in the vicinity of the witness beam, however, are modified by the change in ion density and impact the transverse motion of the beam particles [81, 82, 83].

The ions are pulled towards the center of the wake by the beam fields due to their positive charge. The resulting charge density in the ion column is no longer uniform, ρ is modified and Eq. (2.100) is no longer the solution to Eq. (2.92). Refs. [81, 82, 83] describe a simple model for the ion motion that can be used to understand how the transverse fields within the wake change. The force on the ions is composed of the transverse field from the beam and the radial field from the ions. When r is small compared to the beam size $r \ll \sigma_r$, we can expand $n_b(r, \xi)$ to zeroth order in r and use Eq. (2.170) to find the radial electric field from the beam:

$$E_{r,b}(r, \xi) = -n_b(r=0, \xi) \frac{e}{2\epsilon_0} r. \quad (3.83)$$

We will treat the ions as concentric rings. If we assume no trajectory crossing between sheets, then the electric field from the ions can be found by solving the 2D Poisson like Eq. (2.88) for the scalar potential. The solution is given by 2D Gauss' Law:

$$E_{r,i}(r, \xi) = \bar{n}_i(r) \frac{e}{2\epsilon_0} r, \quad (3.84)$$

where $\bar{n}_i(r)$ is the average ion density within a ring of radius r . Because we assumed no trajectory crossing, the average ion density is simply $\bar{n}_i(r) = n_0 r_0^2 / r^2$. The equation of motion for an ion ring

is then

$$\frac{d^2 r}{d\xi^2} = \frac{Ze^2}{2\epsilon_0 c^2 m_i} \left(-n_b r + n_0 r_0^2 \frac{1}{r} \right), \quad (3.85)$$

where m_i is the ion mass and Z is the ionization level of the ion. In a highly nonlinear blowout, $n_b \gg n_0$ by definition so the ion motion is initially dominated by the beam fields. The phase picked up by the ions from the drive beam is approximately

$$\Delta\phi \approx k_i \Delta\xi = \Delta\xi \sqrt{\frac{Ze^2 n_b}{2\epsilon_0 c^2 m_i}}, \quad (3.86)$$

where k_i is the ion wavenumber from the drive beam and $\Delta\xi$ is the length of the drive beam. As long as $\Delta\phi \ll \pi/2$, the effects of ion motion are negligible. A more detailed calculation of the transverse wakefields is given in Ref. [84]. Fig. 3.23 shows several simulations of ion motion. The ions collapse to a narrow peak near the center of the wake where they distort the focusing force nonlinearly. The change in focusing force does two things: first, the beam distribution is distorted by the nonlinearity; second, the emittance grows because the effective matched size is modified.

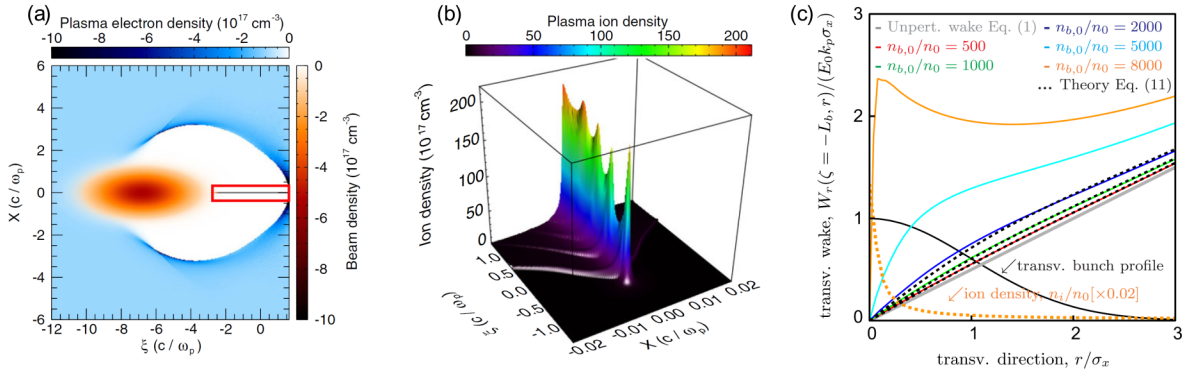


Figure 3.23: Simulations of ion motion. (a) and (b) are reproduced from Ref. [85] and show a simulation of ion motion. (a) shows the beam and plasma density. The witness beam is denoted by the red box. The beams are travelling from right to left. (b) shows the ion density in the x - ξ plane in the vicinity of the witness beam. (c) is reproduced from Ref. [84] and shows the transverse wakefield W_r for different bunch densities. The solid-colored lines show the nonlinearity introduced into the focusing force by the motion of the ions.

Several strategies have been proposed to mitigate the emittance growth caused by ion motion. First, using a plasma with heavy ions significantly reduces the ion motion and formation of nonlinear fields. Second, Gholizadeh *et al.* [86] proposed using a plasma composed of different ions

with reducing mass along s . The technique relies on the fact that the beam distribution can be matched to the nonlinear focusing in such a way the emittance growth is minimized. The difficulty is creating a beam with the desired profile. Slowly reducing the ion mass along the plasma results in an adiabatic change in the nonlinear fields which allows the beam to evolve adiabatically to the matched distribution. Third, An. *et al.* [85] showed simulations that implied over-focusing the beam beyond β_m reduced the emittance growth in the presence of strong ion motion.

3.6 Plasma Source Requirements

Combining everything from this chapter, we have enough information to define the requirements for a plasma source for FACET-II. The bulk plasma density is the first thing to decide. In Sec. 3.3.1, we found that the optimal plasma density for a witness beam with $\sigma_z = 5.2 \mu\text{m}$ was $2.6 \times 10^{16} \text{ cm}^{-3}$. As the FACET-II design matured (see Sec. 5.2.2 for beam parameters), the length of the witness beam was reduced and as a result the optimal plasma density increased to $3.4 \times 10^{16} \text{ cm}^{-3}$ [20]. The bulk plasma density sets the matched beta function to $\beta_m = 6.5 \text{ mm}$. The the FACET-II final focus can achieve at best $\beta^* = 5.0 \text{ cm}$. The entrance ramp will have to reduce β by about a factor of 10. The plasma width should be at least $239 \mu\text{m}$ wide to fully contain the wake. The exit ramp is not as important as the entrance ramp for FACET-II experiments. For convenience, we aim for an inverse of the entrance ramp, an increase in β by a factor of 10.

The length of the bulk plasma is set by the energy loss of the drive beam. We ran a series of simulations at different densities with the drive beam parameters from Table 3.1. The results are shown in Fig. 3.24. Based on the simulations, the drive beam loses energy at a rate of 15.4 GeV/m in a $3.4 \times 10^{16} \text{ cm}^{-3}$ plasma source. Assuming Gaussian ramps which meet our focusing requirements, a 48 cm long plasma will deplete the drive beam energy down to 200 MeV.

We did a lot of analytic work to show that the detailed ramp profile is not important for matching. Using an adiabatic ramp allows us to calculate the beam evolution analytically. We know that small perturbations of such a ramp are unimportant as long as the adiabatic assumption is not violated. The ramp should minimize the integrated plasma density along the ramp to limit

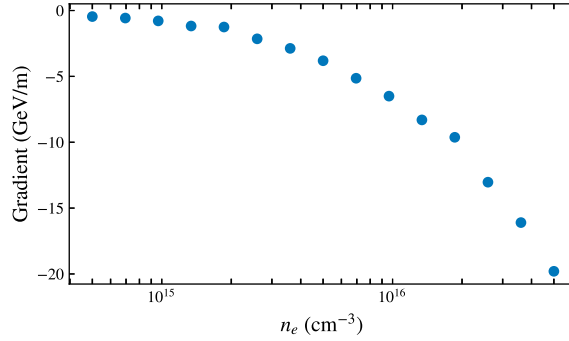


Figure 3.24: Accelerating gradient for the FACET-II drive beam. The gradient determines the length of the plasma source.

the negative impacts on the witness beam's energy and energy spread σ_δ . Even for ramps with significant non-adiabatic sections, the tolerances on the ramp shape are relatively insensitive, as shown for a Gaussian ramp in Fig. 3.25. In the following sections, we show two ramp shapes where the beam evolution can be calculated analytically. These shapes are good candidates for a first pass ramp design with minimal impact on the witness beam. Any other adiabatic ramp shape can, of course, also be handled analytically. Non-adiabatic ramps are still easily designed numerically by back propagating the beam from the start of the uniform plasma to the end of the ramp.

3.6.1 Example Ramp Design

Here, we demonstrate how to use adiabatic and perturbative plasma regions to match a beam originating from a conventional magnetic delivery system into a plasma accelerator. Using these two regions, the beam dynamics can be solved for analytically. In addition, the ramp is robust against errors in the plasma density profile. We assume in this example that the magnetic optics can only focus the beam down to $\beta^* = 10k_{\beta 0}^{-1}$ in vacuum. The plasma ramp must demagnify the β function by a factor of 10 for the beam to be matched to the uniform density, accelerating plasma section. The challenge is to determine an appropriate ramp shape as well as the location of the vacuum focus s^* for the incoming beam.

We will use an adiabatic ramp that is truncated on the upstream side by a short perturbative

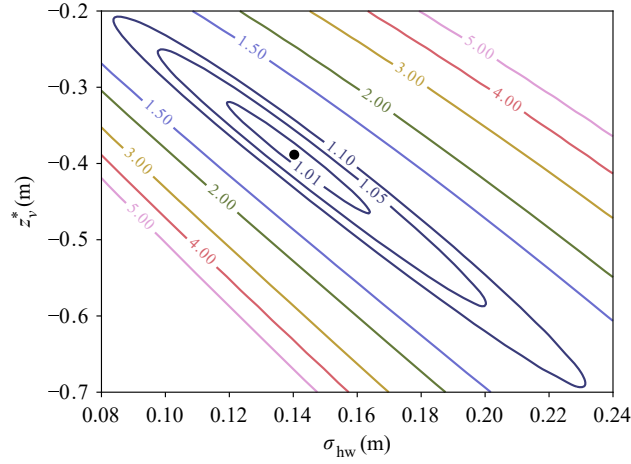


Figure 3.25: Ramp length and electron beam focus tolerances for a Gaussian entrance ramp. Contours are the emittance growth for various combinations of ramp half length σ_{hw} and beam waist locations z_v^* . The beam waist can be adjusted to compensate for errors in the ramp length, which is harder to control. For the 14.0 cm ramp shown here the length tolerance for 1% emittance growth is ± 2.2 cm. Reproduced from [69]

section. The adiabatic section accomplishes most of the beam focusing and the small perturbative section rapidly brings the plasma density up from zero at the start of the ramp. In the adiabatic section, a plasma density given by $\eta(s) = 1/(1 + (as)^2)$ is used because it has a smooth profile at the top and constant α_m in the bulk of the ramp, minimizing the length of the ramp. The perturbative region has a Gaussian plasma density profile given by $\eta(s) = \eta_t \exp(-(s - s_t)^2/2\sigma^2)$, where s_t is the location of the transition between the perturbative and adiabatic section and η_t is the plasma density at that point. The length scale of the Gaussian is set by σ . A Gaussian profile is chosen to make evaluating the integrals in the perturbative section of the ramp straightforward. In the adiabatic section, the beam undergoes adiabatic evolution and remains matched to the plasma $\beta = \beta_m$. To focus the beam by a factor of 10, the plasma density must increase from $\eta = 1/\beta_0^2$ to $\eta = 1$. To keep the ramp as short as possible, the maximum $\mathcal{A}$ is chosen to be 0.1 ($a = 0.2$), which is the approximate upper limit for accurate predictions using the adiabatic approximation. Some focusing will occur in the perturbation section of the ramp so the plasma density increases adiabatically from $\eta = \eta_t = 0.012$ to $\eta = 1$. The lower bound on the plasma density was chosen somewhat arbitrarily to create a perturbative section of reasonable length. In this perturbative

section, β decreases by 8% with respect to β^* .

The length of the adiabatic section is determined by solving $\eta_t = 1/(1 + (as_t)^2)$ for s_t , giving a length of $s_t = -45.4k_{\beta 0}^{-1}$, where $s = 0$ is chosen to be at the top of the ramp. The next step is to determine the length of the perturbative section and the location of the incoming beam's vacuum waist. At the beginning of the adiabatic section, the Courant-Snyder parameters must be $\alpha = \alpha_m = 0.1$ and $\beta = \beta_m = 1/\sqrt{\eta} = 9.13k_{\beta 0}^{-1}$. The length of the perturbation section can be determined by solving Eq. (3.35) for σ . The integrals are taken from 0 to $-\infty$, effectively going backwards from s_t . Using this, the perturbative section needs to have $\sigma = 2.645k_{\beta 0}^{-1}$ to match a $\beta^* = 10k_{\beta 0}^{-1}$ beam into the adiabatic section of the plasma. Next, the transport matrix for the perturbation section can be calculated to find the beam waist. Because the integrals were calculated in the negative direction, the transport matrix propagates in the negative direction. The beam waist can be determined by back propagating the beta function from the end of the perturbative ramp into free space and solving for the vacuum waist from the free space values. Using this method, it is found that the vacuum waist should be placed at $s = -48.04k_{\beta 0}^{-1}$.

The emittance evolution in the ramp can be found using the chromatic amplitude. As an example, we start with an incoming beam with a uniform energy distribution between 1% and -1%. All energy components focus to the same β at the vacuum waist, but α at the waist depends linearly on energy $\alpha(s^*) = r\delta$. Here, r is an arbitrary parameter representing the chromatics in the incoming beam. The δ dependent transport matrix for the perturbative section can be used to find the beam's initial state at the beginning of the adiabatic section. The δ dependent adiabatic solution can be used to find the chromatic amplitude, and thus the emittance growth, at all points in the adiabatic section. Figure 3.26 shows the beam evolution in the plasma for an incoming beam with $r = 5.5$. The centroid energy slice does not display any significant oscillations about the matched size, demonstrating that the beam is correctly matched to the adiabatic section of the ramp. The off energy components undergo significant oscillations in the ramp, but the ramp itself has negligible impact on the beam emittance. The initial projected emittance is larger than the slice emittance due to the chromatics in the beam before it reaches the plasma ramp.

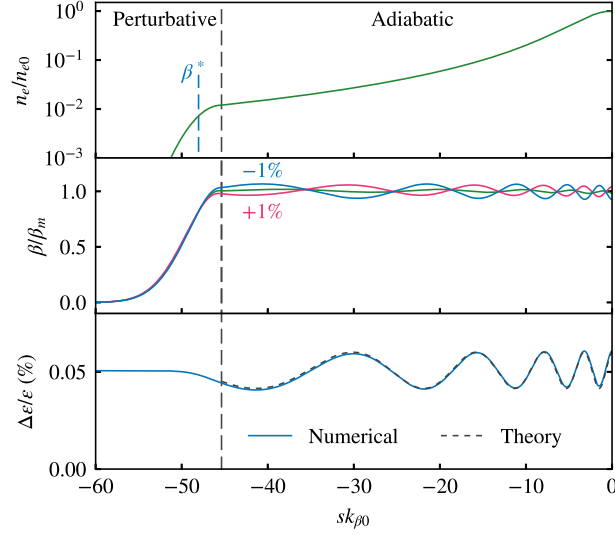


Figure 3.26: An example of a perturbative and adiabatic ramp designed to focus a beam into an accelerating plasma. The incoming beam is assumed to focus all energy components to the same size at the vacuum waist. The length of the perturbation region and the incoming beam's vacuum waist location were determined using the ramp theory presented in Sec. 3.2. The emittance evolution of the beam is well described by the analytic solution found using the chromatic amplitude function.

The theoretical calculation carried out in this example shows good agreement with the numerical solution for both the transverse dynamics and the emittance evolution. The theory presented in Sec. 3.2 provides a simple framework with which to design and evaluate plasma ramps in order to preserve beam quality.

A ramp with an identical functional form can be used to control the divergence of the beam at the exit of the plasma-based accelerator. At the exit, however, the beam energy is greater than at the entrance. This results in a decrease of $k_{\beta 0}$ for the exit ramp, causing an increase in the length scale of the ramp. If the plasma density is uniform throughout the accelerating region (the bulk), then the idealized ramp length increases by a factor of $\sqrt{\gamma_{bf}/\gamma_{bi}}$, where the subscripts i and f denote pre and post acceleration, respectively. After traversing the exit ramp, the beam would behave as if it originated from a vacuum waist with $\beta^* = \beta_0^* \sqrt{\gamma_{bi}/\gamma_{bf}}$, where β_0^* is the vacuum waist produced by the upstream magnetic focusing optics.

3.6.2 Uniform Plasma Ramp

Matching can be achieved in a very short distance by using a low density, uniform plasma in front of the accelerating plasma section. The electron beam enters the low density plasma mismatched and undergoes one quarter betatron oscillation, entering the accelerating plasma at the minimum of the oscillation where $\alpha = 0$. An appropriate choice of plasma density ensures that $\beta = \beta_m$ where the uniform plasma ramp meets the accelerating section. An example of such a ramp is shown in Fig. 3.27 where an $\eta = 0.05$ density plasma produces a factor of 20 reduction in the β function over a distance of $7/k_{\beta 0}$.

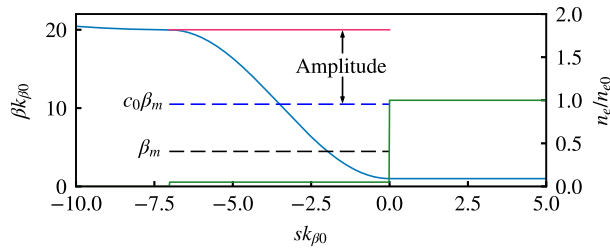


Figure 3.27: A mismatched beam in a uniform density plasma ramp undergoes half an oscillation of β , rapidly focusing the beam. If the plasma ramp is truncated after a quarter betatron period, it can be used to match the beam into a high density plasma in a short distance. The beam oscillates about a center value of $c_0\beta_m$, which is offset from the matched beam size β_m . The amplitude of the oscillation is given by $\sqrt{c_1^2 + c_2^2}$.

The transverse beam evolution in the uniform density “ramp” is described exactly by the adiabatic theory and the results from the preceding section can be used to design the ramp. The plasma density required for focusing an incoming beam is given by

$$\eta = \frac{1 - \gamma_0}{\beta_0 - 1}. \quad (3.87)$$

The necessary length of the ramp can be calculated by solving the phase equation

$$\cos 2\theta = \frac{2\beta_0 - 1 - \beta_m}{1 - \beta_m}, \quad (3.88)$$

with β_m determined from Eq. (3.87) and $\theta = L/\beta_m$ where L is the length of the ramp.

If the beam is at a waist at the beginning of the ramp ($\alpha_0 = 0$), then the length of the ramp must be $L = \pi/(2\sqrt{\eta})$. To the best of our knowledge, this is the shortest ramp yet considered in

the literature. Very little chromatic phase spread develops in such a short ramp so the emittance growth is negligible. The ramp's length and low plasma density mitigate energy modulation of the witness beam before the accelerating section. Though the step function density profile of the theoretical model is nonphysical, the same qualitative behavior will be retained if the change in density occurs over a distance that is small compared to k_β^{-1} .

Chapter 4

Laser Ionized Plasma Sources

The density profile of the plasma used in a plasma-based accelerator has a large impact on the performance of the accelerator. The bulk density impacts the final energy spread and transformer ratio, while the width must be sufficient to fully contain the wake, otherwise the accelerating field is degraded. The plasma should be long enough to fully deplete the drive beam and maximize the energy transferred to the witness beam. Finally, the profile and length of the entrance and exit ramps play a critical role in matching the beam into the plasma in order to preserve the transverse emittance.

The challenge is to produce a plasma that meets all these requirements. The beam parameters at FACET-II make this particularly difficult. In Ch. 3 we developed specific requirements for the FACET-II plasma source that must be satisfied in order to produce a high quality beam. The requirements are demanding, the plasma needs to be 0.5 m long, 250 μm wide with 15 cm shaped ramps on the entrance and exit. Our solution is to laser ionize the plasma from a neutral gas, forming the ramps by partial ionization.

This chapter starts with a discussion of optical field ionization. Several models are presented which allow us to calculate the required laser intensity given a target ionization fraction. Next, we take a brief interlude to introduce the optics background necessary to understand the rest of the chapter. We start with the wave equation and proceed through Fourier optics and relevant nonlinear optics. Then, we introduce the formalism we use to calculate the transverse phase and intensity a laser pulse must have upstream of the plasma source to produce the target intensity

profile. We discuss the distortions introduced in the beam as it ionizes the plasma source. After this, we develop an optical setup capable of shaping a high-power laser. We touch on the fabrication of the optics and present an example design for FACET-II. Finally, we discuss several limitations of the plasma source and possible ways to circumvent these limitations.

4.1 Ionization Models

A high-power laser is required to produce the large plasma sources used for PWFA. These systems are typically 800 nm Ti:sapphire lasers with photon energy around 1.55 eV. Even the atom with the lowest ionization energy, Cesium at 3.89 eV, cannot be resonantly ionized by a single laser photon. Instead the ionization occurs via a process called field ionization, which can be divided into three regimes [87]. At low laser intensity, the atom ionizes in the *multi-photon regime* by absorbing several photons from the laser simultaneously. At high laser intensity, the ionization occurs in the *tunnelling regime*, where the electric field from the laser sufficiently distorts the atomic potential that the probability of the electron tunneling out of said potential becomes significant.¹ As the laser intensity is increased, the electric field becomes strong enough that the potential barrier binding the electron to the atom is removed and the atom is ionized by *barrier suppression ionization*. The three ionization regimes are depicted schematically in Fig. 4.1. The ADK model describes ionization in the tunneling regime. The PPT theory models ionization in both the multi-photon and tunneling regime. Barrier suppression can occur in either the multi-photon or tunneling regime and a time dependent Schrodinger equation (TDSE) simulation is typically required to calculate the ionization rate.

Keldysh [89] wrote the most complete early treatment of field ionization, spanning from the multi-photon regime to the tunneling regime. He defined the Keldysh parameter

$$\gamma = \frac{\omega}{\omega_t}, \quad (4.1)$$

¹ Tunnelling ionization still occurs via the absorption of multiple photons, but the ionization rate can be determined by calculating the tunneling probability and ignoring the photons.

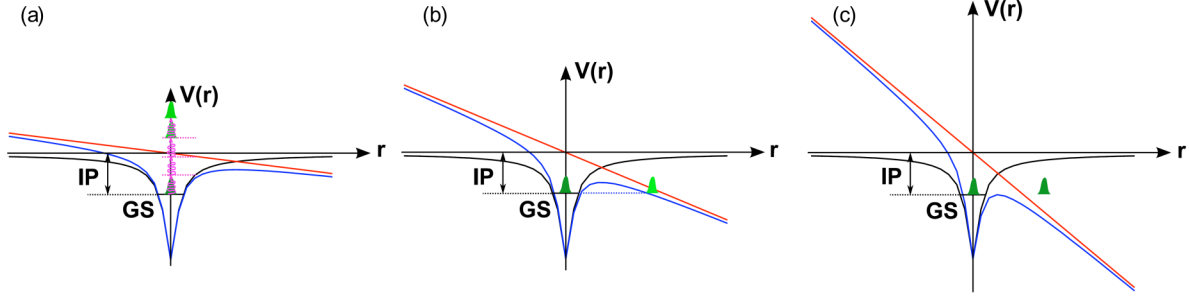


Figure 4.1: Field ionization regimes in order of increasing electric field strength. (a) multi-photon ionization: a single photon does not have enough energy to ionize an electron and multiple photons must be absorbed at once. (b) tunneling ionization: the electric field from the laser distorts the atomic potential enough for the electron to tunnel away from the atom. (c) border suppression ionization: the atomic potential is distorted so much that the electron is no longer bound. Reproduced from Ref. [88].

where ω is the laser frequency and ω_t is the tunneling frequency, defined as

$$\omega_t = \frac{eE}{\sqrt{2m_e\xi_i}}. \quad (4.2)$$

Here, E is the magnitude of the electric field and ξ_i is the ionization energy of the electron. When $\gamma \ll 1$, the field is approximately constant during the time it takes an electron to tunnel out and ionization occurs in the tunneling regime. When $\gamma > 1$ the ionization occurs in the multi-photon ionization (MPI) regime.

The threshold for BSI ionization occurs when the potential barrier containing the electron is completely suppressed. The threshold is [90]

$$E_{BSI} = \frac{1}{e\alpha\hbar} \frac{\xi_i^2}{4Z}, \quad (4.3)$$

where α is the fine structure constant and Z is the ionization state: $Z = 1$ for first ionization, $Z = 2$ for second, etc. The BSI threshold can be crossed in either the multi-photon or tunneling regime depending on the laser frequency; although, for most high ionization threshold gases, the threshold is reached in the tunneling regime.

All the ionization theories calculate an ionization rate. To get the fraction of neutral gas

ionized we need to solve the differential equation

$$\frac{dn_g}{dt} = -n_g w(t), \quad (4.4)$$

where n_g is the neutral gas density and $w(t)$ is the ionization rate as a function of time. The solution is

$$n_g(t) = n_g(t_0) e^{-\int_{t_0}^t w(t') dt'}. \quad (4.5)$$

Letting $n_0 = n_e + n_g$, the ionization fraction is given by

$$\frac{n_e}{n} = \frac{n - n_g}{n} = \frac{1}{n} \left(n - n_g(t_0) e^{-\int_{t_0}^t w(t') dt'} \right) = \left(1 - e^{-\int_{t_0}^t w(t') dt'} \right), \quad (4.6)$$

where we have assumed the ionization fraction is initially 0, i.e., $n_g(t_0) = 0$. If we know the temporal pulse shape of the laser, then we can calculate $\int_{t_0}^t w(t') dt'$ as a function of the peak laser intensity. Inverting this function (typically numerically) allows us to solve for the laser intensity required to produce a particular plasma density.

4.1.1 ADK Tunneling Ionization

In the tunneling regime, field ionization can be described analytically using the ADK ionization theory. The ADK theory comes from a 1986 paper by Ammosov, Delone, and Krainov (ADK) [91]. The theory is based off work originally by Landau and Lifshitz, where they calculated the ionization rate of the hydrogen atom in a constant electric field. Far from any atom, the outer electron sees an effective ion formed from a combination of the nucleus and inner electrons. The potential of this effective ion is a Coulomb potential, so the electron's wave function asymptotically approaches a wave function similar to a hydrogen atom. Smirnov and Chibisov [92] extended the calculation of Landau and Lifshitz to arbitrary asymptotic Coulomb wave functions. Ref. [92] has some errors which Perelomov, Popov, and Terent'ev (PPT) [93] fixed before they time averaged the ionization rate over a sinusoidal electric field. The ADK theory is an extension of this work that introduced the use of Stirling's approximation in determining the unknown asymptotic expansion coefficient. In practice, the use of Stirling's approximation, and thus the contribution of ADK, is

not used, but the theory is still called the ADK ionization theory (For a detailed discussion see Ref. [94]).

We start with an atom that has a valence electron in a state with ionization energy ξ_i , orbital angular momentum l , and its projection m . Note that all units are atomic units: $m_e = \hbar = e = 1$. The electron momentum is $\kappa = \sqrt{2\xi_i}$, the intra-atomic electric field strength is given by $E_0 = \kappa^3$. As long as the external electric field is small compared to the intra-atomic electric field, $E \ll E_0$, then only the tail of the wave function sees the external potential and contributes to ionization. The ionization rate in a constant electric field, from Ref. [93], is given by

$$w_{stat}(E) = |C_{n^*l^*}|^2 \frac{(2l+1)(l+|m|)!}{2^{|m|}(|m|)!(l-|m|)!} \xi_i \left(\frac{2E_0}{E} \right)^{2n^*-|m|-1} \exp\left(-\frac{2E_0}{3E}\right), \quad (4.7)$$

where $n^* = Z/\kappa$ is the effective principle quantum number and $C_{n^*l^*}$ is a constant specific to the atom and the state of the electron being ionized. Z is the charge of the atomic residue, the effective charge of the ion seen by the electron. For the ionization of a neutral atom $Z = 1$, for the ionization of the second electron $Z = 2$, etc.

An analytic solution for $C_{n^*l^*}$ can be found by using effective quantum numbers in the Coulomb wave function:

$$|C_{n^*l^*}|^2 = \frac{2^{2n^*}}{n^* \Gamma(n^* + l^* + 1) \Gamma(n^* - l^*)}, \quad (4.8)$$

where $l^* = n_0^* - 1$ and n_0^* is the effective principal quantum number of the ground state. We never deal with ionization from excited states so $n_0^* = n^*$ according to Refs. [95, 96]. The ADK theory is specifically the use of Stirling's approximation to simplify the expression to

$$|C_{n^*l^*}|^2 = \left(\frac{4e}{n^{*2} - l^{*2}} \right)^{n^*} \left(\frac{n^* + l^*}{n^* - l^*} \right)^{l^*+1/2} \frac{1}{2\pi n^*}. \quad (4.9)$$

Popov [94] pointed out that the use of Stirling's approximation leads to errors of 18% for hydrogen and 25% for helium. There is no practical reason to ever use the approximate version.

The last piece in the ADK theory [91] is to time average over an alternating electric field. Consider an external electric field oscillating according to

$$E(t) = E \cos(\omega t). \quad (4.10)$$

The laser frequency ω must be sufficiently small compared to the tunneling frequency ω_t such that the field is approximately constant in the time it takes the electron to tunnel out. This condition is equivalent to requiring $\gamma \ll 1$. The integration results in an additional constant factor in front of the static rate. For linearly polarized light, the averaged ionization rate is

$$w_{lin}(E) = \left(\frac{3E}{\pi E_0} \right)^{1/2} w_{stat}(E). \quad (4.11)$$

The general formula for elliptically polarized light is given in Refs. [91, 93].

The ionization rate from the $m = 0$ state is generally significantly higher than the ionization rates from the other m states. Specifically,

$$\frac{w_{m=0}}{w_{|m|=1}} = \frac{2E_0}{E}. \quad (4.12)$$

Since it was assumed at the beginning that $E \ll E_0$, this ratio is very large and only the $m = 0$ state needs to be taken into account.

In Fig. 4.2, we compare the ionization rates of three gases we use—hydrogen, helium, and, argon—as a function of laser intensity. The figure also includes the ionization rates of the second electron for helium and argon. We use gas species with high ionization energy for the FACET-II plasma sources to minimize ionization by the electron beam, discussed in Sec. 4.6. Sometimes, we want the beam to ionize the plasma, in which case lithium is used because it has a low ionization energy.

4.1.2 PPT Ionization Theory

The ADK theory is only valid in the tunneling regime where $\gamma \ll 1$. Perelomov, Popov, and Terent'ev (PPT) [93] presented a general solution for the ionization rate of an atom in an alternating electric field that extends from the multi-photon regime through the tunneling regime. They expanded the work of Keldysh [89] and rectified the discrepancy between the Keldysh theory and the well known result of Landau and Lifshitz for the ionization of hydrogen in a uniform field. As in the ADK section, the formulas are all given in atomic units. PPT assumes $E \ll E_0$ and

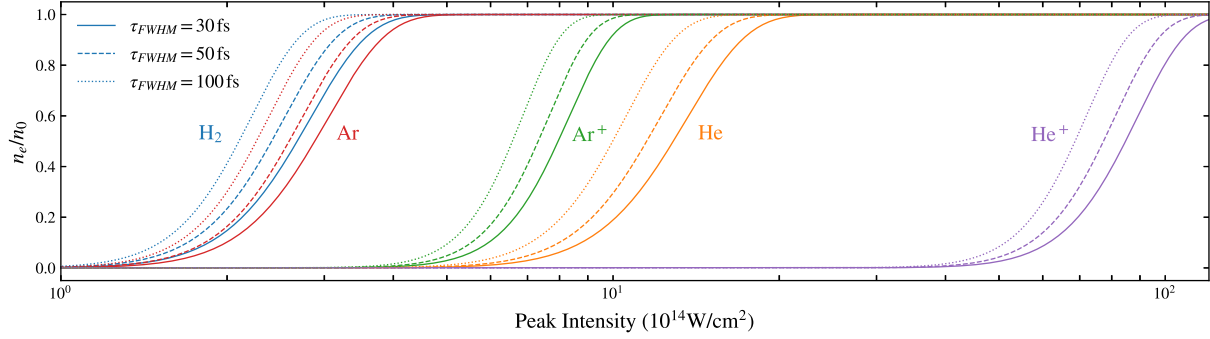


Figure 4.2: Ionization fraction of various gases as a function of peak laser intensity and laser pulse length. The laser pulse is assumed to have a Gaussian temporal profile. Inverting these curves gives the laser intensity required to ionize a target plasma density.

$\omega \ll \xi_i$. Additionally, the theory assumes that the electric field of the laser is constant across the size of the atom and that the electron is non-relativistic so the magnetic field can be ignored.

The PPT theory describes the ionization rate of an atom through the range of field strengths from the multi-photon regime to the tunneling regime [93]. The original theory uses the time dependent Schrödinger equation to directly calculate the ionization rate under the assumptions stated previously, with one addition: the valence electron must be in a short-range potential—it must fall off faster than $1/r$. The initial theory will not give the correct expression if the potential has a Coulomb tail. Essentially, the initial theory is limited to the case of negative ions. The ionization rate for an atom in a linearly polarized wave is given by

$$w_{lm}(E, \omega) = |C_{n^*l^*}|^2 \left(\frac{6}{\pi}\right)^{1/2} f(l, m) \xi_i \left(\frac{E\sqrt{1+\gamma^2}}{2E_0}\right)^{|m|+3/2} A_m(\omega, \gamma) \exp\left(-\frac{2E_0}{3E}g(\gamma)\right), \quad (4.13)$$

where

$$f(l, m) = \frac{(2l+1)(l+|m|)!}{2^{|m|}(|m|)!(l-|m|)!}, \quad (4.14)$$

$$A_m(\omega, \gamma) = \frac{4}{\sqrt{3\pi}} \frac{1}{|m|!} \frac{\gamma^2}{1+\gamma^2} \sum_{n \geq v}^{\infty} \exp\{-\alpha(n-v)\} w_m\left(\sqrt{\beta(n-v)}\right), \quad (4.15)$$

in which

$$w_m(x) = e^{-x^2} \int_0^x e^{y^2} (x^2 - y^2)^{|m|} dy = \frac{x^{2|m|+1}}{2} \int_0^1 \frac{e^{-x^2 t} t^{|m|}}{\sqrt{1-t}} dt, \quad (4.16)$$

and v is the minimum number of photons needed in a multi-photon ionization process:

$$v = \frac{\xi_i}{\omega} \left(1 + \frac{1}{2\gamma^2} \right). \quad (4.17)$$

The parameters α , β , and g are functions of γ , which is itself a function of frequency and field strength. Here, γ is the Keldysh parameter. The parameter definitions are

$$\alpha(\gamma) = 2 \left[\text{Arsh } \gamma - \frac{\gamma}{\sqrt{1 + \gamma^2}} \right] \quad (4.18)$$

$$\beta(\gamma) = \frac{2\gamma}{\sqrt{1 + \gamma^2}} \quad (4.19)$$

$$g(\gamma) = \frac{3}{2\gamma} \left[\left(1 + \frac{1}{2\gamma^2} \right) \text{Arsh } \gamma - \frac{\sqrt{1 + \gamma^2}}{2\gamma} \right]. \quad (4.20)$$

The definition of α and g can be simplified using β

$$\alpha(\gamma) = 2 \left[\text{Arsh } \gamma - \frac{\beta}{2} \right] \quad (4.21)$$

$$g(\gamma) = \frac{3}{2\gamma} \left[\left(1 + \frac{1}{2\gamma^2} \right) \text{Arsh } \gamma - \frac{1}{\beta} \right]. \quad (4.22)$$

The mean probability of ionization for all states with a given orbital angular momentum l is

$$w_l(E, \omega) = |C_{l^*m^*}|^2 \left(\frac{6}{\pi} \right)^{1/2} \xi_i \left(\frac{E\sqrt{1 + \gamma^2}}{2E_0} \right)^{3/2} A_0(\omega, \gamma) \exp \left(-\frac{2E_0}{3E} g(\gamma) \right) \quad (4.23)$$

The original derivation assumed the potential did not have a long Coulomb tail. The tunneling limit was derived assuming a long Coulomb tail, which is why the PPT formula does not reduce to the ADK formula when $\gamma \ll 1$. A later paper by Perelomov and Popov [97] added in the Coulomb correction to make the formula useful for an arbitrary atom. For linearly polarized light, the correction is a multiplicative factor that takes the form

$$\left(\frac{2E_0}{E} \right)^{2n^*}. \quad (4.24)$$

The PPT model for an arbitrary atom can then be written as

$$w_{lm}(E, \omega) = |C_{n^*l^*}|^2 \left(\frac{6}{\pi}\right)^{1/2} f(l, m) \xi_i \left(\frac{2E_0}{E}\right)^{2n^* - |m| - 3/2} \left(\sqrt{1 + \gamma^2}\right)^{|m| + 3/2} \times A_m(\omega, \gamma) \exp\left(-\frac{2E_0}{3E} g(\gamma)\right). \quad (4.25)$$

With the Coulomb correction, the PPT formula of Eq. (4.25) reduces to the time averaged ADK formula when $\gamma = 0$, as expected. It should be noted that the Coulomb correction is only correct if the field satisfies $E \gg n^*(\omega/\xi_i)^2 E_0$. Further, information on how to find the momentum spectrum of the ionized electrons is given in Refs. [97, 98].

The Coulomb correction from [97] has a condition for validity that can also be written as $\gamma \ll \gamma_c$, where

$$\gamma_c = \sqrt{\frac{E_0}{n^* E}}. \quad (4.26)$$

At a large enough γ , the Coulomb correction breaks down. There is an improved expression for the Coulomb correction given by [99]

$$\left(\frac{2E_0}{E}\right)^{2n^*} \left(1 + \frac{2\gamma}{e}\right)^{-2n^*}, \quad (4.27)$$

where e is the base of the natural logarithm. This correction is only valid assuming the parameter $\mu \ll 1$ where $\mu = Z\omega/(2\xi_i)^{3/2}$.

The use of PPT in the literature can often be confusing and appear inconsistent. In Appendix F, we compare and rectify difference between the theory as presented in different sources. The PPT formula presented in Eq. (4.25) is already time averaged, the time averaging works the same way as in the ADK formula and can be removed using Eq. (4.11).

Lai *et al.* [100] compared the PPT and ADK models to ion yields measured in ionization experiments. The resulting ion yield is compared against the theoretical predictions in Fig. 4.3. Notice the excellent agreement between the PPT model and the measured data for ionizing argon with an 800 nm laser. The version of PPT used did not contain the γ dependent Coulomb correction, it used Eq. (4.24).

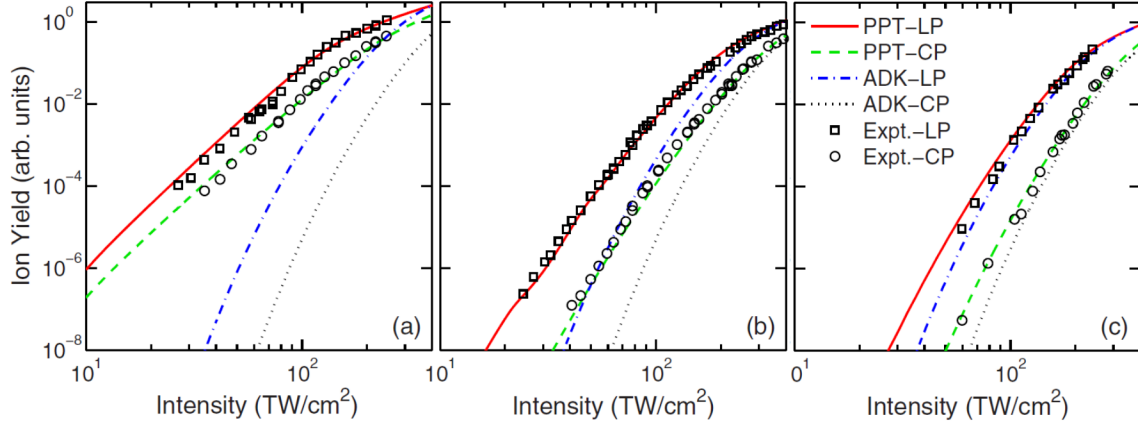


Figure 4.3: Measured ion yield of argon compared to the PPT and ADK ionization theories. In (a) the gas is ionized with a 400 nm laser, (b) 800 nm, and (c) 1.3 μm . Reproduced from Ref. [100]

4.1.3 Ionization Formulas in Practical Units

Both the ADK and the PPT models are given in atomic units. In this section we convert them to more practical units.

All quantities with units are replaced by the SI equivalent except for the ionization energy, which is given in eV. First, the ionization energy in Hartree is replaced by

$$\xi_i[\text{h}] \rightarrow \frac{e}{\alpha^2 c^2 m_e} \xi_i[\text{eV}]. \quad (4.28)$$

The ionization rate replacement is given by

$$w[1/\tau_a] \rightarrow \frac{\hbar}{\alpha^2 c^2 m_e} w[1/\text{s}], \quad (4.29)$$

and the field strength is replaced by

$$E[E_a] \rightarrow \frac{e\hbar}{\alpha^3 c^3 m_e^2} E[\text{V/m}]. \quad (4.30)$$

The ratio E/E_0 is unit-less, but E_0 needs to be given in V/m:

$$E_0[\text{V/m}] = 2^{3/2} \frac{\sqrt{m_e c}}{\hbar} \xi_i^{3/2}[\text{eV}]. \quad (4.31)$$

n^* itself is unit less, but it has a numerical factor for calculating it from $\xi_i[\text{eV}]$

$$n^* = \alpha c \sqrt{\frac{m_e}{2e}} \frac{Z}{\sqrt{\xi_i[\text{eV}]}}, \quad (4.32)$$

With these replacements, it is a simple matter of substitution to convert the ADK and PPT formulas into more useful units.

ADK in useful units is given by

$$w_{stat}(E) = |C_{n^*l^*}|^2 \frac{e}{\hbar} \frac{(2l+1)(l+|m|)!}{2^{|m|}(|m|)!(l-|m|)!} \xi_i \left(\frac{2E_0}{E} \right)^{2n^*-|m|-1} \exp \left(-\frac{2E_0}{3E} \right). \quad (4.33)$$

The time averaging factor is unit-less and the functional form does not change from Eq. (4.11).

The primary difference between atomic and SI units comes from the new expression for E_0 given by Eq. (4.31).

For the PPT theory, notice that γ is dimensionless and so are α , β , and g . As we have Eq. (4.1) in SI units, the formulas for α , β , and g remain unchanged. The expression for A_m is also dimensionless and requires no conversion. The version of the PPT formula in useful units is given by

$$w_{lm}[1/s] \frac{\hbar}{\alpha^2 c^2 m_e} = |C_{n^*l^*}|^2 \left(\frac{6}{\pi} \right)^{1/2} f(l, m) \frac{e}{\alpha^2 c^2 m_e} \xi_i[\text{eV}] \left(\frac{2E_0}{E} \right)^{2n^*-|m|-3/2} \times \left(\sqrt{1+\gamma^2} \right)^{|m|+3/2} A_m(\omega, \gamma) \exp \left(-\frac{2E_0}{3E} g(\gamma) \right). \quad (4.34)$$

This trivially simplifies down to

$$w_{lm}[1/s] = |C_{n^*l^*}|^2 \frac{e}{\hbar} \left(\frac{6}{\pi} \right)^{1/2} f(l, m) \xi_i[\text{eV}] \left(\frac{2E_0}{E} \right)^{2n^*-|m|-3/2} \times \left(\sqrt{1+\gamma^2} \right)^{|m|+3/2} A_m(\omega, \gamma) \exp \left(-\frac{2E_0}{3E} g(\gamma) \right). \quad (4.35)$$

For all our calculations, we will be using the version of PPT with $m = 0$. In this case the formula reduces to

$$w_l[1/s] = |C_{n^*l^*}|^2 \frac{e}{\hbar} \left(\frac{6}{\pi} \right)^{1/2} (2l+1) \xi_i[\text{eV}] \left(\frac{2E_0}{E} \right)^{2n^*-3/2} \left(\sqrt{1+\gamma^2} \right)^{3/2} A_0(\omega, \gamma) \exp \left(-\frac{2E_0}{3E} g(\gamma) \right), \quad (4.36)$$

where

$$A_0(\omega, \gamma) = \frac{4}{\sqrt{3\pi}} \frac{\gamma^2}{1+\gamma^2} \sum_{n \geq v}^{\infty} \exp\{-\alpha(n-v)\} w_0 \left(\sqrt{\beta(n-v)} \right), \quad (4.37)$$

in which

$$w_0(x) = e^{-x^2} \int_0^x e^{y^2} dy = \frac{x}{2} \int_0^1 \frac{e^{-x^2 t}}{\sqrt{1-t}} dt. \quad (4.38)$$

$w_0(x)$ is now just Dawson's integral, a common special function.

4.1.4 SAE-TDSE Ionization Calculations

If the electric field is large enough, then the potential barrier binding the electron to the atom is completely suppressed and neither ADK nor PPT strictly apply. Theoretical models exist for handling the transition from the tunneling regime to the BSI regime. Surprisingly, the use of the tunneling formula (ADK) in the BSI regime overestimates the ionization rate. The BSI threshold is typically irrelevant for our experiments, with one exception. We use four different gas species in our plasma sources: hydrogen, helium, argon, and lithium. The first three have large ionization energies. As a result, the ionization stays in the tunneling regime, or at least the ionization chance has reached close to 100% early in the laser pulse, before the intensity becomes large enough to reach the BSI regime. Lithium is a special case; the BSI threshold is reached when $\gamma = 3.66$, and ionization is occurring in the multi-photon regime. Over the BSI threshold, we should treat the ionization rate calculated with PPT with skepticism.

No analytic theory exists to calculate the BSI ionization rate when the threshold occurs in the multi-photon regime. Instead, we turned to numerical solutions to the time dependent Schrodinger equation (TDSE) using the single active electron (SAE) approximation. Our colleagues at CU, in the Becker group, performed the calculations for us using a laser field we provided them. From the simulation, we received the ionization probability as a function of time [101]. Combined with the pulse envelope, we can extract the time averaged ionization rate using

$$w = -\frac{d}{dt} \left(1 - \frac{n_e(t)}{n_0} \right) \quad (4.39)$$

Fig. 4.4(a) shows a plot of the ionization rate as a function of E_0 for several different pulse shapes. A simple exponential fit does a good job of estimating the ionization rate:

$$w_{lin}[\text{fs}^{-1}] = 2.18 \times 10^{-6} E[\text{GV/m}]^{5.87} \quad (4.40)$$

The ionization rate is less than that predicted by either PPT or ADK as shown in Fig. 4.4(b).

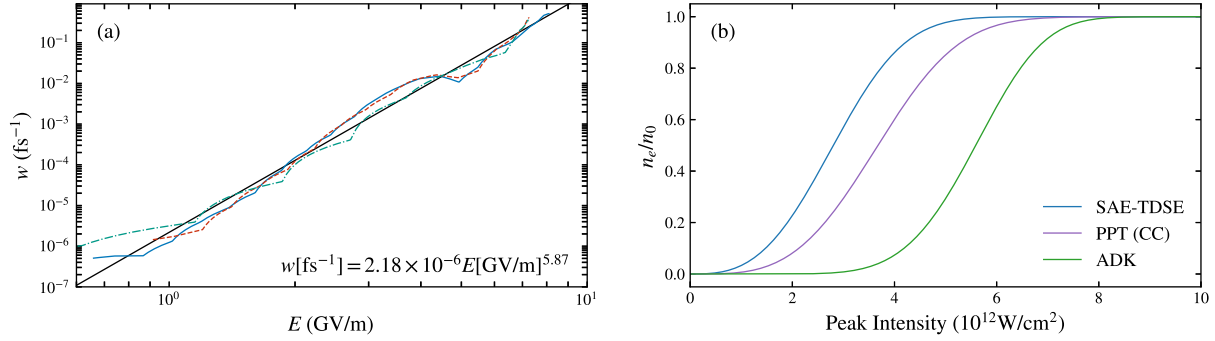


Figure 4.4: Ionization rate of lithium calculated numerically using SAE-TDSE. (a) shows the numerical rate for three different temporal pulse shapes (colored curves). An exponential fit was used to model the rate (black curve). (b) compares the ionization fraction as a function of laser intensity calculated using the SAE-TDSE ionization rate, the PPT rate, and the ADK rate.

4.2 Lasers and Optics

From the beam dynamics chapter, Ch. 3, we know what the requirements on the plasma density are to produce high quality beams from a PWFA. From Sec. 4.1 on ionization models, we know how to figure out what the magnitude of the electric field should be to produce the required plasma density in a uniform gas. All we need to do now is create an electric field with the desired magnitude. This section contains the optics background necessary to understand our approach to laser ionization and beam shaping. As the laser ionizes the plasma source it will interact with the plasma changing its evolution. We also include the necessary optics background to understand this process.

4.2.1 Basics of Light

Light is an electromagnetic wave propagating with frequencies in the infrared or higher. We will be interested in light propagating in a partially ionized plasma. The fields are then described by the inhomogeneous wave equation derived in Sec. 2.1.3. We rewrite Eq. (2.25) here for convenience

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{E} = \frac{1}{\epsilon_0} \nabla \rho + \mu_0 \frac{\partial \mathbf{J}}{\partial t}. \quad (4.41)$$

In vacuum ρ and $\mathbf{J}$ are zero, and we get the homogeneous wave equation

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) \mathbf{E} = 0. \quad (4.42)$$

One particular solution to the wave equation is the real part of the plane wave:

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t + \phi_0)}, \quad (4.43)$$

where $\mathbf{E}_0$ must be perpendicular to $\mathbf{k}$,² $\mathbf{E}_0$ is composed of the amplitude $E_0 = |\mathbf{E}_0|$ and polarization $\hat{\mathbf{E}}_0$,³ $\mathbf{k}$ is the wave vector, ω is the angular frequency, and ϕ_0 is the phase offset of the wave. The frequency and wavenumber are related by the dispersion relation $\omega = c|\mathbf{k}|$. The wave propagates in the $\mathbf{k}$ direction with velocity c . The general solution to the wave equation is a superposition of plane waves with different $\mathbf{E}$, $\mathbf{k}$, and ϕ_0 .⁴ The magnetic field has an identical wave equation, the solutions of which are given by plane waves with the same form as Eq (4.43). The magnetic field and electric field are not independent, but are coupled by Faraday's Law (Eq. (2.11)) and Ampere's Law (Eq. (2.12)). The result is that the plane wave for the electric field is accompanied by a plane wave for the magnetic field with the same $\mathbf{k}$ and ω but an amplitude given by $\mathbf{B}_0 = \frac{1}{\omega} \mathbf{k} \times \mathbf{E}_0$ [32].

As light propagates through a medium, it will polarize it, resulting in a charge and current density of

$$\rho_{pol} = -\nabla \cdot \mathbf{P}, \quad \mathbf{J}_{pol} = \frac{\partial \mathbf{P}}{\partial t}. \quad (4.44)$$

Inserting Eq. (4.44) into Eq. (4.41) results in the wave equation in terms of the polarization

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = -\frac{1}{\epsilon_0} \nabla(\nabla \cdot \mathbf{P}) + \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{P}}{\partial t^2}. \quad (4.45)$$

In an isotropic linear medium, such as a gas or plasma, the polarization depends linearly on the electric field

$$\mathbf{P} = \epsilon_0 \chi^{(1)} \mathbf{E}, \quad (4.46)$$

² Equivalently, $\mathbf{E}_0 \cdot \mathbf{k} = 0$ a direct result of Eq. (2.9).

³ We choose to let $\mathbf{E}_0$ be a vector of reals. Often $e^{i\phi_0}$ is included as part of $\mathbf{E}_0$ making it a vector of complex numbers.

⁴ Plane waves are one particular choice of basis for writing the general solution. Many other are possible, we will discuss the Bessel function basis in Sec. 4.3.

where $\chi^{(1)}$ is the first order susceptibility of the medium. Using Eq. (4.46) to replace $\mathbf{P}$ in the wave equation gives

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = -\chi^{(1)} \nabla(\nabla \cdot \mathbf{E}) + \frac{\chi^{(1)}}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (4.47)$$

Inserting the plane wave solution of Eq. (4.43), we immediately notice that $\nabla \cdot \mathbf{E} = 0$ because the gradient of $\mathbf{E}$ is only non-zero in the $\mathbf{k}$ direction but $\mathbf{E}$ must be perpendicular to $\mathbf{k}$. The remaining expression is

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\chi^{(1)}}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad (4.48)$$

which can be rearranged to give

$$\nabla^2 \mathbf{E} - \frac{1 + \chi^{(1)}}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0, \quad (4.49)$$

which is the wave equation, but with a wave velocity of $v = c/\sqrt{1 + \chi^{(1)}}$. Clearly the plane wave solution satisfies the differential equation; this justifies setting $\nabla \cdot \mathbf{E} = 0$. The only change is in the dispersion relation, which is given by $\omega = v|\mathbf{k}|$, where v can vary with ω because $\chi^{(1)}(\omega)$ is, in general, frequency dependent.

The index of refraction is defined as the ratio of the speed of light in vacuum to the phase velocity of the electromagnetic wave in a medium,

$$n = \frac{c}{v}. \quad (4.50)$$

Thus, for a linear material, the index of refraction is given by

$$n^2 = 1 + \chi^{(1)}. \quad (4.51)$$

Using the relation $\epsilon = 1 + \chi^{(1)}$, the index of refraction can also be written as

$$n = \sqrt{\epsilon} \quad (4.52)$$

A linear medium is equivalent to free space propagation but with a velocity of c/n . At the densities we work at, the partially ionized plasma has a nearly uniform index of refraction. A sum of plane wave solutions is a good first approximation for the propagation of light within the plasma-gas mixture.

The intensity is the power per unit area of the wave. It is given for a plane wave by

$$I = \frac{cn\epsilon_0}{2} |\mathbf{E}_0|^2 \quad (4.53)$$

This expression applies to most of the optical waves we work with if we replace $|\mathbf{E}_0|$ with the amplitude of the field.⁵

4.2.2 The Laser

Lasers produce the high intensity pulses of light needed to field ionize the gas in our PWFA plasma sources. A laser is a device that amplifies light through the mechanism of stimulated emission. Stimulated emission is the process by which a photon can drive an electron from a high energy state to a low energy state if the photon energy matches the energy difference between the two states. In the process, the electron will emit a photon as it transitions. The emitted photon will have the same polarization and phase as the stimulating photon. As a result, the electric fields of the two photons will add coherently. The medium within which the electronic transition occurs is called the gain medium; it can be a solid, liquid, or gas. Absorption will occur on the same transition if a large number of electrons exist in the low energy state. In order for emission to dominate, the transition must have a population inversion, more electrons in the high energy state than the low energy state. There are various ways of creating a population inversion, for details consult a laser textbook, such as Ref. [102]. Typically, only a small amount of energy is gained each time a group of photons passes through the gain medium. To produce significant amounts of energy, the gain medium is placed in a resonant cavity.⁶ Lasers are able to produce the extremely strong electric fields required to directly field ionize a gas into a plasma.

Typically, we model the optical pulse produced by a laser as a Gaussian temporal pulse and a super-Gaussian transverse pulse. Letting the pulse propagate in the z direction, the electric field

⁵ An exception is evanescent waves, which are oscillatory, but do not transport any energy.

⁶ Resonant cavities exist at the beginning of high-power systems. The laser pulse in a high-power amplifier, however, has so many photons, the pulse might only pass through the gain medium four or five times.

is given by

$$\mathbf{E}(x, y, t) = \mathbf{E}_0 \exp \left[-2 \ln 2 \frac{t^2}{\tau_p^2} - \left(\frac{r}{w_0} \right)^m \right] \exp [i(\phi_0 - bt^2 - \omega_0 t)], \quad (4.54)$$

where $\mathbf{E}_0$ is the electric field amplitude of the pulse,⁷ τ_p is the FWHM pulse length, w_0 is the super-Gaussian spot size, m is the order of the super-Gaussian and must be an even number, and b describes the chirp of the pulse. Including wavefront errors in the pulse is accomplished by making ϕ_0 a function of transverse position. The energy in the pulse is found by integrating over the intensity

$$\mathcal{E} = \frac{cn\epsilon_0}{2} \int dx dy dt |\mathbf{E}_0|^2 \exp \left[-4 \ln 2 \frac{t^2}{\tau_p^2} - 2 \left(\frac{r}{w_0} \right)^m \right]. \quad (4.55)$$

Evaluating the integral gives

$$\mathcal{E} = \frac{cn\epsilon_0 \pi^{3/2}}{2\sqrt{\ln 2}} \frac{4^{-1/m}}{m} \Gamma \left(\frac{2}{m} \right) |\mathbf{E}_0|^2 \tau_p w_0^2. \quad (4.56)$$

Typically the pulse energy, pulse length, and spot size are known for a given laser. Eq. (4.56) can then be used to find $|\mathbf{E}_0|$. The beam evolution in z can be found using the Fourier techniques presented in the next section.

The effect of optics on the pulse can be modelled by multiplying the electric field of Eq. (4.54) by a transmission function. For a spherical lens, the transmission function is a quadratic phase function:

$$T(r) = e^{-ikr^2/(2f)}, \quad (4.57)$$

where f is the focal length of the lens.

4.2.3 Fourier Optics

From the previous section, we know how to write down the electric field from the laser at $z = 0$ and add the phase imparted by a lens. To find how the field changes as the laser pulse propagates in z , we use Fourier optics. Any solution to the wave equation in free space, or in a linear isotropic medium, can be given as a sum of plane waves. The plane wave basis naturally

⁷ Here, $\mathbf{E}_0$ can be complex to encode the polarization of the pulse. Circularly polarized light will have a phase shift between the x and y components of the field. Because the pulse is propagating in z , $E_{0z} = 0$.

lends itself to calculation in Fourier space. We follow the angular spectrum approach of Goodman [103] and Ratcliffe [104]. Notice that the wave equation in free space, Eq. (4.49), is identical for each component:

$$\left(\nabla^2 - \frac{n^2}{c^2} \frac{\partial^2}{\partial t^2}\right) U = 0, \quad (4.58)$$

where U is called the field and can represent any of the three components of the electric field.⁸

The use of Eq. (4.58) is called the *scalar approximation*.

Imagine that we have a monochromatic wave propagating with some component of $\mathbf{k}$ in the z direction. The field on the $z = 0$ plane, $U(x, y, z = 0, t)$, is known and we want to find the field at all positions in z . If we evaluate the plane wave of Eq. (4.43) in the scalar approximation at $z = 0$ we get

$$U_k(x, y, 0, t) = U_0 e^{-i\omega t} e^{i(k_x x + k_y y)}, \quad (4.59)$$

which is a simple complex exponential in the transverse plane with a coefficient given by $U_0 e^{-i\omega t}$.⁹ The expression $e^{i(k_x x + k_y y)}$ is the basis function for Fourier space written in position space. This means that if we take the 2D Fourier transform of $U(x, y, 0, t)$, then we have decomposed the field into a sum of plane waves. The Fourier transform is then the amplitude and phase of each plane wave¹⁰

$$\hat{U}(k_x, k_y) = U_0(k_x, k_y) e^{-i\omega t}, \quad (4.60)$$

where the Fourier transform is defined as

$$\hat{U}(k_x, k_y) = \mathcal{F}_{xy}\{U(x, y, 0, t)\}(k_x, k_y) = \frac{1}{(2\pi)^2} \iint_{-\infty}^{\infty} dx dy U(x, y, 0, t) e^{-i(k_x x + k_y y)}. \quad (4.61)$$

The wave is monochromatic so we can write $U(x, y, z = 0, t) = e^{-i\omega t} U(x, y, 0, 0)$. To make things simple, let $U_{\perp}(x, y) = U(x, y, 0, 0)$, in which case the phase and amplitude of each Fourier compo-

⁸ Although the individual electric field components appear independent of each other, they are coupled to the magnetic field. Because of this, two electric field components can be dependent on each other indirectly through the magnetic field. The coupling only occurs in free space in the vicinity of charges and currents, i.e., near a boundary, and dies off over distances on the order of the wavelength. It is therefore unimportant for lasers [103].

⁹ In contrast to Eq. (4.43), U_0 is complex and includes the phase offset.

¹⁰ The electric field must be perpendicular to $\mathbf{k}$ in free space. The use of the scalar approximation here introduces an error because all possible value of K_x and k_y cannot have the same electric field direction. As long as the wave is paraxial, $K_x, k_y \ll k_z$, the error is negligible.

nent is

$$\hat{U}_\perp(k_x, k_y) = U_0(k_x, k_y). \quad (4.62)$$

The field is given in terms of its Fourier transform $\hat{U}_\perp(k_x, k_y)$, called the *angular spectrum*, by

$$U(x, y, 0, t) = e^{-i\omega t} \iint_{-\infty}^{\infty} dx dy \hat{U}_\perp(k_x, k_y) e^{i(k_x x + k_y y)}, \quad (4.63)$$

where the integral makes it clear the solution is a sum of plane waves.

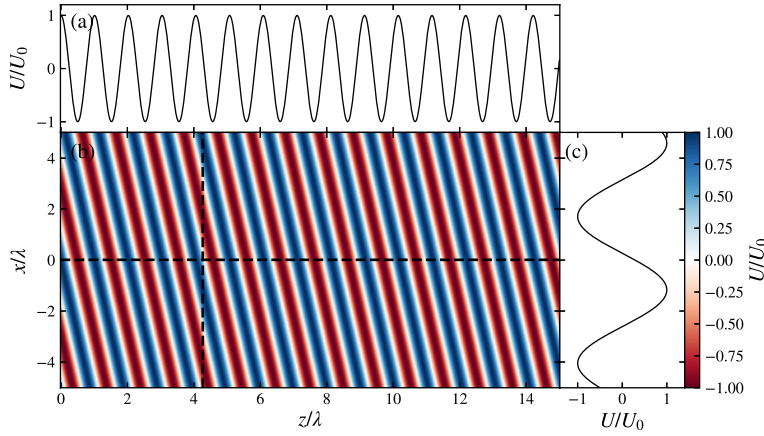


Figure 4.5: Illustration of the fundamental idea behind Fourier optics. The plane wave solution (b) to the wave equation can be sliced along a transverse plane (c). In this plane, the wave has a transverse frequency that depends on its angle of incidence. We can thus use the Fourier transform of the field on any transverse plan to decompose the wave into a sum of plane waves. Each plane wave picks up phase in z depending on its angle of incidence (a). The phase can be applied in Fourier space to solve for the field everywhere.

We can take advantage of the fact that we know how plane waves evolve in z to find $U(x, y, z, t)$. We can see from Eq. (4.43) that the plane wave of Eq. (4.59) evolves according to

$$U(x, y, z, t) = U(x, y, 0, t) e^{ik_z z}, \quad (4.64)$$

where $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$. Since Eq. (4.63) is a sum of plane waves with $z = 0$, the solution for all z is the same sum, but of the full z dependent plane waves

$$U(x, y, z, t) = e^{-i\omega t} \iint_{-\infty}^{\infty} dk_x dk_y \hat{U}_\perp(k_x, k_y) e^{izk_z} e^{i(k_x x + k_y y)}. \quad (4.65)$$

In the theory of linear, invariant systems, $H(k_x, k_y; z) = e^{izk_z}$ is called a transfer function. Using the definition of the inverse Fourier transform

$$\mathcal{F}^{-1}\{\hat{U}(k_x, k_y)\}(x, y) = \iint_{-\infty}^{\infty} dk_x dk_y \hat{U}(k_x, k_y) e^{i(k_x x + k_y y)} \quad (4.66)$$

the field can be written as

$$U(x, y, z, t) = e^{-i\omega t} \mathcal{F}^{-1}\{\mathcal{F}[U_{\perp}(x, y)]H(k_x, k_y; z)\}. \quad (4.67)$$

Using the transfer function, we can find the phase and amplitude of a laser beam at arbitrary z given the phase and amplitude on a transverse plane.

A finite duration laser pulse cannot be monochromatic. Propagating a pulse is a simple matter of applying Eq. (4.65) to each frequency component and summing the results together. The medium the wave is propagating through must be linear and isotropic, but it can be dispersive. Assume we start with a field given by $U(x, y, 0, t)$ as before, but now the time dependence is non-trivial. Our basis function in real space is now $e^{i(k_x x + k_y y - \omega t)}$ so the amplitude of each plane wave is given by the Fourier transform

$$\hat{U}(k_x, k_y, \omega) = U_0(k_x, k_y, \omega) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} dx dy dt U(x, y, 0, t) e^{-i(k_x x + k_y y - \omega t)} \quad (4.68)$$

As before, each plane wave is propagated by the transfer function $H(k_x, k_y, \omega; z) e^{izk_z(\omega)}$ where k can depend on ω through the dispersion function. The field at all locations and times is given by:

$$U(x, y, z, t) = \iiint_{-\infty}^{\infty} dk_x dk_y d\omega \hat{U}(k_x, k_y, \omega) e^{izk_z(\omega)} e^{i(k_x x + k_y y - \omega t)}. \quad (4.69)$$

The transfer function H is called the Rayleigh-Sommerfeld transfer function when it uses $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$. Applying the paraxial approximation $k_{x,y} \ll k_z$ gives $k_z = k^2 - \frac{1}{2}k_x^2 - \frac{1}{2}k_y^2$, which, when used in H , results in the Fresnel transfer function.

Fourier optics techniques lend themselves to numerical propagation of laser pulses. The Fourier transform is implemented numerically through the fast Fourier transform (FFT) algorithm, which allows for fast evaluation of the forward and reverse Fourier transforms. In addition, the spatial discretization only needs to resolve the highest transverse frequency rather than the laser

wavelength, like in a FDTD solution to Maxwell's equations. Laser propagation problems are typically paraxial and thus have large spatial frequencies allowing relatively coarse grids to be used. We will make extensive use of Fourier optics to simulate optics and as part of the algorithm for simulating laser propagation through plasma.

4.2.4 Nonlinear Optics

Nonlinear optics deals with light propagation through materials within which the polarization depends nonlinearly on the electric field. In these materials, the index of refraction is no longer sufficient to describe the impact of the material on the propagation of light. Laser ionization of a plasma is a special case of nonlinear optics. The polarization depends linearly on the electric field at a given plasma density, but the plasma density depends nonlinearly on the time history of the electric field, so the propagation of light is nonlinear. In this section, we will focus on the first type of nonlinearity, which appears in the optical systems we use as well as the plasma, although it is much less important in the plasma than the second type of nonlinearity. We generally follow the work of Boyd [105].

The full wave equation with the polarization terms is given by Eq. (4.45). Generally the spatial variation of the polarization is small and the $\nabla(\nabla \cdot \mathbf{P})$ term can be ignored. The wave equation is then given in the simplified form as

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{P}}{\partial t^2}. \quad (4.70)$$

Electromagnetic wave propagation in matter depends on how the polarization $\mathbf{P}$ is related to the electric field $\mathbf{E}$. We now have to be more careful with our complex representation of the electric field because the inhomogeneous wave equation is, in general, nonlinear.¹¹ The electric field can be written as a Fourier series

$$\mathbf{E} = \sum_n \mathbf{E}(\omega_n) e^{-i\omega_n t}, \quad (4.71)$$

¹¹ Treating the electric field as the real part of a complex number works because the homogeneous wave equation is linear. To show why, we can write the wave equation in terms of the d'Alembert operator $\square U = 0$. If we know that a complex number $U = a + ib$ is a solution: $\square(a + ib) = 0$, then $\square a + i\square b = 0$. The real and imaginary components must each satisfy the equation independently, so the real part is also a solution $\square a = 0$.

where $\mathbf{E}(-\omega_n) = \mathbf{E}^*(\omega_n)$, ensuring that the summation is real. All the spatial variation is included in $\mathbf{E}(\omega_n)$ terms. The polarization can also be expressed as a sum of frequency components

$$\mathbf{P} = \sum_n \mathbf{P}(\omega_n) e^{-i\omega_n t}. \quad (4.72)$$

Each term in the polarization can be expressed in terms of the electric field by [105] (vector Taylor series in $\mathbf{E}$)

$$P_i(\omega) = \epsilon_0 \left(\sum_j \chi_{ij}^{(1)}(\omega) E_j(\omega) + \sum_{jk} \sum_{(nm)} \chi_{ijk}^{(2)}(\omega; \omega_n, \omega_m) E_j(\omega_n) E_k(\omega_m) + \dots \right), \quad (4.73)$$

where $\chi^{(n)}$ is the n th order susceptibility tensor and (nm) means sum over all n and m such that $\omega_n + \omega_m = \omega$. In an isotropic material, the susceptibility tensors are diagonal with identical elements. If the material is lossless the tensors are real.

We are not concerned with the nonlinear processes of frequency conversion, but with distortions to the laser pulse that occur due to nonlinear effects. In particular, we will be concerned with two nonlinear processes: the Kerr effect and relativistic self-focusing.

The optical Kerr effect occurs when light is propagating through a medium with nonzero $\chi^{(3)}$. Again, we adopt the scalar approximation to simplify the math because we only consider isotropic media. In addition, we assume that $\chi^{(3)}$ is independent of frequency, a good approximation for the systems we consider, where the laser frequency is much lower than the resonant frequencies of the system. Consider a plane wave with frequency ω_0 : $U = \frac{1}{2}U_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega_0 t)} + \frac{1}{2}U_0 e^{-i(\mathbf{k} \cdot \mathbf{x} - \omega_0 t)}$. For the plane wave, $U(\omega_0) = \frac{1}{2}U_0 e^{i\mathbf{k} \cdot \mathbf{x}}$. The third term in Eq. (4.73) is

$$P^{(3)}(\omega) = \epsilon_0 \sum_{(onm)} \chi^{(3)} U(\omega_o) U(\omega_n) U(\omega_m), \quad (4.74)$$

where (onm) denotes the sum over all indices where $\omega = \omega_o + \omega_n + \omega_m$. Inserting our plane wave we find that the polarization has two non-zero, positive frequency components:

$$P^{(3)}(\omega_0) = \frac{3}{8} \epsilon_0 \chi^{(3)} U_0^3 e^{i\mathbf{k} \cdot \mathbf{x}} \quad (4.75)$$

$$P^{(3)}(3\omega_0) = \frac{1}{8} \epsilon_0 \chi^{(3)} U_0^3 e^{i3\mathbf{k} \cdot \mathbf{x}} \quad (4.76)$$

The negative frequency components are simply the complex conjugates, $P(-\omega_0) = P(\omega_0)^*$ and $P(-3\omega_0) = P(3\omega_0)^*$. The total polarization is given by Eq. (4.72). The two terms with the same frequency as the driving wave describe the Kerr effect. Including the first order susceptibility and the Kerr terms, the polarization is

$$P = \epsilon_0 \left(\chi^{(1)} + \frac{3}{4} \chi^{(3)} U_0^2 \right) U \quad (4.77)$$

Inserting P into Eq. (4.70) gives

$$\nabla^2 U - \frac{1}{c^2} \left(1 + \chi^{(1)} + \frac{3}{4} \chi^{(3)} U_0^2 \right) \frac{\partial^2 U}{\partial t^2} = 0. \quad (4.78)$$

we see that $\chi^{(3)}$ acts like an intensity dependent index of refraction. Using Eq. (4.53) we can write

$$n = n_0 + n_2 I = n_0 + \frac{3\chi^{(3)}}{4n_0^2 \epsilon_0 c} I, \quad (4.79)$$

where n_2 is the nonlinear index of refraction.

In high power lasers, the Kerr effect can cause significant distortions of the pulse if not accounted for. The center of the pulse, with intensity $I(z)$, picks up phase as it propagates through the medium according to

$$\phi = k \int_0^L n dz = k \int_0^L n_0 dz + k \int_0^L n_2 I(z) dz. \quad (4.80)$$

The edges of the pulse, where the intensity is small, only pick up the phase from the first term. The phase difference between the center of the pulse and the edges is called the *B-integral*. and is given by the second term

$$B = k \int_0^L n_2 I(z) dz. \quad (4.81)$$

Two things happen if the B-integral is large. Self-focusing occurs because the center of the pulse picks up more phase than the edges, acting like a positive lens. Self-phase modulation occurs because the temporal center of the pulse picks up more phase than the beginning and end, leading to a change in frequency along the length of the pulse. Both processes are best avoided in optical systems, a good rule is to keep $B < 1$ [102].

Plasmas inherently have $\chi^{(3)} = 0$ at low laser power because the plasma oscillations are linear in E (see Sec. 2.1.2). At high laser power, the oscillation amplitude becomes large enough that the electron velocities start to become relativistic. The result is an increase in the effective mass of the electron which changes the oscillation frequency. We derive the index of refraction in a plasma in Sec. 4.4, taking that result and making the substitution $m_e \rightarrow \gamma_b m_e$ gives the index with relativistic electron motion

$$n^2 = 1 - \frac{\omega_p^2}{\gamma_b \omega^2}. \quad (4.82)$$

The new factor, γ_b , however, changes throughout the oscillation period of the electron. In Refs. [106, 107], it is shown that an effective value can be used, and the nonlinear index of refraction can be written as

$$n_2 = \frac{\omega_p^2 e^2}{2n_0^2 \epsilon_0 m_e^2 c^3 \omega^4}, \quad (4.83)$$

where n_0 is the normal index of refraction of the plasma Eq. (4.112). The beam will again undergo self-focusing, now called relativistic self-focusing after the source of the nonlinearity.

4.3 Bessel Beams

PWFA experiments require plasmas with lengths close to a meter. Traditional spherical lenses produce short, narrow, high intensity spots with depths of focus given by the Rayleigh length $z_R = \pi w_0^2 / \lambda$. Here w_0 is the spot size of a Gaussian laser pulse at the focus. Long plasmas can be produced using spherical lenses with very long focal lengths, for a 40 cm plasma source the beam waist would be $320 \mu\text{m}$ [64]. This focusing system has three issues. First, the divergence of the beam is $w_0 / z_R = 0.8 \text{ mrad}$ requiring a focal length of 25 m, which significantly increases the footprint of the accelerator. The length can be improved by increasing the laser power. Increasing the power, however, leads to the second issue; the intensity near the focus of the spherical lens can be much larger than the intensity required to fully ionize the gas. The result is strong nonlinear effects that distort the pulse. Third, the plasma is dumbbell shaped, ionizing a much wider plasma near the ends than in the middle; this requires a large amount of laser energy.

A better focusing solution for long plasmas is to use Bessel like beams. Bessel beams are solutions to the wave equation in cylindrical coordinates. The size and intensity of the central focus spot remains constant in z . For this reason, they are often called non-diffracting beams. Just like a plane wave, they are only non-diffracting if the beam has infinite transverse extent. Realizable Bessel beams can only maintain a uniform longitudinal intensity over a finite distance. The easiest way to produce Bessel beams is to use an axicon lens. Essentially a cylindrical prism, an axicon imprints a linear phase in r on the beam. But the real usefulness of Bessel beams comes when trying to shape the longitudinal intensity of the focus, using a superposition of Bessel beams, we can create near arbitrary on-axis intensity profiles.

4.3.1 Wave Equation in Cylindrical Coordinates

In free space, under the scalar approximation, Maxwell's equations for the electric and magnetic fields reduce to the scalar wave equation,

$$\nabla^2 U - \frac{1}{c^2} \frac{\partial^2 U}{\partial t^2} = 0, \quad (4.84)$$

where U is the scalar field representing one component of either the electric or magnetic field. In cylindrical coordinates the Laplace operator is given by

$$\nabla^2 U = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r} \frac{\partial U}{\partial r} + \frac{1}{r^2} \frac{\partial^2 U}{\partial \phi^2} + \frac{\partial^2 U}{\partial z^2}. \quad (4.85)$$

Using separation of variables, $U(r, \phi, z, t) = U_r(r)U_\phi(\phi)U_z(z)U_t(t)$, the wave equation in cylindrical coordinates reduces to four ordinary differential equations,

$$\begin{aligned} k_r^2 r^2 \frac{\partial^2 U_r}{\partial (k_r r)^2} + k_r r \frac{\partial U_r}{\partial (k_r r)} + (k_r^2 r^2 - m^2) U_r &= 0 \\ \frac{\partial^2 U_\phi}{\partial \phi^2} &= -m^2 U_\phi \\ \frac{\partial^2 U_z}{\partial z^2} &= -k_z^2 U_z \\ \frac{\partial^2 U_t}{\partial t^2} &= -\omega^2 U_t, \end{aligned} \quad (4.86)$$

where $\omega = ck$ and $k^2 = k_z^2 + k_r^2$. The first equation is Bessel's equation in $k_r r$. The last three equations describe simple harmonic oscillators in ϕ , z , and t . Several boundary conditions constrain

the solutions. The fields must be finite at the origin, so only the Bessel functions of the first kind are retained. The field must be continuous in the ϕ direction, requiring $m \in \mathbb{Z}$. Additionally, since the beams propagate in the positive z direction, ω is assumed to be positive. Combining the solutions, we get the expression for a Bessel beam, the cylindrical coordinate equivalent to a plane wave

$$U = U_0 J_m(k_r r) e^{i(k_z z + m\phi - \omega t)} \quad (4.87)$$

The field structures of a $m = 0$ and $m = 2$ Bessel beam are shown in Fig. 4.6. Much like how the general solution in Cartesian coordinates (Eq. (4.69)) is a sum of plane waves, the general solution in cylindrical coordinates is a superposition of Bessel beams:

$$U(r, \phi, z, t) = \sum_m \int_{-\infty}^{\infty} d\omega \int_0^{\infty} k_r dk_r S_m(k_r, \omega) J_m(k_r r) e^{i(k_z z + m\phi - \omega t)}, \quad (4.88)$$

where $S_n(k_\rho, \omega)$ specifies the weighting of each Bessel function at each temporal frequency. As with Fourier optics, we are interested in how a beam specified on the $z = 0$ plane propagates in z .

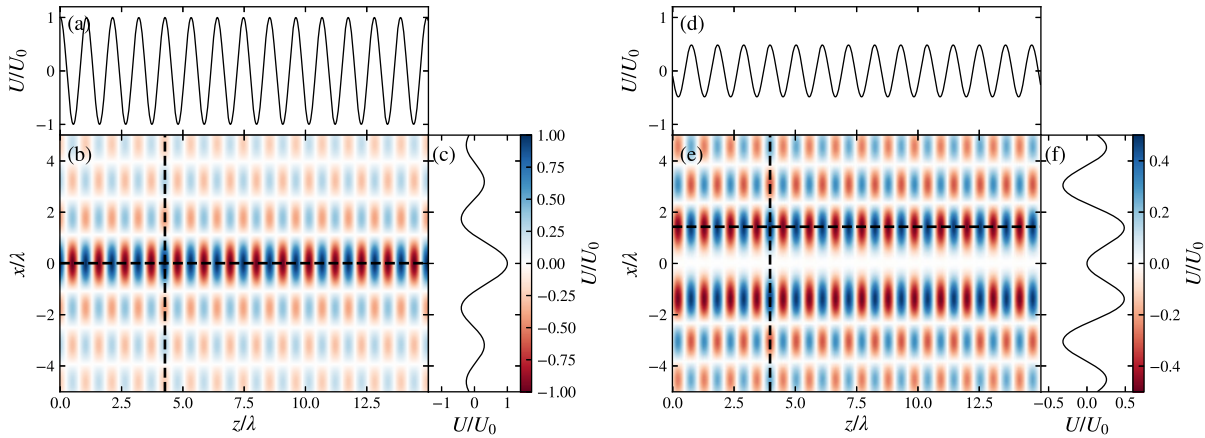


Figure 4.6: The electric field of a Bessel beam. An $m = 0$ beam is shown in (b), (a) is the field along $r = 0$ (c) is a transverse line-out showing the Bessel function. (e) shows an $m = 2$ Bessel beam with (d) being a longitudinal line-out at the r where the peak electric field is and (f) shows the transverse Bessel function.

We want to find $S_n(k_\rho, \omega)$ using the fields on the $z = 0$ plane. Just like in Fourier optics k_z is a function of k_r given by $k_z = \pm \sqrt{k^2 - k_r^2}$, the positive solutions correspond to waves propagating

in $+z$, the negative solutions in $-z$, and the imaginary solutions are evanescent waves. Starting with Eq. (4.88) and evaluating at $z = 0$ gives

$$U(r, \phi, z = 0, t) = \sum_m \int_{-\infty}^{\infty} d\omega \int_0^{\infty} k_r dk_r S_m(k_r, \omega) J_m(k_r r) e^{i(m\phi - \omega t)}. \quad (4.89)$$

The solution can be found by first multiplying both sides by $J_n(k'_r r) e^{-i(n\phi - \omega' t)}$ and integrating over r , ϕ , and t . The expression can be simplified using orthogonality relationships:

$$\int_0^{\infty} r dr J_n(k_r r) J_n(k'_r r) = \frac{\delta(k_r - k'_r)}{k_r}. \quad (4.90)$$

$$\int_0^{2\pi} d\phi e^{i(m-n)\phi} = 2\pi \delta_{nm} \quad (4.91)$$

$$\int_{-\infty}^{\infty} dt e^{-i(\omega - \omega')t} = 2\pi \delta(\omega - \omega') \quad (4.92)$$

The result is

$$\begin{aligned} \int_{-\infty}^{\infty} dt \int_0^{2\pi} d\phi \int_0^{\infty} r dr U(r, \phi, 0, t) J_n(k'_r r) e^{-i(n\phi - \omega' t)} \\ = (2\pi)^2 \sum_m \int_{-\infty}^{\infty} d\omega \int_0^{\infty} dk_r S_m(k_r, \omega) \delta(k_r - k'_r) \delta(\omega - \omega') \delta_{nm}. \end{aligned} \quad (4.93)$$

Solving for the spatial spectrum and making the substitutions $k_r \rightarrow k'_r$, $n \rightarrow m$, and $\omega' \rightarrow \omega$ gives

$$S_m(k_r, \omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} dt \int_0^{2\pi} d\phi \int_0^{\infty} r dr U(r, \phi, 0, t) J_m(k_r r) e^{-i(m\phi - \omega t)}. \quad (4.94)$$

The field at all z is given by Eq. (4.88).

For circularly symmetric beams, only the zeroth order Bessel function remains. Looking at a monochromatic beam, the spatial spectrum is given by

$$S(k_r) = \int_0^{\infty} r dr U(r, z = 0) J_0(k_r r), \quad (4.95)$$

which is the Hankel transform of $U(r, z = 0)$. The solution for arbitrary z reduces to

$$U(r, z, t) = \int_0^{\infty} k_r dk_r S(k_r) J_0(k_r r) e^{i(k_z z - \omega t)}. \quad (4.96)$$

Along the optical axis the field is

$$U(r=0, z) = \int_0^\infty k_r dk_r S(k_r) e^{i(k_z z - \omega t)}. \quad (4.97)$$

Changing the integration variable to k_z using $k_z = \sqrt{k^2 - k_r^2}$, and ignoring the evanescent waves we get

$$U(0, z) = \int_0^k k_z dk_z S(\sqrt{k^2 - k_z^2}) e^{ik_z z}. \quad (4.98)$$

The on-axis field is the inverse Fourier transform of the spatial spectrum evaluated at $\sqrt{k^2 - k_z^2}$ [108]. We will use this fact to create pulses with arbitrary on-axis intensity in Sec. 4.3.3.

4.3.2 The Axicon Lens

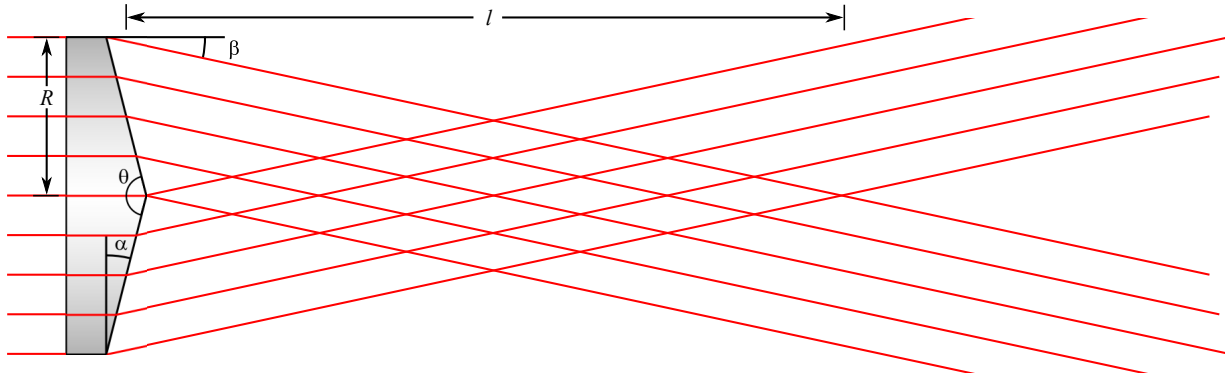


Figure 4.7: An axicon lens is described by the cone angle α and radius R . Parallel incident rays are deflected at an angle of β , producing a long focus of length l .

The easiest way to produce a quasi-Bessel beam is with an optic called an axicon lens.¹² An axicon is a cylindrical prism (a cone shaped lens). The geometry is defined by the angle of the prism α and the radius of the lens R ; Fig. 4.7 shows a schematic of a typical axicon. Sometimes the axicon is specified by the included angle of axicon tip $\theta = \pi - 2\alpha$ instead. The deflection angle of the rays β is determined from α and the index of refraction of the axicon. Using Snell's Law:

$$\beta = \arcsin\left(\frac{n_a}{n} \sin \alpha\right) - \alpha, \quad (4.99)$$

¹² We say quasi-Bessel because a true Bessel beam requires infinite transverse extent.

where n is the index of refraction of the medium surrounding the axicon and n_a is the index of refraction of the axicon. The length of the focus from the simple ray tracing picture is given by

$$l = \frac{R}{\tan \beta} \quad (4.100)$$

The intensity along the focus is not uniform, as an axicon does not produce a true Bessel beam. The intensity rises linearly from the lens before dropping off rapidly at the end of the focus. The linear growth can be easily understood from the ray optics picture using an energy argument. The rays in an annulus of radius r and thickness dr are all focused to $z = \frac{r}{\tan \beta}$ over a distance $dz = \frac{dr}{\tan \beta}$. Assuming uniform incident intensity on the axicon, the energy incident on an annulus as a function of r is $E \propto r dr$; therefore, the intensity on-axis is proportional to z . A numerical Fourier optics calculation of the intensity from an axicon is shown in Fig. 4.8 demonstrating the linear rise in intensity immediately after the axicon and the rapid reduction in intensity at the end of the focus.

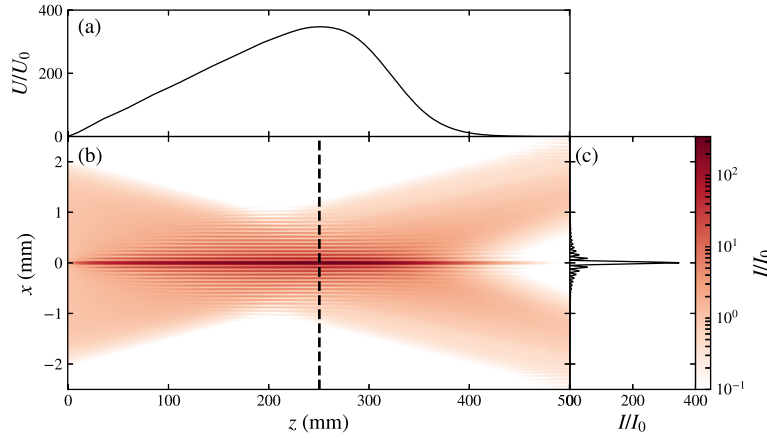


Figure 4.8: Intensity produced by an axicon lens illuminated by a super-Gaussian laser beam. The axicon has an angle of $\alpha = 0.7^\circ$ and is illuminated by a 4 mm diameter beam. The intensity grows linearly from zero immediately after the optic, before dropping off rapidly.

The transverse intensity profile of an axicon is a Bessel function. To understand why, consider the Fourier transform of a $m = 0$ Bessel beam on the $z = 0$ plane. The beam is circularly symmetric. It ends up that the 2D Fourier transform of a circularly symmetric function is the Hankel transform

of that function [103]:

$$\hat{U}(k_x, k_y) = \frac{1}{(2\pi)^2} \iint dx dy U_0 J_0(k_r r) e^{-i(k_x x + k_y y)} = \int_0^\infty r dr U_0 J_0(k_r r) J_0(k'_r r). \quad (4.101)$$

Using Eq. (4.90), the spatial frequencies of a Bessel beam are then

$$\hat{U}(k_x, k_y) = U_0 \frac{\delta(k_r - k'_r)}{k_r}, \quad (4.102)$$

where $k'_r = \sqrt{k_x^2 + k_y^2}$. In Fourier space, a Bessel beam is simply an infinitely thin ring of spatial frequency with radius k_r . Remember that the spatial frequency distribution corresponds to the distribution of k vectors in the plane wave decomposition of a given optical field. A ring corresponds to a cone of k vectors pointing inwards. An axicon produces exactly such a ring; the light rays are refracted into an inward facing cone all with the same angle relative to the optical axis.

In the thin lens approximation, an axicon can be described by a phase transmission function:

$$T(r) = e^{-ik \sin(\beta)r}. \quad (4.103)$$

Axicons were used heavily at FACET to create plasma sources longer than a meter; however, they suffer from several drawbacks. They provide no control over the ramps at the beginning and end of the plasma and they tend to be very energy inefficient.

4.3.3 Shaping the On-axis Intensity

Axicon lenses are able to produce a long enough and wide enough focus to satisfy our plasma requirements, but they provide no control over the intensity used to create the ramps. We want full control over the longitudinal intensity. This has been demonstrated with a low-power CW laser by Cizmar *et al.* [108], and later, with a low-power ultrashort laser by Ouadghiri-Idrissi *et al.* [109]. In this section, we describe how to calculate the phase and intensity the laser pulse must have upstream of the focus to produce a near arbitrary on-axis intensity profile. In Sec. 4.5, we describe a way to produce the required transverse intensity and phase. For now, we focus on finding the field at $z = 0$. The approach is straightforward. We already know the on-axis field of a general

circularly symmetric beam from Eq. (4.98). Assume we have some target intensity $I(z)$. Using Eq. (4.53), we can find the target electric field amplitude, $E(z)$. The full electric field will also have an oscillatory component that can, in general, be z dependent. Combining, the target electric field is

$$U_g(z) = E(z)e^{ik_{z0}(z)z}. \quad (4.104)$$

Inserting $U_g(z)$ into the left hand side of Eq. (4.98), we see that the on-axis field is the inverse Fourier transform of the spatial spectrum—up to a difference in integration bounds. If we assume the frequency components of U_g are negligible outside of the range 0 to k (we will come back to this later), we can Fourier transform both sides and solve for S

$$S(\sqrt{k^2 - k_z^2}) = \frac{1}{2\pi k_z} \int_{-\infty}^{\infty} E(z)e^{ik_{z0}(z)z} e^{-ik_z z} dz. \quad (4.105)$$

The oscillating term $k_{z0}(z)$ determines the width of the focus since $k_\rho = \sqrt{k^2 - k_z^2}$. The first root of J_0 occurs at $r = 2.4048/k_r$. The radial width of the main lobe (radius of the first zero) of the Bessel beam is approximately given by

$$R_b(z) \approx \frac{2.4048}{\sqrt{k^2 - k_{z0}(z)^2}} \quad (4.106)$$

For our plasma sources, we want a beam of uniform width, $k_{z0}(z) = k_{z0}$. In this case, we can use the Fourier shift theorem to simplify the spatial spectrum,

$$S(\sqrt{k^2 - k_z^2}) = \frac{1}{2\pi k_z} \hat{E}(k_z - k_{z0}). \quad (4.107)$$

The coordinates can then be transformed back to k_r

$$S(k_r) = \frac{1}{2\pi \sqrt{k^2 - k_r^2}} \hat{E}(\sqrt{k^2 - k_r^2} - k_{z0}). \quad (4.108)$$

The field on the boundary is given by inserting Eq. (4.108) into Eq. (4.89):

$$U(r, z = 0) = \mathcal{H}_0^{-1}\{S(k_r)\}, \quad (4.109)$$

from which the phase can be extracted from the argument of the field and the intensity using Eq. (4.53). If a target intensity pattern is chosen with frequency components outside of the range

$(0, k)$, then $S(k_r)$ can still be found with Eq. (4.108), and the required phase and intensity at $z = 0$ plane can be found using Eq. (4.109). However, for $k_r > k$ the waves are evanescent and will decay exponentially in z because k_z is imaginary. They will not contribute to the on-axis field. An example of an on-axis intensity profile is shown in Fig. 4.9 along with the radial phase and intensity required to create it.

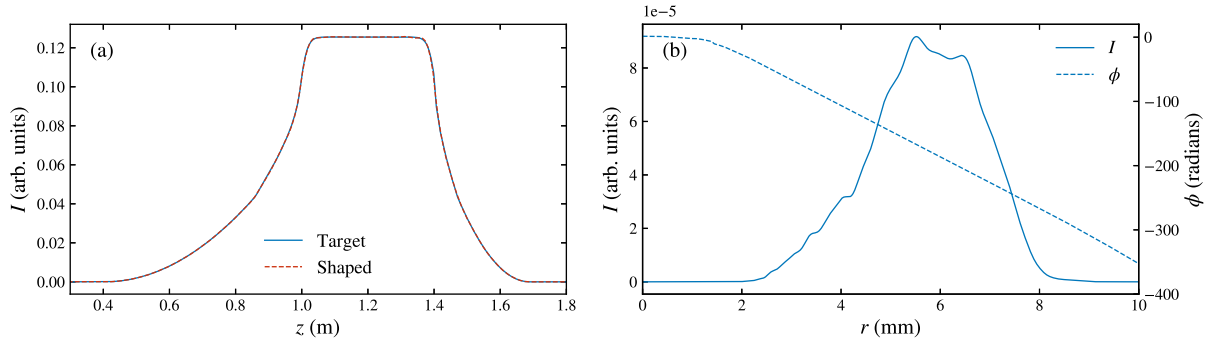


Figure 4.9: Example of shaping the on-axis intensity. (a) shows a typical intensity target for a PWFA plasma source. (b) plots the transverse phase and intensity required at $z = 0$ to produce the target intensity. The dash line in (a) is the result of numerically propagating the transverse field from (b).

4.4 Laser Propagation Through a Plasma

At this point, it seems we have everything we need to create a plasma source with controlled density ramps. We can take a target plasma density, calculate the required laser intensity, and shape the laser to produce this longitudinal intensity. Alas, it is not so simple. As the laser pulse propagates it interacts with the plasma modifying the intensity of the pulse.

The index of refraction of a plasma is less than the index of neutral gas. As the laser pulse traverses the gas, the front of the pulse begins to ionize the gas and form a plasma. The plasma changes the index of refraction experienced by the rear of the pulse. As a result, the rear of the pulse is defocused as the beam propagates. We call this process plasma refraction. It is inherently nonlinear because the index of refraction depends on the plasma density, and thus, the time history of the electric field at any given location.

4.4.1 Index of Refraction in a Plasma

The laser pulse refracts off the gas-plasma mixture because the plasma has a different index of refraction than the gas. We are interested in finding the index of refraction of the mixture as a function of ionization fraction. Recall that the index of refraction depends on the polarization in a linear medium. A simple model for the polarization is given in Jackson Sec. 7.5 [32]. The model treats the electron in the gas atom or plasma as bound to some point and driven by the electric field, the equation of motion is

$$m_e(\ddot{\mathbf{x}} + \gamma\dot{\mathbf{x}} + \omega_0^2\mathbf{x}) = -e\mathbf{E}(t), \quad (4.110)$$

where we have assumed the oscillation amplitude is small compared to the optical wavelength so we can ignore the $\mathbf{x}$ dependence of $\mathbf{E}$. Assuming the electric field is a linearly polarized plane wave, the field goes as $\mathbf{E} = \mathbf{E}_0 e^{-i\omega t}$. The particle position is given by the well-known solution for a driven damped harmonic oscillator:

$$\mathbf{x} = -\frac{e}{m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma} \mathbf{E}. \quad (4.111)$$

The dipole moment contributed by a single electron is then

$$\mathbf{p} = \frac{e^2}{m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma} \mathbf{E}. \quad (4.112)$$

If the electrons in the medium behave this way, then the material is clearly linear and isotropic. In a plasma, the electrons are not bound to any particular location so $\omega_0 = 0$. The dampening term is negligible as well $\gamma = 0$. If the electron density is n_e then the first order susceptibility is

$$\chi^{(1)} = -\frac{n_e e^2}{\epsilon_0 m_e} \frac{1}{\omega^2} = -\frac{\omega_p^2}{\omega^2}. \quad (4.113)$$

The index of refraction is given by Eq. (4.51)

$$n = \sqrt{1 - \frac{\omega_p^2}{\omega^2}}. \quad (4.114)$$

In a gas, the dipole moment from every electron in an atom has to be taken into account, each with its own ω_0 and γ . The sum of all the terms is called the atomic polarizability $\alpha(\omega)$

defined by

$$\mathbf{P} = \epsilon_0 n_g \alpha(\omega) \mathbf{E}. \quad (4.115)$$

The susceptibility is given by

$$\chi^{(1)} = n_g \alpha(\omega), \quad (4.116)$$

and the index of refraction is

$$n = \sqrt{1 + n_g \alpha(\omega)}. \quad (4.117)$$

The total index in a partially ionized plasma is given by adding the polarization vectors together

$$n = \sqrt{1 + n_g \alpha(\omega) - \frac{\omega_p^2}{\omega^2}}. \quad (4.118)$$

For the laser wavelength, plasma density, and gas species we use, the total index of refraction is very close to 1, so we can expand the square root to get

$$n(\omega) \approx 1 + \frac{1}{2} n_g \alpha(\omega) - \frac{\omega_p^2}{2\omega^2} = 1 + \frac{1}{2} n_g \alpha(\omega) - \frac{n_e e^2}{2\omega^2 \epsilon_0}. \quad (4.119)$$

4.4.2 Split Step Fourier Simulation

Solving for the beam evolution in the plasma requires finding solutions to the set of differential equations:

$$\nabla^2 U - \frac{n^2(\mathbf{x})}{c^2} \frac{\partial^2 U}{\partial t^2} = 0, \quad (4.120)$$

$$\frac{dn_g(\mathbf{x})}{dt} = -n_g(\mathbf{x}) w(|U(\mathbf{x}, t)|), \quad (4.121)$$

where $n(\mathbf{x})$ is related to n_g by Eq. (4.119) with $n_e = n_0 - n_g$. The spatial dependence of n depends on $\mathbf{E}$ non-trivially through the ADK formula, making the wave equation nonlinear. Fourier optics techniques can no longer be used to solve it directly. Instead, we can take advantage of the fact that the index of refraction is very close to 1. We can write the index as a constant with an $\mathbf{x}$ dependent perturbation.

$$n = n_h + n_{ih} = n_h - \frac{1}{2}(n_0 - n_g)\alpha - \frac{n_e e^2}{2\omega^2 \epsilon_0} \quad (4.122)$$

where $n_h = 1 + \frac{1}{2}n_g\alpha$. The wave equation can then be broken into two parts. In the first part, the index of refraction does not depend on $\mathbf{x}$, and it can be solved using Fourier optics. In the second part, the index is a small perturbation and can be treated as adding a real space dependent phase. The solution can be found by alternating between Fourier space and real space. The pulse envelope is propagated diffractively in Fourier space, and the spatially dependent index adds local phase in real space. As long as the spatial step is small enough that the envelope does not diffract significantly in a single step, and the change in ray angle does not result in too large of a transverse propagation, the error in the solution remains small.

The algorithm for the electric field proceeds iteratively

$$U(x, y, z + \Delta z, t) = \mathcal{F}_{txy}^{-1} \left\{ \mathcal{F}_{txy} \left[U(x, y, z, t) e^{ik_0 n_{ih}(t, x, y, z + \Delta z/2) \Delta z} \right] e^{ik_z \Delta z} \right\}, \quad (4.123)$$

where k_z is a function of ω , k_x , and k_y . Now, we need to include the ionization of the plasma. To do this, the rear part of the pulse needs to be able to see the index of the plasma created by the front part of the pulse. By ignoring dispersion, which is a secondary effect to refraction, we can rewrite the algorithm as

$$U(x, y, z + \Delta z, t) = \mathcal{F}_{xy}^{-1} \left\{ \mathcal{F}_{xy} \left[U(x, y, z, t) e^{ik_0 n_{ih}(t, x, y, z + \Delta z/2) \Delta z} \right] e^{ik_z \Delta z} \right\}. \quad (4.124)$$

Since t is no longer part of the Fourier transform, we can solve for $U(x, y, z, t_0)$ using $n_{ih}(t_0)$. We can then numerically integrate Eq. (4.121) to find $n_{ih}(t_1)$, and use this to solve for $U(x, y, z + \Delta z, t_1)$. A diagram of the algorithmic structure is shown in Fig. 4.10.

4.4.3 Optimizing the Plasma Width

The refraction of the rear of the pulse off the plasma formed by the front of the pulse can have a significant impact on the ionization process. The plasma reduces the index of refraction, leading to a defocusing effect on the rear of the pulse. Unsurprisingly, this will reduce the peak intensity, but it also tends to widen the focus. The reduction in peak intensity will change the ionization fraction in the plasma density ramps, modifying the ramp shape. In the bulk, the reduction in

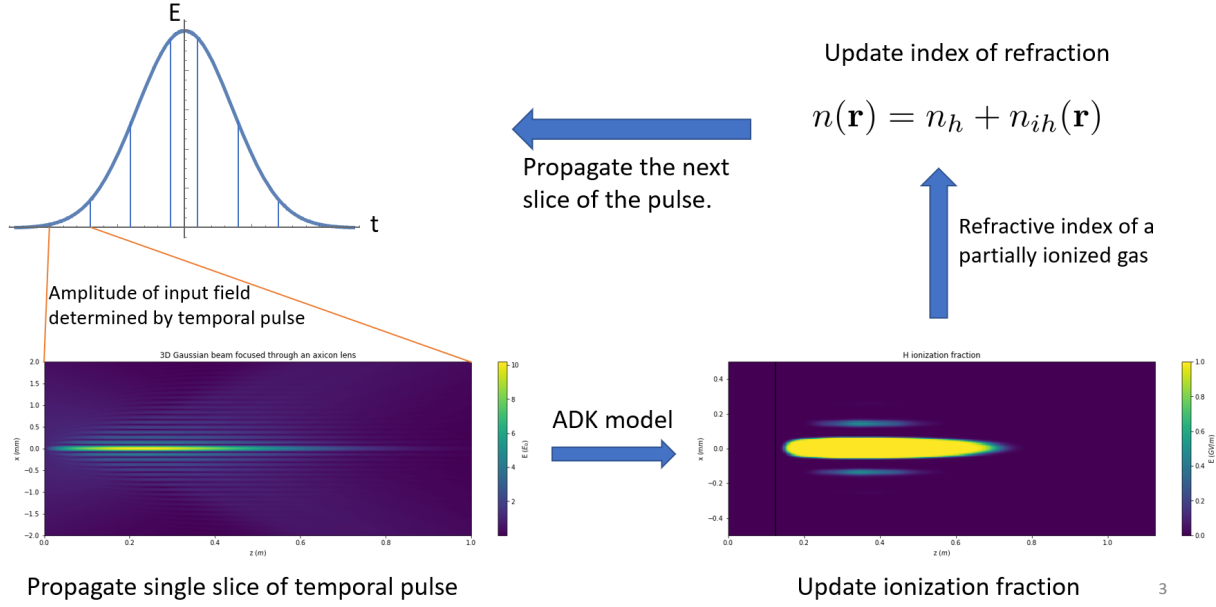


Figure 4.10: Structure of the split-step Fourier algorithm for simulating a laser ionized plasma. The pulse is divided up temporally and the first slice is propagated through the gas using the split step technique. The ionization from the first slice is calculated and the index of refraction is updated. The next slice is propagated through this new index of refraction and the algorithm continues iteratively.

intensity can be significant enough to prevent 100% ionization. At the same time, the widening of the focus results in a wider plasma and can be advantageous for reaching the required plasma width.

At our target density of $3.4 \times 10^{16} \text{ cm}^{-3}$, the plasma should be $239 \mu\text{m}$ wide. Naively, we would think to use a Bessel beam with $R_b = 120 \mu\text{m}$. If we did, we would find that the plasma would not be fully ionized but would be much wider than our requirement. Reducing R_b will increase the plasma density, while reducing the diameter. To show this, we ran a series of refraction simulations using the longitudinal intensity profile for the E300 optic, detailed in Sec. 4.7.2. We varied R_b at two different energies to see how the plasma changed. The results are shown in Fig. 4.11. As expected, the plasma width increases with R_b , while the plasma density is reduced.

We see a general trend across all the simulations we have run. If the width of the Bessel R_b starts small, then there is minimal refraction, and the plasma will be ionized as expected from ADK

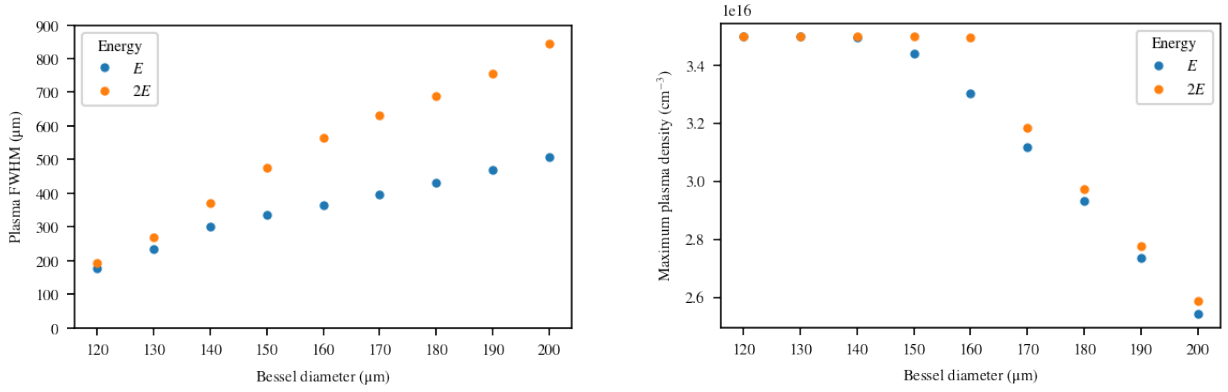


Figure 4.11: Properties of a laser ionized plasma formed in a $3.5 \times 10^{16} \text{ cm}^{-3}$ lithium vapor. The left plot shows the FWHM of the plasma as a function of the width of the central Bessel peak, $2R_b$. The right plot shows the peak plasma density as a function of Bessel width. Refraction widens the laser focus resulting in a plasma that is wider than the Bessel peak. As the Bessel width is increased, refraction lowers the peak intensity and prevents the plasma from being fully ionized.

using the un-refracted intensity. Increasing R_b results in increased refraction leading to a wider, but lower intensity, pulse. At some point, the pulse intensity drops too low to fully ionize, but the width of the partially ionized plasma continues to increase, see Fig. 4.11. Increasing the energy does little when R_b is small—the plasma stays fully ionized. Surprisingly, at large R_b , the peak intensity, and thus ionization fraction, is almost independent of the pulse energy. The increase in energy only causes the refraction and defocusing to occur earlier in the pulse, but the peak intensity remains unchanged. Instead, increasing the energy causes the plasma to get even wider as the rear of the pulse is significantly refracted. Both effects can be seen in the orange curves of Fig. 4.11. There is an optimal R_b for each combination of density and gas species, where the plasma is still fully ionized, and the width is maximized. For lithium gas at $3.4 \times 10^{16} \text{ cm}^{-3}$ the optimum is about $R_b = 60 \mu\text{m}$.

4.5 Tandem Lens Optics

Given a target longitudinal intensity profile, Eq. (4.109) tells us how to find the laser's transverse intensity and phase at a plane upstream of the target. In this section, we present the optical technique we chose to produce such a laser pulse. Axial intensity shaping has recently

been demonstrated experimentally [108, 109]. These demonstrations used a phase only spatial light modulator (SLM) to produce a hologram. When combined with Fourier spatial filtering at the focus of a lens, phase only SLMs can create a pulse with shaped phase and intensity. The laser fluence required for field ionization, however, is too large for any existing SLM. Even if an SLM could be used, the efficiency of the technique is very low. Most of the laser energy ends up in diffraction orders that are rejected by the spatial filter. We turn to an optical technique proposed by Honkanen and Turunen [110] called the tandem lens. As the name implies, the optical system is composed of a pair of lenses. The first lens is placed upstream of the second lens and shapes the intensity on the second lens so that it matches the target intensity. The second lens removes the residual phase left by the first lens and adds the target phase. A schematic of the tandem lens technique is shown in Fig. 4.12. By solving for the transverse field using a superposition of Bessel beams, we fully account for diffractive effects and significantly improve the performance of the optics compared to the original idea.

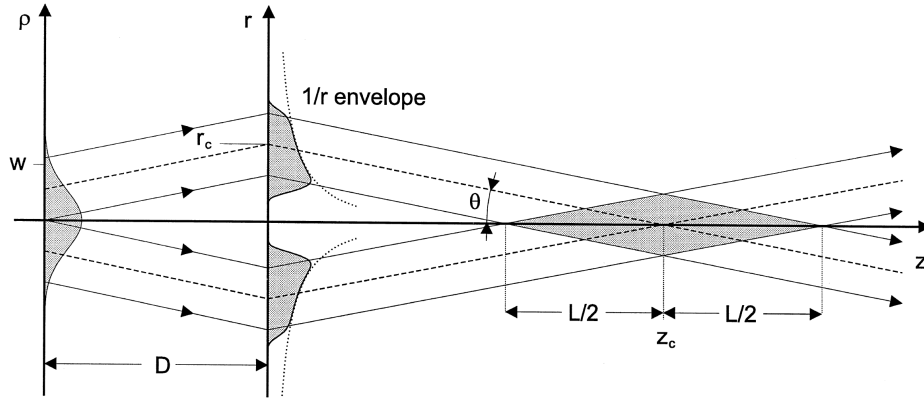


Figure 4.12: Principle of operation of the tandem lens. The first lens shapes the intensity onto the second lens. The second lens removes the residual phase from the first lens and adds the required phase. Reproduced from Ref. [110].

4.5.1 Phase of Each Lens

We use simple ray tracing to design the phase of the first lens. The goal is to create a lens that transforms one intensity profile into another over a given distance L . Call the plane a distance

L from the lens the output plane. The intensity incident on the lens is the input intensity and the intensity on the output plane is the output intensity. The lens design problem is then to find the ray deflection as a function of radius, $\theta(r)$. From there, it is straightforward to find the phase of the lens [111]. This problem is equivalent to mapping each radius on the lens, r , to a radius on the target plane, r' . We can do this mapping using geometric energy conservation which can be stated as

$$2\pi I_{in}(r)rdr = 2\pi I_{out}(r')r'dr'. \quad (4.125)$$

This equation represents the mapping of energy from an annulus of thickness dr on the lens plane to an annulus of thickness dr' on the output plane. The two radii are related by the ray angle according to

$$r' = r + L \tan[\theta(r)]. \quad (4.126)$$

Inserting into Eq. (4.125) gives a differential equation for $\theta(r)$

$$\frac{d\theta}{dr}(L^2\theta + rL) + L\theta + r \left(1 - \frac{I_{in}(r)}{I_{out}(r + L\theta)}\right) = 0 \quad (4.127)$$

The boundary condition is slightly difficult to determine. If both lenses have non-zero intensity at $r = 0$ then the boundary condition is $\theta(0) = 0$. For the tandem lenses, I_{out} is often an annular profile. The initial value of θ must map the center of the input lens to the inner radius of the annulus, where there is appreciable intensity. The differential equation must be integrated numerically so “appreciable” intensity in this case depends on the details of the step size and the intensity profile.

In practice, we use a different numerical technique, based off Eq. (4.125), to avoid these issues. We start with I_{out} and r' on a grid and calculate the corresponding r for each r' . We start by writing Eq. (4.125) as single terms in a Riemann sum. We need the result to depend only $I_{in}(r_{i-1})$ and not r_i , as we do not know r_i , it is what we are solving for. We use a left Riemann sum for I_{in} so it only depends on r_{i-1} and a middle Riemann sum for everything else

$$I_{in}(r_{i-1})\frac{1}{2}(r_i + r_{i-1})(r_i - r_{i-1}) = \frac{1}{2}(I_{out}(r'_i)r'_i + I_{out}(r'_{i-1})r'_{i-1})(r'_i - r'_{i-1}). \quad (4.128)$$

The input radius is then given by

$$r_i^2 = (I_{out,i}r'_i + I_{out,i-1}r'_{i-1}) \frac{\Delta r'}{I_{in}(r_{i-1})} + r_{i-1}^2, \quad (4.129)$$

where we have assumed we can evaluate I_{in} at arbitrary r , either through interpolation or analytically. We now have a recursion relation for r_i starting from $r_0 = 0$ and r'_0 . r'_0 is chosen so that the intensity is appreciable at r'_0 . Each ray goes from r_i to r'_i , so the ray angle is given, in the paraxial limit, by

$$\theta(r_i) = \frac{r'_i - r_i}{L} \quad (4.130)$$

Once the ray angles are known, it is straightforward to find the phase function of the lens. The derivative of the phase is related to the ray angle according to

$$\frac{d\phi(r)}{dr} = -\sin[\theta(r)]. \quad (4.131)$$

The phase of the first tandem lens is found by integrating Eq. (4.131).

The phase imparted by the first lens at the output plane is found by taking the incoming beam profile, applying the phase of the first lens, and propagating to the second lens. The phase of the second lens is then the difference between the target phase and the incoming phase. Fig. 4.13 shows the tandem lens phases required to produce the transverse intensity and phase from Fig. 4.9(b), with $L = 1$ m. The distance between the lenses is a free parameter, but numerical tests show it does not have a significant impact on lens performance if it is sufficiently long to keep all the rays paraxial while remaining short enough that diffraction is not significant.

4.5.2 Fabrication

The tandem optics can be created either using refractive or diffractive lenses. Because the lenses are rotationally symmetric, they could be created as a refractive optic using single point diamond turning. The surface relief of the optic would follow the phase profile. From Fig. 4.13, we can see that the first tandem lens would have a sharp corner at the center that would be difficult to fabricate with a turning process. We could of course ignore this section when designing the

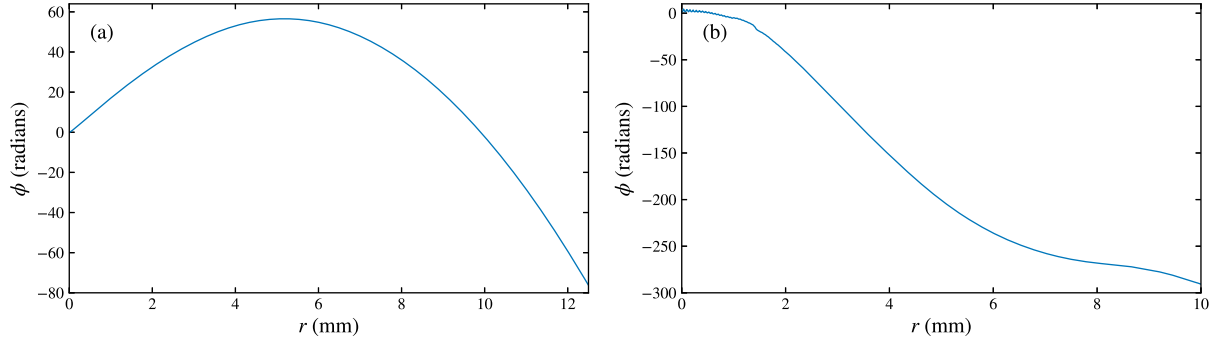


Figure 4.13: Phases of both tandem lenses. (a) is the first lens, (b) is the second. The lenses are designed to produce the transverse phase intensity profile of Fig. 4.9(b) from an incoming beam with a super-Gaussian intensity profile defined by $w_0 = 12$ mm and $n = 8$.

phase of the first lens, but this would lower the efficiency. The lens also must be thin to limit the self-phase modulation introduced into the laser pulse. Thin lenses are difficult to turn without introducing distortion into the lens, but the primary difficulty with refractive optics is the expense to manufacture. For this reason, we use diffractive optics for the tandem lenses.

Diffractive optics work by exploiting the wave nature of light. In our case, we will be using a binary optic, also called a kinoform. The surface profile of the optic is found by taking the phase of the lens modulo 2π , and then dividing it into 2^m uniform, discrete steps. The discretized phase profile forms the surface profile of the optic. The physical height of each step corresponds to the difference in material thickness required to produce a $2\pi/2^m$ phase change. Assuming the material has index of refraction n , the height of each step is

$$d_{step} = \frac{\lambda}{2^m(n-1)}. \quad (4.132)$$

For an 800 nm laser passing through an 8 layer optic made of fused silica, $n = 1.4533$, the step height is 220 nm. Etch depths of several 100 nm in fused silica are within the standard range of reactive ion etching (RIE) techniques.

The disadvantages of a refractive optic are twofold: efficiency and wavelength sensitivity. The discretization of the phase causes some amount of light to be transmitted into higher diffraction orders. In the language of Fourier optics, the discretized phase profile has spatial frequency

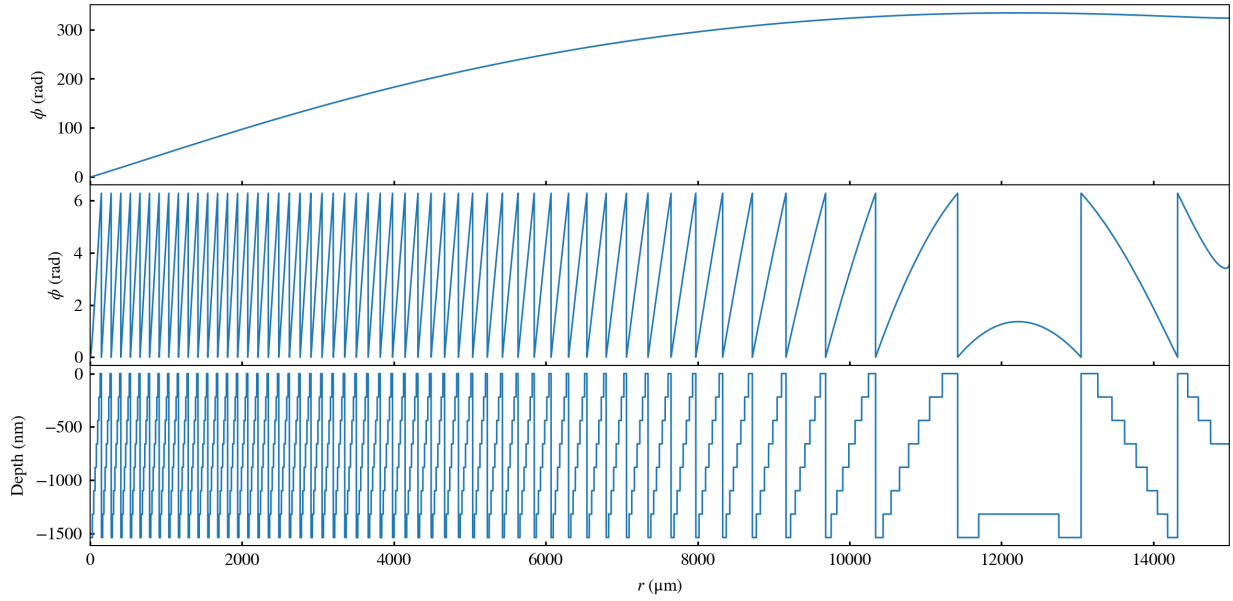


Figure 4.14: The surface profile of the binary optic is created by taking the lens phase modulo 2π (top plot to center plot) and discretizing into 2^m layers (center to bottom plot). In this case, $m = 3$ and the lens has been discretized into 8 layers. The depth of each layer depends on the index of refraction of the lens material according to Eq. (4.132).

harmonics at the spatial frequencies of the discrete steps. The first harmonic has a transverse wavenumber k_r corresponding to the width of the steps, there are also higher harmonics at half the step size and up. These correspond to the first and higher diffraction orders. Light in the first and higher diffraction orders has a larger spatial frequency than the light we want in the 0th order. The efficiency of the optic is the fraction of the incoming laser power in the 0th diffraction order. It increases with increasing m , at $m = 3$ it is over 95% for the design wavelength [112].

The step height of Eq. (4.132) depends on λ , called the design wavelength. Other colors of light will not have perfect 2π phase wrapping and will have reduced diffraction efficiency and a phase error. Looking at different colors from a ray optics perspective, the angle of the ray after the optic will be shifted proportional to the frequency. For detailed calculations of the focal length shift with frequency, see Refs. [113].

We attempted to make our own binary optics in the University of Colorado's nanofabrication lab, called the Colorado Shared Instrumentation in Nanofabrication and Characterization (COS-

INC). The fabrication process starts by laser writing a mask for each etch step. For our optics, we use $m = 3$ etches, which produces $2^3 = 8$ layers. The masks start as 4 in squares of 0.06 in thick soda lime glass with a 100 nm layer of chrome and then a 530 nm layer of AZ1518 photoresist deposited onto one side. The masks are loaded into a direct laser writer (Heidelberg DWL 66FS) where an ultraviolet laser exposes the photoresist in the regions that will be etched. Alignment marks and serial numbers are also exposed during this step. The mask is then submerged in a developer solution to dissolve the exposed resist. The mask is rinsed off, submerged in an acid bath to etch away the chrome that is no longer covered by resist, and rinsed again before entering an ultrasonic solvent bath that removes the undeveloped resist. Finally, the mask is rinsed, dried, and inspected.

The optics start as 1 mm thick polished wafers of fused silica, 2 in in diameter. Each optic is rinsed in solvent and then water to clean it before fabrication. After cleaning, the wafers are heated on a hot plate at 110° C for 120 s to drive off moisture. The wafers are then spin coated with HMDS, an adhesion promoter, at 6000 RPM for 30 s and thermally activated by heating at 110° C for 120 s to improve the HMDS performance. The wafers are then spin coated with AZ9260 photoresist at 6000 RPM for 30 s producing a $5.7\ \mu\text{m}$ thick coating for the two shallow etches. For the deepest etch, the wafers are coated twice at 4000 RPM for 30 s with a softbake of 110° C for 180 s in between to help evaporate the solvent from the first coating. The resulting photoresist is $16.6\ \mu\text{m}$ thick. The wafers are then heated at 110° C for 180 s to evaporate the rest of the solvent from the resist. The wafers are allowed to rest for 30 min to re-hydrate and prevent cracking of the photoresist layer.

The coated wafers are exposed through the masks using a Karl Suss MJB3 mask aligner. The mask and wafer are placed into the machine and the mask is aligned to any previous etches manually through a microscope. The mask aligner then clamps the mask to the substrate using vacuum and exposes the backside of the mask to the light from an ultraviolet lamp. The thin photoresist layers were given 3.5 min of exposure, the thick layers 5 min. After exposure, the wafers are submerged in a container of developer for 2.5 min and 5 min for the thin and thick layers respectively. The

developer dissolves the exposed photoresist. The wafers are then rinsed, dried, and heated on a hot plate at 90° C for 180 s. The wafers are then ready for the etch.

Wafers are etched in a reactive ion etcher using CF_4 and O_2 process gases at 14 sccm and 4 sccm, respectively. The chamber is kept at 180 mTorr with 300 W of RF power used to drive the plasma. The etch rate for this recipe is approximately 36 nm/min. The remaining photoresist is stripped from the wafers in a solvent ultrasonic bath. The wafers are then rinsed and dried, and the process starts over again for the next etch. The etches are performed in order of increasing depth starting with a 220 nm etch, then a 440 nm etch, and finishing with a 880 nm etch. A profilometer trace after the first two etches is shown in Fig. 4.15. The measured etch depth of the second etch was 430 ± 52 nm over 2 mm of the sample demonstrating good control of the etch depth, but less than ideal etch uniformity.

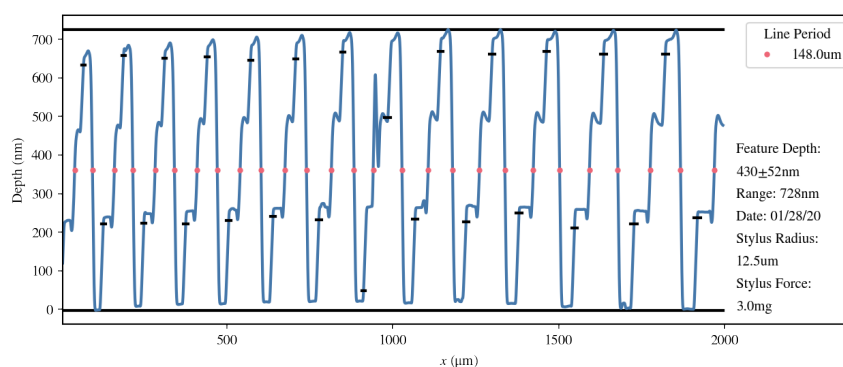


Figure 4.15: Profilometer trace of a binary optic after performing a 220 nm etch and a 440 nm etch. The feature depth is the spacing between the black lines, and measures the depth of the second etch. The surface relief has the stair step pattern expected from a binary optic.

Two issues plagued our attempts to make optics at COSINC. First, the RIE etcher produced inconsistent results. It struggled to maintain process parameters, and as a result, the etch depths and selectivity had large variation. Second, the mask aligner was in a state of disrepair making it almost impossible to align multiple layers. The first etch is only a 220 nm deep feature in a clear substrate. Resolving the etch under the mask aligner's microscope was exceedingly difficult. The result was a very low success rate with aligning multiple etches. In the end we had the optics

manufactured by an external company, NILT Technologies. Microscope images of the binary optics are shown in Fig. 4.16.

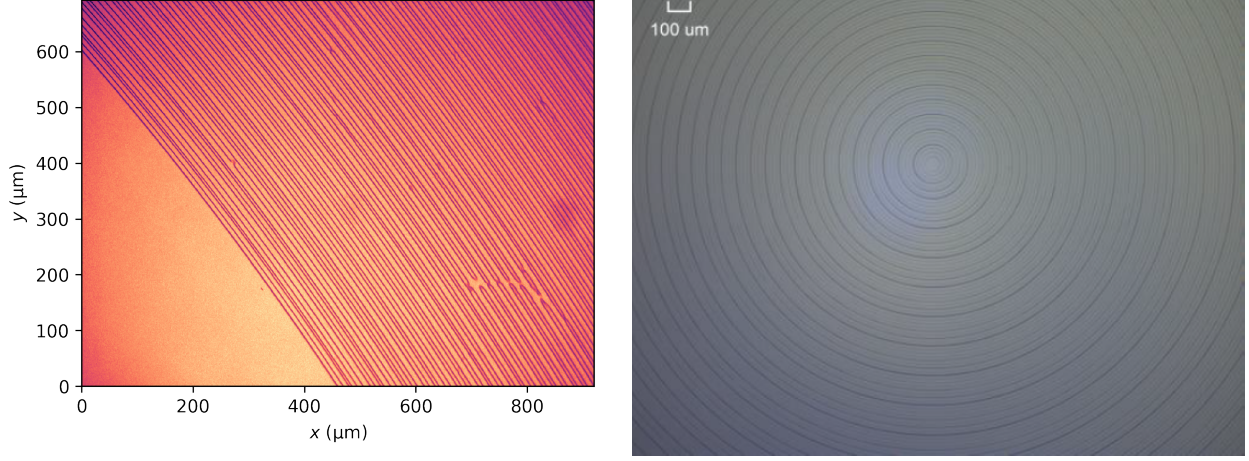


Figure 4.16: Microscope images of etched binary optics. Left is a false color image of a single etch carried out in the nanofabrication lab at University of Colorado. Right is an image of a purchased diffractive optic.

4.5.3 E301 Plasma and Optics

As an example of the optics design process presented in the previous sections, consider a plasma for the E301 experiment at FACET-II. E301 is the designation for plasma wakefield experiments using a laser ionized gas in a flooded chamber as the plasma source. The plasma source is designed for the beam parameters presented in Sec. 5.2.2. The expected FACET-II beam parameters have changed slightly since we defined the plasma source requirements. For E301, the initial design uses a $3.5 \times 10^{16} \text{ cm}^{-3}$ bulk plasma with a length of 48 cm. The ramps on both ends have a Gaussian profile. Numerical optimization was performed separately for both the x and y direction to determine the optimal ramp length and ramp location relative to the focus of the electron beam (called the interaction point or IP). We cannot have different ramps in both transverse directions, so the average of the two solutions was taken. The resulting entrance ramp has a length of $\sigma_{ent} = 4.9 \text{ cm}$, with the top of the ramp located 15 cm downstream of the IP. The exit ramp length is less critical and was chosen to have $\sigma_{exit} = 6 \text{ cm}$.

To find the target on-axis intensity, we started by finding the ionization fraction as a function of peak intensity for a $\tau_p = 70$ fs laser pulse using Eq. (4.5), with the ionization rate given by the time averaged ADK formula of Eq. (4.11) and (4.33). The gas species is helium to minimize beam ionization. We then numerically inverted the ionization fraction-intensity relation to get the peak intensity as a function of ionization fraction. Using this, we converted the target plasma density into a target laser intensity. We do not want the laser to have barely enough intensity to fully ionize in the bulk, so we modify the intensity target by doubling the intensity in the fully ionized part of the plasma. The intensity ramps up from the top of the ramps to ensure the overall target profile is smooth. The final target intensity profile is shown in Fig. 4.17(d) where the tops of the ramps are clearly visible.

The longitudinal intensity is converted into the required transverse phase and intensity of the laser at the second tandem optic using Eqs. (4.108) and (4.109). The Bessel width was chosen to be $R_b = 75 \mu\text{m}$, which maximizes the plasma width. Fig. 4.17 shows the transverse intensity at the second tandem optic. The phase of both tandem lenses is found using the procedure outlined in Sec. 4.5.1. The locations of the lenses relative to the plasma are dictated by the layout of the IP area at FACET-II, the second lens is 1.75 m upstream of the IP and the first lens is 1 m upstream of the second. For more details of the IP layout see Sec. 5.2.3. The phases of the two lenses are shown in Fig. 4.17(b) and (c).

The actual plasma formed will not match the target plasma due to refraction during the ionization process. We ran a split-step Fourier simulation of the ionization process to see how the ramps were distorted. The results are presented in Fig. 4.17(f). The distortion is minimal in the entrance ramp but is significant near the end of the plasma and in the exit ramp. We calculated the emittance growth through the plasma in order to see how well the actual plasma source would perform. We included energy gain using the simple energy gain model of Eq. (3.76) with $E_0 = 18.7 \text{ GV/m}$ determined from PIC simulation and $E_1 = 0$. The resulting emittance evolution for both the x and y directions is shown in Fig. 4.18. The emittance growth is only 4% in both directions, which is about the minimum possible due to the difference in waist position

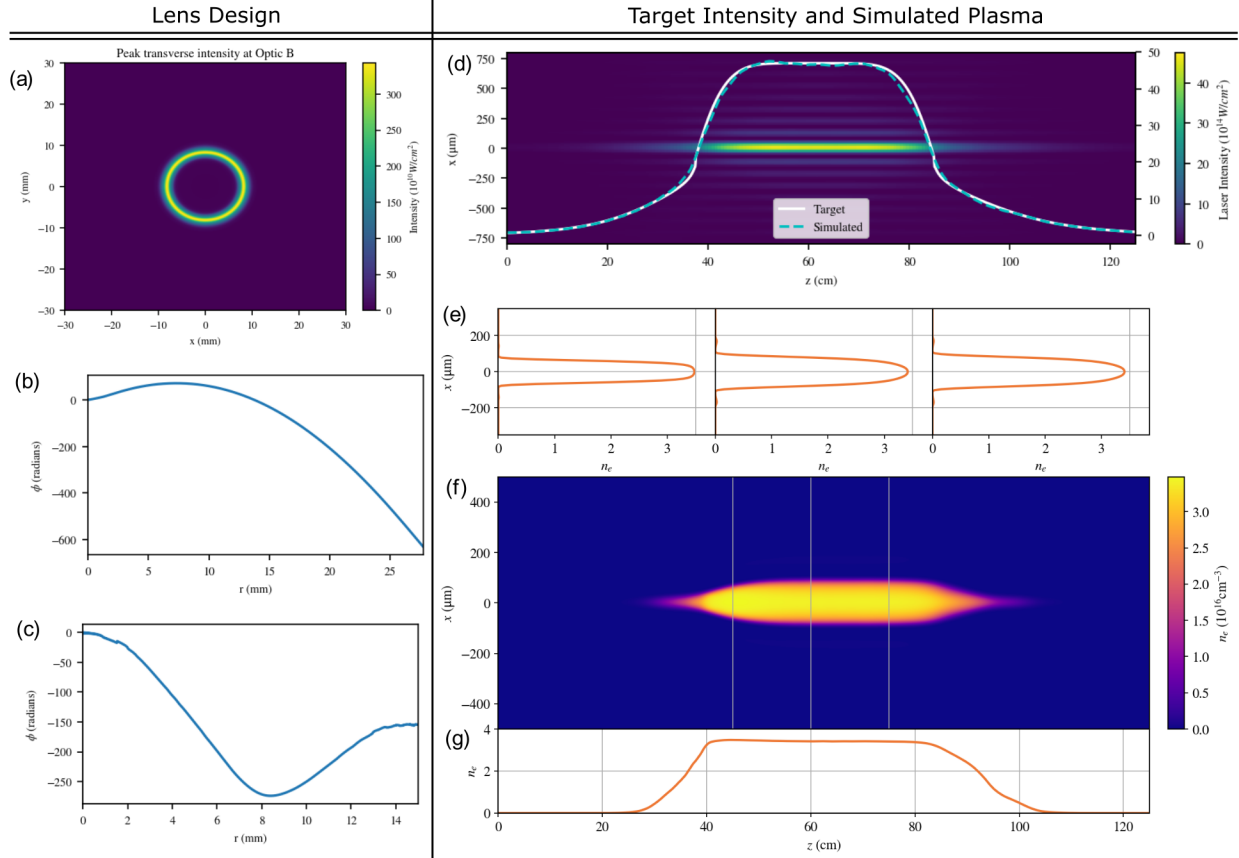


Figure 4.17: Initial lens design for E301. (a) peak laser intensity at lens B as a result of the phase imparted by lens A. (b) phase imparted by lens A. (c) phase imparted by lens B. (d) simulated laser intensity in vacuum, the lineouts show the on-axis intensity from simulation (dashed) and the target (solid). (e) transverse plasma density at the start, middle, and end of the bulk plasma. (f) simulated plasma density including refraction. (g) longitudinal plasma density on-axis.

between the two directions.

4.6 The Problem of the Electron Beam

We worked out the electric field produced by the electron beam in Sec. 2.3.5. Using Eq. (2.174) to calculate the fields from the FACET-II drive beam, we see that the beam easily has fields on the order of 100 GV/m when matched to the plasma. Such a strong field is more than sufficient to field ionize helium. The problem is made worse by the fact that the location of the drive beam focus is downstream of the witness beam focus. As a result, both the drive and witness beam can be simultaneously matched to the plasma. For emittance preservation, the witness beam must be

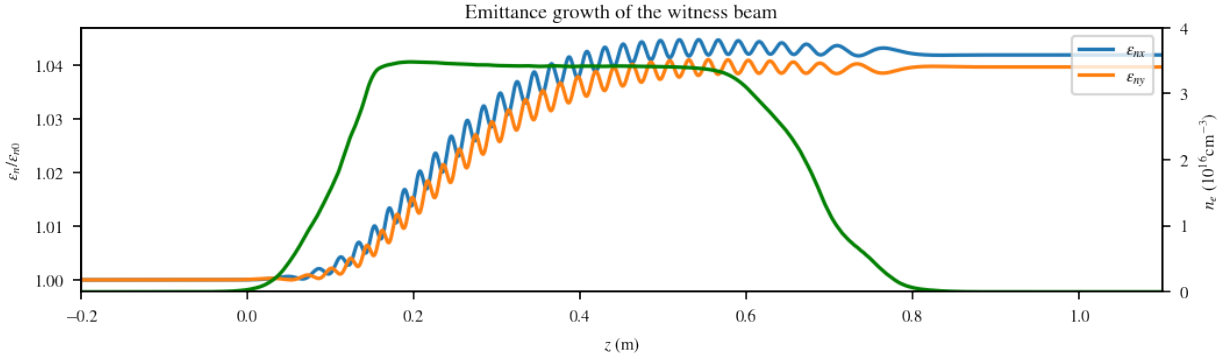


Figure 4.18: Growth of the witness beam’s emittance in the E301 plasma source. The plasma density profile is taken from the output of a refraction simulation, distortion is obvious in the down ramp.

matched, and thus, the drive beam cannot be matched and will undergo betatron oscillations in the ramp. At the minimum of the oscillations, the beam will pinch down below the matched size, increasing the strength of the electric field and resulting in even more ionization. The end result is a significant, potentially nonlinear, distortion of the partially ionized ramp that will prevent emittance preservation at the μm level. An example of the drive beam ionization in the E301 plasma source is shown in Fig. 4.19. For initial experiments, the emittance of the drive beam can be spoiled by introducing foils into the beam line. The resulting beam fields are reduced below the ionization threshold. The spoiled beams will at least allow proof of principle emittance preservation experiments, but new plasma sources will have to be developed to preserve the beam emittance at the μm level.

4.7 Shaping the Gas Density

Beam ionization distorts the partially ionized ramps in a laser ionized plasma source, preventing emittance preservation at the μm level. The obvious solution is to form the ramps using a shaped gas density and then fully ionize the plasma source along its entire length. LWFA experiments routinely use this approach either with a capillary discharge plasma, a gas cell, or a small gas jet. From Sec. 3.6, we know that we need a plasma that is about 50 cm long. It is most likely

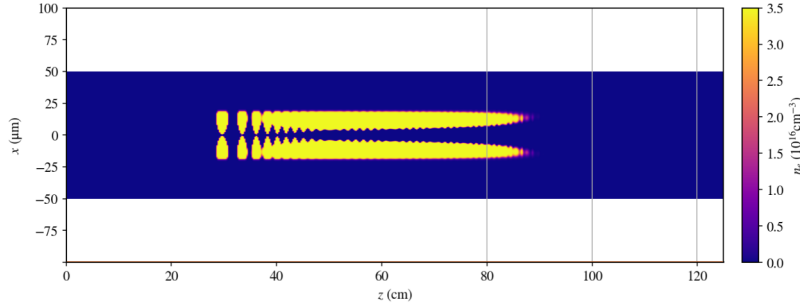


Figure 4.19: Beam ionization fraction in E301. The beam fields will ionize the remaining gas in the ramp every time the beam pinches down. By the time the beam reaches the bulk plasma it is continuously ionizing a annulus around itself.

possible to build a capillary discharge plasma this long, the longest capillary discharge built for LWFA is 40 cm long [114], but the uniformity of the plasma along the length of the discharge is unknown. Ramps are formed by the expansion of gas from the gas inlet to the end out the capillary and out into the surrounding vacuum. Even with a long enough capillary, the ramp lengths are short and difficult to control and there is no guarantee that the gas is fully ionized in the ramps [64, 115]. The next option is a gas cell, a chamber filled with gas with small apertures at the entrance and exit.¹³ Again, the expansion of the gas as it flows out of the apertures creates density ramps. The main difference from a capillary discharge is that the gas cell is either beam ionized, or laser ionized. To produce ramps of reasonable length, the inlet and outlet must have diameters on the order of millimeters. Such small apertures do not allow ionization of long plasmas using an axicon or tandem lens, which requires at least a 2 cm of clear aperture.

Both round and slit gas jets have been heavily used in LWFA. The shape of the nozzle can be modified to control the gas density profile and there are no obstructions to laser ionization. However, gas jets typically have sizes on the order of millimeters up to a couple centimeters, using them at FACET-II would require more than an order of magnitude increase in length. To determine if they would be a feasible option, we ran simulations of elongated gas jets to see what the typical longitudinal and transverse profiles would look like for FACET-II densities. The simulations are

¹³ Technically a capillary is a kind of gas cell where the outlet apertures are about the same size as the cell cross section.

discussed in Sec. 4.7.1.

The main plasma source used in PWFA experiments to date is an alkali vapor heat pipe oven. These plasma source can produce very long alkali vapor columns with controlled ramps at the entrance and exit. The alkali vapor is contained in the oven with a cooled buffer gas of helium, for more details on their operation see Sec. 5.2.3 and Ref. [116]. Originally designed for beam ionization, the vapor has a very low first ionization threshold. Even though the beam can ionize the plasma, laser pre-ionizing helps minimize head erosion. Unfortunately, heat pipe ovens still suffer from ramp degradation due to beam ionization of the buffer gas on either side. Nevertheless, the first experiments at FACET-II will utilize a lithium heat pipe oven because the initial beam emittance will be low enough that it will not ionize helium. The lithium oven is a good test bed for developing optics to ionize an elongated gas jet—from the laser’s perspective the setup is the same. In Sec. 4.7.2, we present a tandem lens design for ionizing the lithium oven at FACET-II.

4.7.1 Elongated Gas Jets

One way to avoid the distortion of the plasma ramps due to beam ionization is to shape the gas density into a ramp and fully ionize all the gas. We propose using a very long gas jet to produce a suitable gas density. The density distribution of the gas is controlled by the geometry of the nozzle and the backing pressure behind the nozzle. A set of tandem lenses can then fully ionize the entire length of the plasma source.

The question we must answer is: how does the nozzle geometry impact the gas density profile? We will use a converging-diverging nozzle to expand the gas from a reservoir into vacuum, Fig. 4.20(a) shows the cross section of a typical nozzle. A converging-diverging nozzle directed into vacuum will produce supersonic jets of gas that expand away from the nozzle in a non-isotropic distribution. Within the nozzle, the gas properties are described analytically by 1D isentropic theory [117, 118]. The fluid properties can all be found from the Mach number M , which is related

to the area of the nozzle by

$$\frac{A}{A^*} = \frac{1}{M} \left(\frac{2 + (k-1)M^2}{1+k} \right)^{(k+1)/[2(k-1)]}, \quad (4.133)$$

where A^* is the area of the nozzle throat, A is the nozzle area as a function of x , and $k = c_p/c_v$ is the ratio of specific heat at constant pressure c_p to specific heat at constant volume c_v . The rest of the fluid properties follow from M :

$$\frac{p}{p_0} = \left(\frac{2}{2 + (k-1)M^2} \right)^{k/(k-1)} \quad (4.134)$$

$$\frac{T}{T_0} = \frac{2}{2 + (k-1)M^2} \quad (4.135)$$

$$\frac{\rho}{\rho_0} = \left(\frac{2}{2 + (k-1)M^2} \right)^{1/(k-1)}, \quad (4.136)$$

where p is the pressure, T is the temperature, and ρ is the density. The subscript zero denotes the fluid properties in the reservoir backing the nozzle. The shape of the converging section of the nozzle is unimportant as long as it smoothly tapers to the throat. The gas jet can only produce suitable ramps if the background gas density in the chamber is very low, otherwise the background gas will be ionized and impact the beam evolution. This means the gas must be pumped out between shots and the gas flow into the chamber should be minimized. The nozzles then should have a narrow throat width, $w_t < 1$ mm. Since w_t is much less than the length of the nozzle L , the nozzle behaves as a 2D structure and A in Eq. (4.133) can be replaced by the nozzle width.

1D isentropic theory assumes there is no transverse variation in the gas properties. The fluid velocity, however, always drops to zero at the edge of the nozzle, and there is a transition zone from the uniform flow at the center of the nozzle to the stationary flow at the edge, called the boundary layer. The impact of the boundary layer is quantified by the dimensionless Knudsen number Kn [118]

$$\text{Kn} = \frac{\lambda_{MF}}{L} = \frac{1}{\sqrt{2}\sigma n_g L}, \quad (4.137)$$

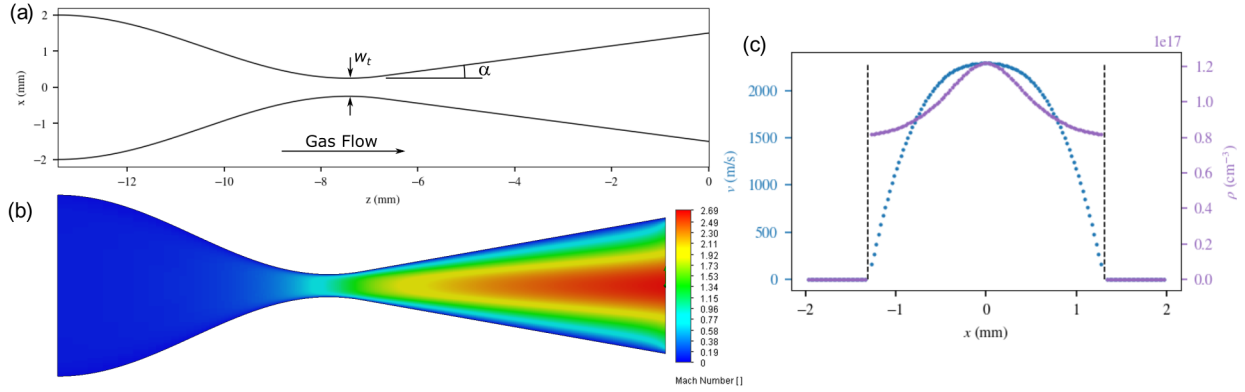


Figure 4.20: A converging-diverging, low pressure 2D nozzle. (a) schematic of the nozzle geometry. The diverging part of the nozzle is linear with half angle α . The geometry of the converging part of the nozzle has little effect on the flow. (b) Mach number within the nozzle. At the nozzle throat, $M = 1$ and increases in the diverging part of the nozzle. At the gas densities relevant for PWFA, the boundary layers occupy a significant portion of the nozzle. (c) velocity and density profile 1 mm before the mouth of the nozzle showing that the boundary layers extend almost to the center of the nozzle.

where λ_{MF} is the mean free path of the gas molecules, L is the characteristic size of the geometry, and σ is the scattering cross section of the molecules. As long $\text{Kn} < 10^{-1}$, continuum fluid mechanics is a valid approximation, but the boundary layer can become significant. Typical gas jets, such as rocket nozzles or high density LWFA jets, operate in a regime where the boundary layer effect is minimal. We want gas densities of around $3.5 \times 10^{16} \text{ cm}^{-3}$ about 1 cm above the gas jet. The resulting density in the jet is low enough that the boundary layer is significant as shown in Fig. 4.20(b) and (c). The boundary layer extends nearly to the center of the jet at the nozzle outlet. We are forced to turn to simulation to determine the flow behavior of the nozzle.

We considered 2D conical nozzles, which have straight sides in the diverging section. The geometry of the nozzle is shown in Fig. 4.20(a). The throat width w_t , diverging angle α , and exit width w_e fully define the nozzle. From Eq. (4.133) we can see that the important parameter is the area ratio w_e/w_t , which we will typically use instead of w_e . We ran 2D computational fluid dynamics (CFD) simulations (Solidworks Flow Simulation) of different conical nozzle geometries to see how the nozzle geometry impacts the density distribution of the plume. The results are shown

in Fig. 4.21. Two different throat sizes were simulated, the larger throat resulted in a plume with lower density at the exit of the nozzle, but the density dropped off more slowly in the axial direction. The transverse profile of the plume 10 mm away from the nozzle was similar for both throat sizes. Larger area ratio has a similar effect as throat width, reducing the density at the nozzle exit with a slower drop off away from the nozzle. The transverse profile is slightly narrower for large area ratio. The diverging angle had a small impact on the longitudinal density, with small diverging angle leading to lower density. The diverging angle has the largest impact on the transverse profile, with larger diverging angles leading to narrower plumes. We also simulated a bell shaped nozzle but the profile was very similar to the 2D conical nozzles due to the large boundary layer [119].

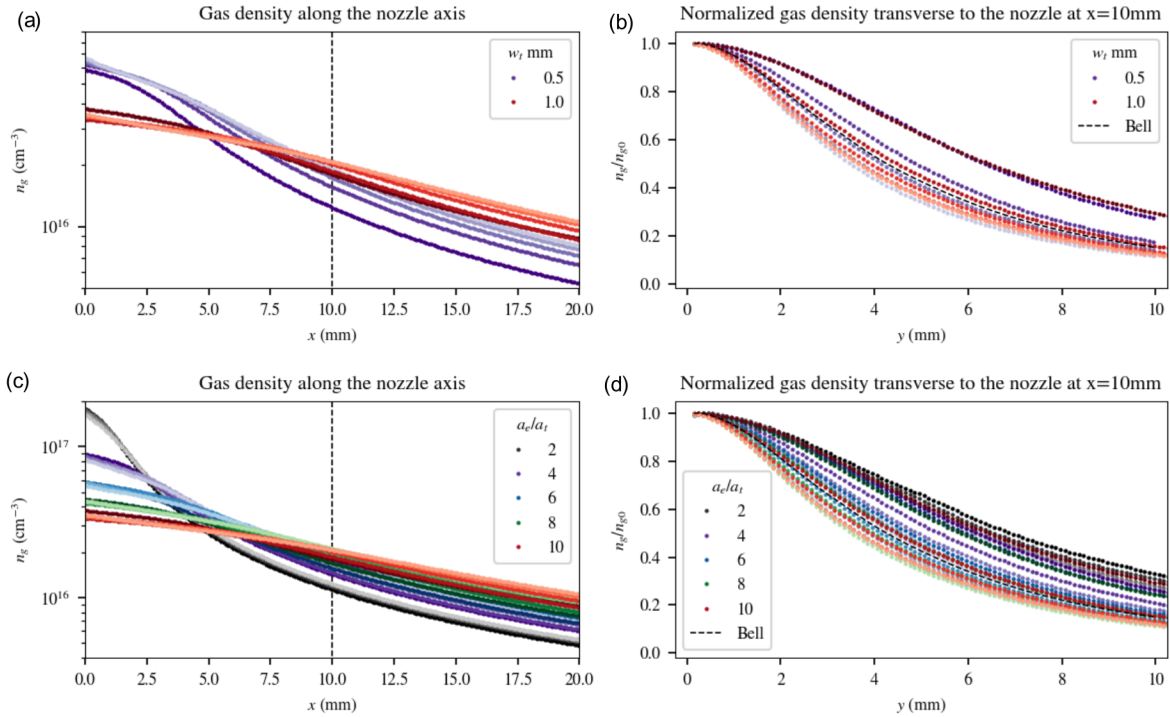


Figure 4.21: Variation in the gas density profile as a function of 2D nozzle geometry. The top plots show the axial gas density (a) and the transverse gas density 10 mm above the nozzle (b) at $AR = 10$ for different throat sizes. The bottom shows the axial (c) and transverse (d) density for a $w_t = 1$ mm nozzle with different expansion ratios (color). In all figures the shading indicates α from 5° to 15° (dark to light).

Remember, the primary goal of our plasma sources is to preserve the beam emittance. The jet needs to produce appropriate ramped gas density profiles on the entrance and exit to achieve

this goal. We ran a series of simulations of a full 40 cm gas jet with different end geometries to see what ramp shapes we could achieve. The end geometries all consisted of a flat plane truncating the end of the nozzle. The angle between the plane and the axial direction was varied between 0° to 15° . The ramp lengths at larger angles are around 5 cm long, about the length we require for matching the witness beam to the bulk plasma. After initial proof of concept experiments with a flooded chamber, the elongated gas jet will be considered as a replacement plasma source at FACET-II.

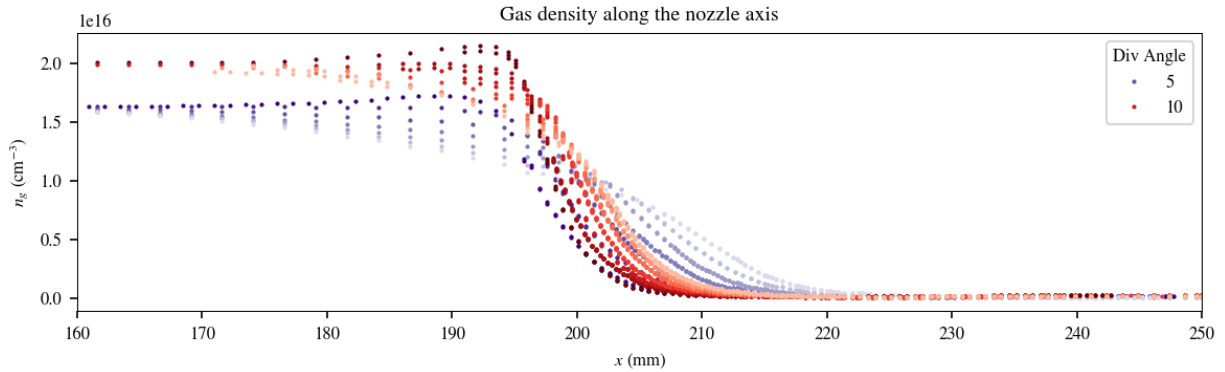


Figure 4.22: Gas density ramp at the end of several elongated gas jets with different end geometries. The purple points use a nozzle with $\alpha = 5^\circ$, the red use $\alpha = 10^\circ$. Both nozzles have $T = 1$ mm and $AR = 10$. The end of the nozzle is a flat plane inclined relative to the nozzle flow direction. Dark to light is inclined angles from 0° to 15° .

4.7.2 Laser Ionization of a Lithium Vapor Oven

The lithium heat pipe oven differs from the E301 plasma source in that the gas species already has the correct density profile and simply needs to be ionized along the entire length of the oven. The lithium oven is good analog for the ionization optics that would be used on an elongated gas jet. We follow the same approach as Sec. 4.5.3, but we have more freedom in choosing the target intensity profile because we no longer need to form ramps. The lithium gas density in the oven rises from zero and drops back down to zero over a length of 80 cm. With experience, we have found that the plasma source width is maximized if the laser intensity grows as it propagates. The target laser intensity is set to be 6 times that required for full ionization at the start of the lithium

oven and grows by 40% over 80 cm. Somewhat arbitrary intensity ramps are added to either side to maintain a smooth intensity profile. The target is shown in Fig. 4.23(d). For lithium, we ran a set of refraction simulations with different R_b to find the value that produced the widest plasma while still fully ionizing across the entire width. The optimal value was $R_b = 60 \mu\text{m}$.

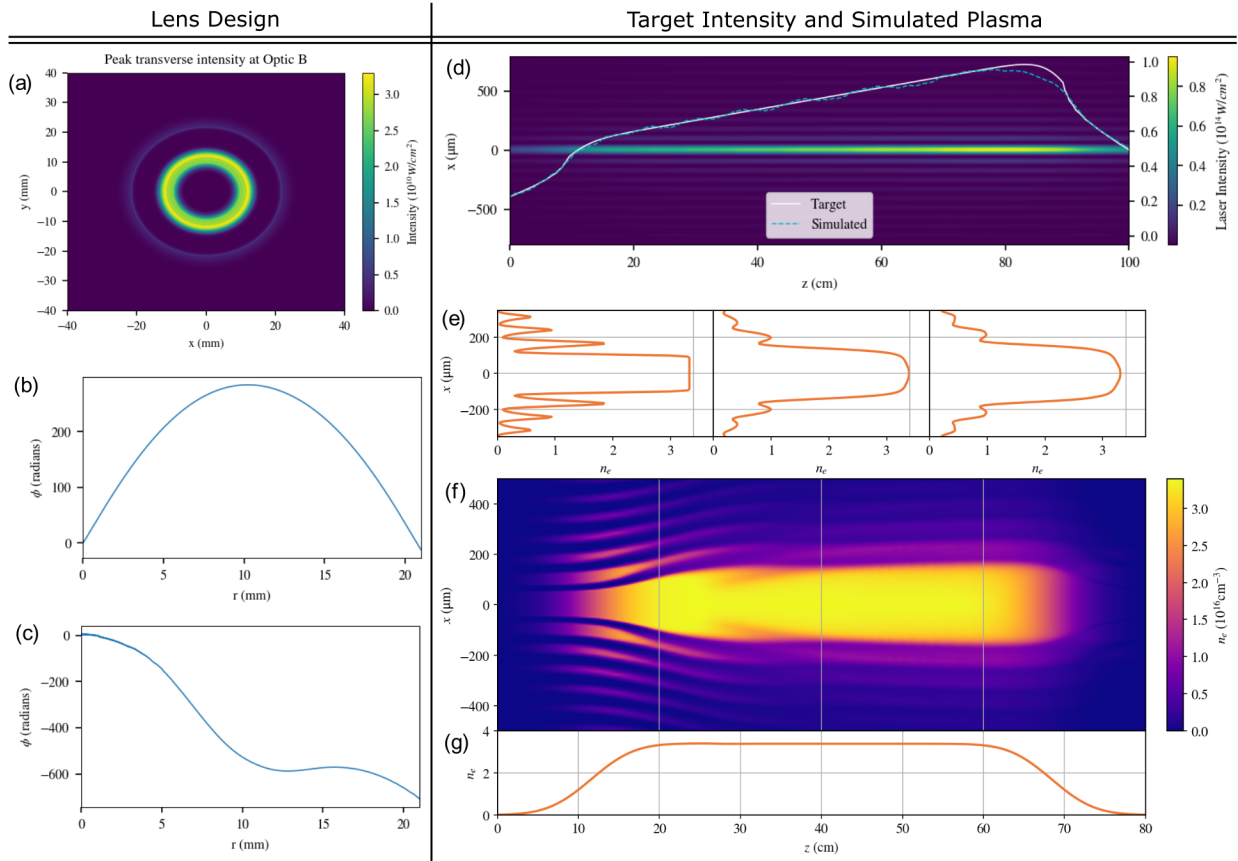


Figure 4.23: Design of the LO8 optic for E300. (a) peak laser intensity at lens B as a result of the phase imparted by lens A. (b) phase imparted by lens A. (c) phase imparted by lens B. (d) simulated laser intensity in vacuum, the lineouts show the on-axis intensity from simulation (dashed) and the target (solid). (e) transverse plasma density at the start, middle, and end of the bulk plasma. (f) simulated plasma density including refraction. (g) longitudinal plasma density on-axis. For the lithium oven, the longitudinal density profile is determined by the lithium density along the oven.

The tandem lens pair sits in the accelerator with the second lens 1.75 m upstream of the start of the 80 cm long lithium oven. The first tandem lens is 1 m upstream of the second. The intensity on the second lens and the phases of the two lenses are shown in Fig. 4.23(a), (b), and (c), respectively. The simulated plasma profile, with refraction, is shown in Fig. 4.23(f). The

widening of the plasma through refraction is clearly visible. The plasma width at the beginning of the entrance ramp is the width of the central Bessel peak in vacuum. As the beam propagates, the plasma widens significantly. In the lithium oven, the plasma width does not need to fully contain the wake because the ionization threshold is low enough that the drive beam will ionize a much wider region than the laser. The pre-ionized plasma is there to minimize head erosion in the drive beam. The longitudinal line-out in Fig. 4.23(g) shows the lithium density profile, which is the plasma density profile in the lithium oven.

Chapter 5

Experimental Setup

The experimental campaign to demonstrate emittance preservation can be broken down into two pieces. The first is demonstrating that the laser ionization optics perform as expected and can produce the desired plasma profile. The second is firing a pair of electron bunches into the plasma and showing that the emittance of the witness beam is preserved during acceleration, preferably with significant energy gain and minimal energy spread. The electron beams are provided by the FACET-II accelerator at SLAC. The accelerator is equipped with a suitable laser for producing laser ionized plasma sources, but the experimental environment at FACET-II is challenging. The experimental area is located 10m underground and 100m from the nearest staircase and space in the tunnel is limited. Because the experiment is located in the accelerator housing, physical access to the experimental area is limited to when the accelerator is shutdown. Making changes or adjusting setups is time consuming and difficult. The best approach is to test and validate the plasma formation away from FACET-II and then install it for the final experiment. For this reason, characterizing the optics for laser ionization is carried out in our lab at University of Colorado Boulder (CU). The laser system at CU is intentionally designed to have similar specifications to the experimental laser at FACET-II.

In this chapter, we describe the experimental setups at both CU and FACET-II. We start with the 15 TW laser system, vacuum system, and control system built at CU for these experiments. Next, we describe the experimental setup used to characterize the tandem lenses; the results of the tests are presented in Ch. 5. Then we move to FACET-II with a brief description of the accelerator

itself, followed by a more detailed description of the experimental area and laser system.

5.1 University of Colorado Laser Lab

The laser lab at the University of Colorado Boulder is designed for plasma source and diagnostic development for FACET-II. The lab revolves around a 15 TW peak power ultrafast Ti:sapph laser, model Centaurus from Continuum Inc. (now Amplitude Laser Group). The laser produces 30 fs, 450 mJ pulses at 10 Hz repetition rate. The laser is coupled to an experimental vacuum system with a deformable mirror (DM) and a large volume that can be filled with gas for producing laser ionized plasmas in. The experimental equipment is controlled by a home-made central data acquisition system and control computer.

5.1.1 The Laser System

The laser is a commercial system that fits on a single 4 ft by 12 ft optical table with a free standing in-vacuum compressor at the foot of the table occupying a 4 ft x 4 ft area. The layout of the system is shown in Fig. 5.1. The laser chain consists of an oscillator (Laser Quantum Vanteon Power) producing 10 fs pulses at 80 MHz repetition rate. The pulses enter the stretcher enclosure where an Öffner triplet stretches the pulses to 300 ps. A fast photodiode just upstream of the stretcher generates a 80 MHz electrical signal which is divided by the laser's timing system to produce the 10 Hz triggers for the rest of the system. After the stretcher is a Faraday isolator followed by a pulse picker. The pulse picker uses a Pockels cell and a polarizing beam splitter to select every 8th million pulse, reducing the pulse train to 10 Hz repetition rate. The rest of the pulses are sent back through the Faraday isolator, where they are dumped.

After stretching and downsampling to 10 Hz, the laser pulses are ready for amplification. The next enclosure contains both the regenerative amplifier (regen) and the first multi-pass amplifier (MP1). Both are pumped by a single Continuum Inlite II producing 150 mJ, 6 ns pulses of 532 nm light at 10 Hz. The regen amplifies the 800 nm seed pulses up to 0.7 mJ using a Brewster cut Ti:sapph crystal. The pulses are further amplified in the 5-pass MP1 to 20.5 mJ before leaving the

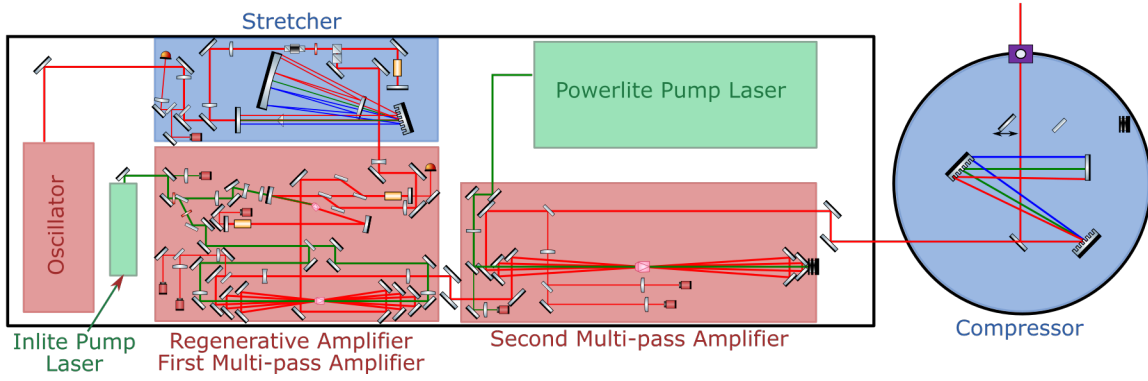


Figure 5.1: Schematic of the 15 TW laser system at the University of Colorado. Ultrashort laser pulses are generated in the oscillator before being stretched and amplified by a chain of three amplifiers: regenerative amplifier to 0.7 mJ, first multi-pass amplifier to 20.5 mJ, and second multi-pass amplifier to 550 mJ. The amplifiers are pumped by two of 532 nm lasers.

enclosure and entering the enclosure for the second multi-pass amplifier (MP2). MP2 is pumped by a Continuum Powerlite DLS which produces 2 J, 6 ns laser pulses of 532 nm light at 10 Hz. The pump pulses pass through a homogenizer before reaching the Ti:sapph crystal in MP2. The homogenizer helps create a uniform transverse intensity in the 800 nm pulse. MP2 is a water cooled, 4-pass amplifier that increases the laser power to 550 mJ. A beam sampler picks off 5% of the laser pulse, most of which is sent into a pair of probe lines. After each multi-pass amplifier is a camera imaging the laser beam at the Ti:sapph crystal, called a near field imaging system, and a camera imaging the focal point of a spherical lens, called a far field imaging system. The near field camera monitors the laser position, the far field camera monitors the pointing angle of the laser, together they act as an alignment reference.

The pulse passes through a collimation lens and then a periscope to raise it to the height of the compressor chamber. Inside the chamber, a pair of gratings compress the pulse down to 30 fs with 80% energy efficiency. The grating spacing is remotely adjustable allowing the compression to be adjusted to compensate for dispersion in downstream optics. The final mirror in the compressor is also motorized to allow alignment through the exit window of the compressor. The window has an aperture of 30.2 mm, which is slightly smaller than the beam diameter. Therefore, alignment adjustability after pumping the chamber down to vacuum is critical to minimize clipping on the

aperture. After passing through the exit window, the laser enters the first experimental chamber where it reflects off a deformable mirror (Imagine Optics Ilao Star). Leakage light from the mirror directly after the deformable mirror passes out of the vacuum chamber through a window and is imaged by a 4f, $M = 0.1$ telescope onto a Shack-Hartmann wavefront sensor. The wavefront sensor acts in closed loop with the DM to produce a target phase profile. The DM can typically reduce the wavefront error to $< 0.2 \mu\text{m}$ P-V and $< 0.03 \mu\text{m}$ RMS. We have seen P-V errors less than $\lambda/10$. A summary of typical laser parameters after the DM is shown in Fig. 5.2.

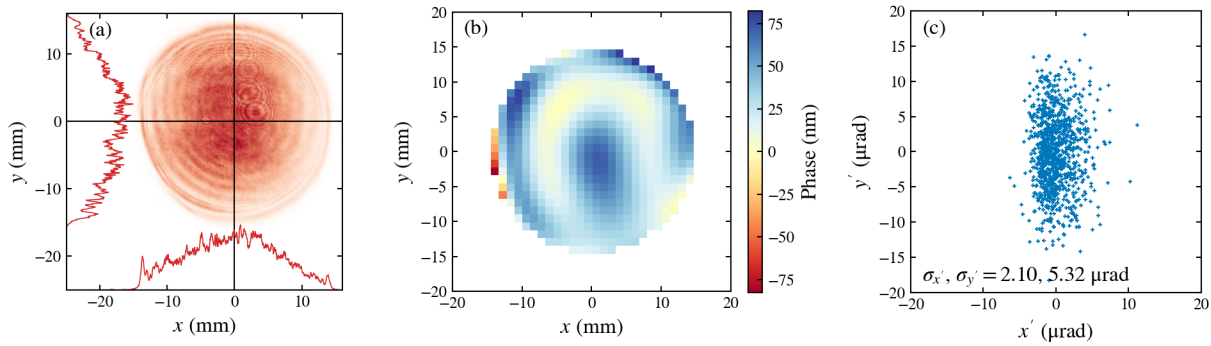


Figure 5.2: Pulse parameters of the laser system. (a) the transverse profile is slightly clipped by the window aperture on the compressor. MP2 was not pumped for this image, the transverse profile is more uniform when pumped. (b) measured wavefront of the laser after correction with the DM, focus and tilts have been removed. The RMS wavefront is 21 nm. (c) laser pointing jitter over 1000 shots taken at the focal plane of a 1 m focal length lens.

5.1.2 The Vacuum System

The vacuum system starts with the vacuum compressor for the laser system. The compressor is pumped by an Agilent TwissTorr 304fs turbomolecular pump and typically operates at a couple 10^{-7} mTorr to protect the surface coatings of the gratings. The compressor has its own roughing pump which also serves as a backing pump for the turbo. All the valves on the system are pneumatically operated which allows the pump down of the chamber to be automated by a controller supplied with the laser system. The compressor chamber is free standing on its own supports and is not coupled to the optical tables. Inside the chamber is an optical breadboard that is decoupled

from the chamber and mounted to the floor through a separate support structure. The compressor vacuum is isolated from the rest of the system by a window installed in the bellows between the compressor and the next chamber, called chamber A. The main experimental system is composed of two chambers, chamber A and chamber B, separated by 39.25 in. In between the two chambers is a six-way vacuum cross with gate valves between it and either chamber. Spools connect the gate valves to the cross. The layout of the system is summarized in Fig. 5.3.

The system is pumped by an Agilent TwissTorr 304fs turbomolecular pump mounted on top of the cross, a gate valve is placed between the turbo and the cross so that it can be isolated from the vacuum system. The system has a separate roughing pump that is connected via right angle valves to both the chambers and the six-way cross. The gate valves separate the system into three vacuum volumes, chamber A, six-way cross, and chamber B, each of which can be roughed individually. The six-way cross can be pumped by the turbo pump directly while chambers A and B can be pumped to high vacuum by opening the gate valve between the chamber and the six-way cross. The chambers can be flooded with either nitrogen, argon, or helium from a set of manifolds plumbed into each of the three vacuum volumes. The layout of the system is summarized in Fig. 5.3.

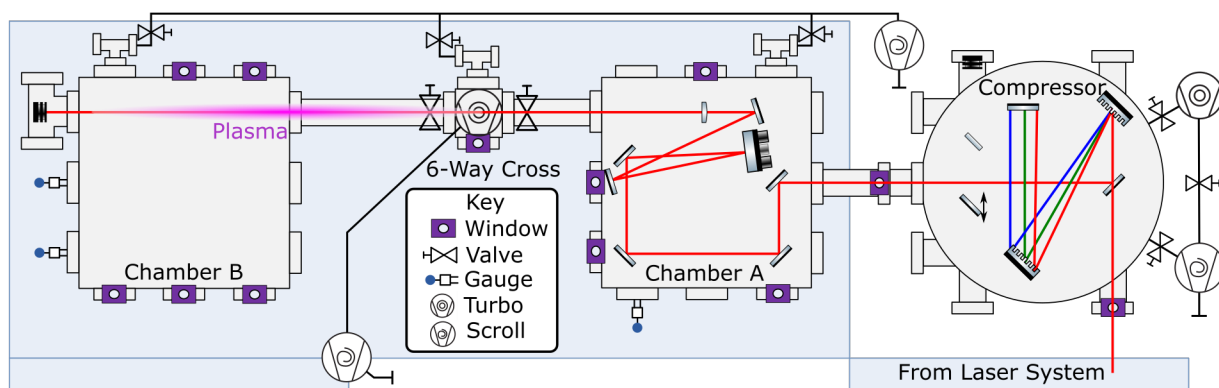


Figure 5.3: Diagram of the vacuum system. The laser pulse is compressed in the compressor chamber which is vacuum isolated from the rest of the system by a thin window. Laser preparation occurs in chamber A and the plasma is formed in chamber B.

Chamber A is primarily used for preparing the laser. Inside the chamber is the deformable

mirror for correcting the laser wavefront. In plasma generation experiments, the focusing optics are typically located in chamber A. The chamber is equipped with a selection of windows and electronics feedthroughs. All the mirrors in the chamber are motorized to allow remote alignment while the system is pumped down. Being able to align the system under vacuum is important because the system tends to shift during pump-down. Leakage light from several mirrors exits the chamber through the windows and can be used for diagnostics. Plasma formation typically occurs either in the six-way cross, chamber 2 or both. The cross and chamber B are equipped with windows to accommodate optical diagnostics of the plasma.

5.1.3 DAQ

The data acquisition system (DAQ) handles recording data from instruments, controlling the experimental hardware, and displaying data in live time. The system is a custom python program and GUI that operates on a single computer and connects to instruments in the lab either over a dedicated local area network or via USB. The program is heavily multi-processed to prevent one instrument from holding up data acquisition from other instruments. Each device interfacing with the DAQ, whether it be a turbo pump controller or a camera, has two dedicated processes for data acquisition and control. The first process handles sending and receiving commands from the device, and the second handles saving data to disk. With this architecture, a device that takes a long time to reply to a request will not prevent commands from being sent to any other device. Similarly, saving the data to disk will not interrupt data gathering. In fact, the computer has enough memory, 256 Gb, to temporarily store a significant number of shots in memory before saving them to disk. On top of the DAQ processes is a GUI which provides a convenient way to control the system as well as displaying live data from the instruments. Simple post-processing applications are built into the GUI, such as reference crosses on images, to aid in running experiments.

Synchronization of the experimental systems is handled by a signal delay generator (Stanford Research Systems DG645). The master timing signal comes from the laser system. The signal delay generator (SDG) can then generate triggers for measurement devices such as cameras and

spectrometers. The SDG can be enabled and disabled by a home-built timing controller that interfaces with the DAQ. One output of the SDG is fed back into the timing controller to keep track of the shots in the current dataset. This allows shot numbers to be assigned to the data from multiple instruments and be confident the shot numbers are synchronized across instruments. The system works as follows. When a new dataset is started, all instruments are stopped, and the SDG is disabled turning off the trigger signals. The instruments are reset and told to wait for a trigger and that they will be taking N shots. Once all instruments are ready, the SDG is enabled, and the timing controller starts counting the trigger signals sent out by the SDG. Once the timing controller has counted to N , it disables the SDG ensuring that only N triggers have been sent out. The number of shots recorded by each instrument is then counted. As long as each instrument recorded N shots, we can be confident that all the shots are synchronized. If any shots are missing, the dataset is retaken. Typical datasets range from 10 to 100 shots, which take 1-10 seconds when the laser is firing at its nominal rate of 10 Hz.

5.1.4 Setup for Testing Lenses

Optical characterization of the tandem lens was performed outside of vacuum for convenience. Fig. 5.4 is a schematic diagram of the experiment. The setup uses the leakage light normally used for the wavefront sensor. The leakage from the first mirror after the DM exits the vacuum through a window on the chamber. The light passes through a periscope, bringing it to table height, and is reflected off a beamsplitter then a mirror to reach the first tandem lens, a second mirror directs the beam through the second tandem lens and down the optical table. The first tandem lens is mounted on a kinematic mount with four degrees of freedom: translation in both transverse directions and rotation in pitch and yaw. The second tandem lens mount is on a slightly different kinematic mount that allows translation in the horizontal and longitudinal directions as well as pitch and yaw rotation. The pitch degree of freedom is provided by a micrometer rotation stage to ensure known angular adjustments. This is used for finding the angular alignment tolerance of the optic. A custom built 2.2 m travel translation stage is placed after the tandem lens to measure the

evolution of the pulse.

The light transmitted through the first beamsplitter is used to characterize the laser before it reaches the optics. A 4f imaging telescope with $M = 0.1$ is used to image the surface of the DM onto a near field camera and a Shack-Hartmann wavefront sensor. A second diagnostic arm focuses the laser with a 1 m focal length lens. The focal spot is imaged onto a far field camera using a $10\times$ microscope objective. In addition, a spectrometer can be inserted into the laser before the first tandem lens to measure the spectrum of the pulse. The near and far field allow measurements of the beam profile and pointing stability. The far field was used to take the pointing jitter measurements in Fig. 5.2(c).

5.2 FACET-II Accelerator

The Facility for Advanced Accelerator Experimental Tests (FACET) operated from 2011 to 2016 producing the highest energy electrons and positrons in the world. The facility was designed to provide the high peak current, high energy electron beams needed to develop plasma wakefield accelerators. The accelerator occupied the first 2 km of the 3 km SLAC linac, delivering 20 GeV electron and positron beams to a group of experiments tucked away in the middle of the linac tunnel [7]. In 2016 LCLS-II began construction in the first kilometer of the accelerator, requiring the removal of the first half of the FACET accelerator. This provided an opportunity to install a new electron source and upgrade FACET into FACET-II with half the energy, but significantly lower emittance and increased peak current. The peak current is so high, it opens up research opportunities in high energy physics as well as advanced accelerator research. The experimental area of the accelerator has been redesigned to handle a selection of diverse experiments in both fields [120].

In this section, we give a brief overview of the FACET-II accelerator from the electron source to the experimental area. We discuss the expected beam parameters for PWFA experiments and go over the parts of the experimental system relevant to emittance preservation, in particular the laser ionized plasma source, lithium oven, and diagnostics.

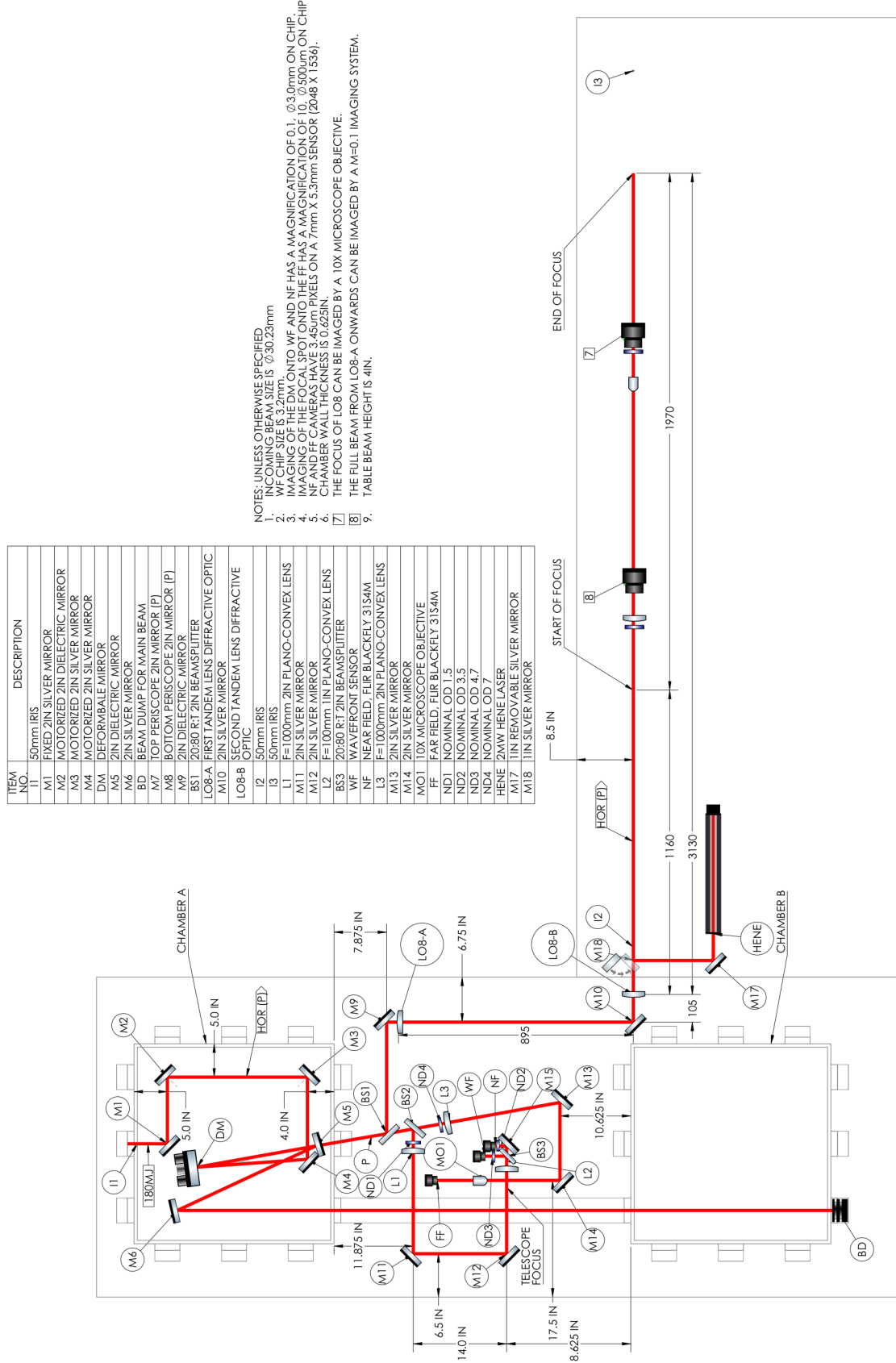


Figure 5.4: Experimental setup for characterizing tandem lenses. After reflecting off the deformable mirror, the leakage from M5 leaves the vacuum system and is split into two arms: the tandem lens arm and a diagnostic arm. The tandem lens arm has a 2.2 m translation stage for measuring the laser propagation. The diagnostic arm contains a near and far field camera and a Shack-Hartmann wavefront sensor.

5.2.1 Brief Linac Overview

The FACET-II accelerator starts in sector 10, 1 km from the start of the original SLAC linac, where the electron gun is located in a bunker off the side of the main accelerator housing. The accelerator housing, also called the tunnel, is located 10 m underground and contains the magnets and RF cavities of the accelerator. At ground level above the accelerator is the Klystron gallery, stretching the full length of the accelerator. The gallery contains the RF amplifiers (Klystrons) used to feed the accelerating cavities, it also contains support equipment such as electronics and vacuum pumps for the linac. After the gun, the beam is accelerated to 135 MeV before entering the main linac where it is accelerated and then longitudinally compressed in three stages. Finally, the beam reaches the final focus and experimental area. An overall schematic of the FACET-II linac is shown in Fig. 5.5. The accelerator can produce beams with a range of parameters; in the following paragraphs, we consider a typical beam for PWFA experiments.

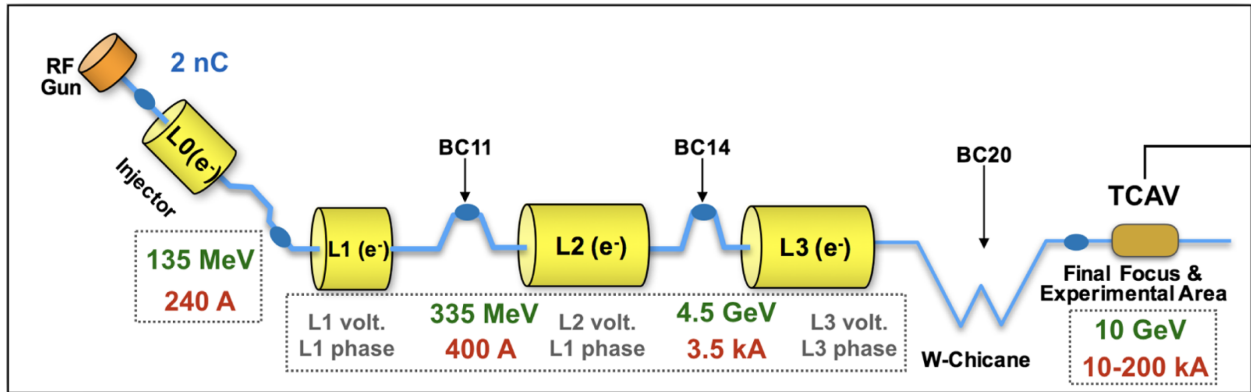


Figure 5.5: Overall schematic of the FACET-II accelerator. The beam is generated in a photo-injector (RF gun) before being sent through a series of RF accelerating structures (yellow) and bunch compressors before reaching the final focus and experimental area. Reproduced from Ref. [121].

A photocathode source provides the electrons for FACET-II. Within the source, a 400 μJ UV laser pulse at 253 nm is sent onto a copper cathode within a 1.6 cell RF cavity. The laser pulse releases 2 nC worth of electrons from the copper. The phase of the RF cavity is synchronized to the laser so that the electrons are injected near the peak of the RF field. As a result, the particles

quickly accelerate to 4.2 MeV before the emittance has a chance to grow significantly due to space charge effects. The bunch immediately passes through a solenoid magnet integrated onto the end of the RF cavity that focuses the bunch. Fig. 5.6 is a schematic of the RF gun showing the RF cavity, solenoid magnet, and the path the laser takes to the cathode. The beam then enters a pair of 3 m S-band accelerating structures to increase the energy to 135 MeV. These first two accelerating structures are called L0 and, combined with the gun, form the part of the accelerator called the injector.

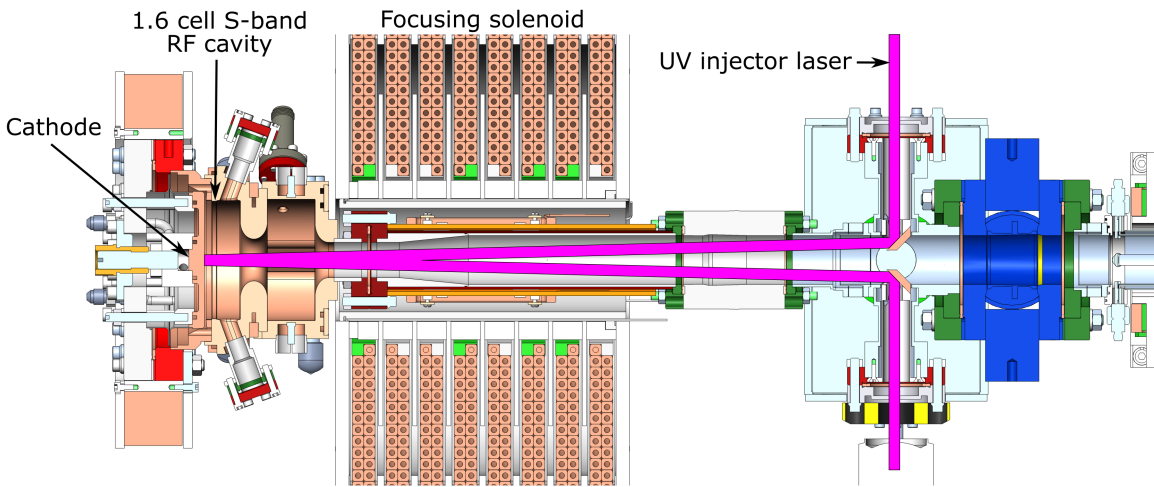


Figure 5.6: FACET-II photo-injector electron source. A UV laser pulse enters the vacuum system from the top of the image and is reflected onto the copper cathode. The UV photons free electrons from the copper and into the fields of a 1.6 cell S-band RF cavity. The electrons are accelerated in the cavity to 4.2 MeV before entering a focusing solenoid.

The beam is injected onto the main accelerator axis by a pair of dipole magnets. The main accelerator accelerates and compresses the beam in three stages, shown in Fig. 5.7. Each acceleration section is composed of many 3 m long S-band traveling wave structures, with every three powered by a Klystron RF amplifier. In the first acceleration stage (L1), the beam is accelerated to 335 MeV while travelling through the accelerating structures -20.5° off-crest of the RF wave, generating an energy chirp for compression. This is followed by the first round of compression in a 4-bend chicane in sector 11 (BC11), reducing the beam length to $468 \mu\text{m}$. The second acceleration stage (L2) follows. The beam is boosted to 4.5 GeV at -39.0° before being compressed by a second

4-bend chicane (BC14) to $96\ \mu\text{m}$. Typically, the beam is slightly over-compressed to provide the required chirp for the final bunch compressor. In the third and final stage of acceleration, the beam reaches the full 10 GeV energy, accelerated on the RF crest [120]. Before delivery to the experimental area, the beam is compressed to its final length of $1.8\ \mu\text{m}$ in the W-chicane shown in Fig. 5.8.

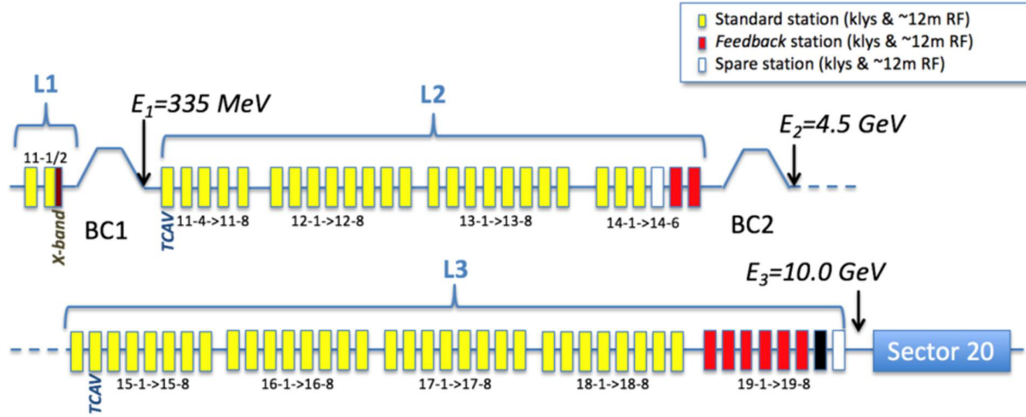


Figure 5.7: Overview of the FACET-II linac. The beam is accelerated in a series of S-band travelling-wave RF structures. Each box represents four 3 m long RF sections fed by a single Klystron. Reproduced from Ref. [120] .

Fig. 5.8 also shows a top level schematic of the experimental area, located after the final bunch compressor. The experimental area will be covered in more detail in Secs. 5.2.3 and 5.2.4, but, in brief, it contains the transverse deflecting cavity, final focus magnets, interaction point (IP) area with experimental hardware—targets and plasma source—and laser, imaging spectrometer for the electron beam, gamma ray/x-ray diagnostics and the beam dump. From the linac side, the experimental area is dedicated to shaping the beam for experiments and measuring the beam parameters before and after the experiment.

In our description so far, we have only described the process to generate and accelerate a single beam. Plasma wakefield acceleration requires two bunches, a drive and witness. Two ways are available to produce the two bunch structure at FACET-II. The first technique is illustrated in the top of Fig. 5.8. The beam is dispersed transversely according to energy within the W-chicane.¹

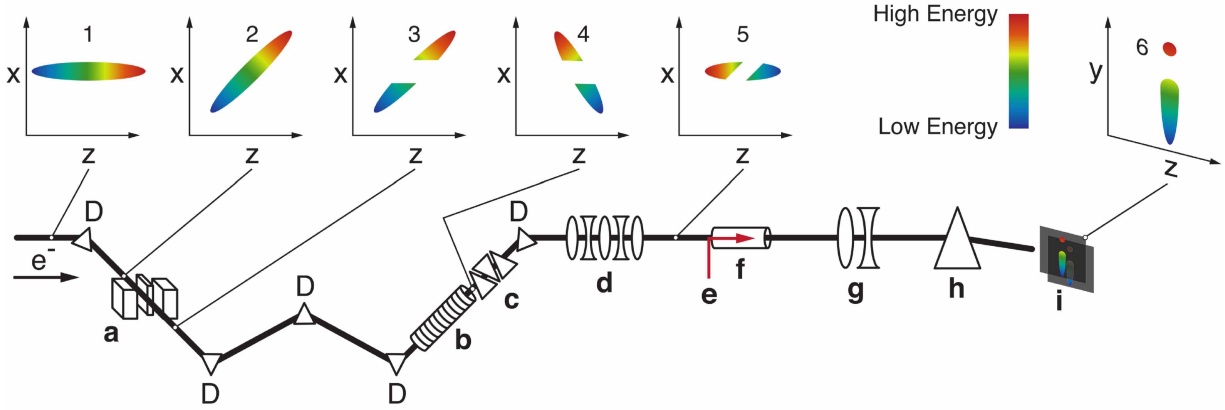


Figure 5.8: Final bunch compressor and IP area of FACET. Dipole magnets are labelled with (D). (a) is the notch collimator, (b) the transverse deflecting cavity, (c) a spectrometer, (d) final focus magnets, (e) plasma ionization laser, (f) plasma source, (g) image quadrupoles for the spectrometer, (h) spectrometer dipole magnet, and (i) screens for imaging the beam. For FACET-II the transverse deflecting cavity has been moved after the chicane and an additional quadrupole has been added to the imaging spectrometer. Reproduced from Litos *et al.* [8].

If the beam has some initial energy chirp, which it must for compression, then each energy slice is located at a different longitudinal position in the beam. In the chicane, the beam is displaced transversely according to energy and, therefore, displaced transversely according to longitudinal position. The beam can then be sliced in two transversely by placing a dense metal in the way of the center part of the beam. The transverse two bunch structure will be rotated into a two bunch longitudinal structure through the rest of the chicane. The device that blocks the center of the beam is called the notch collimator and it includes adjustable pieces to block the tails of the beam as well. The second way to produce two beams is to send two, closely spaced, laser pulses into the photocathode. The beams are then accelerated and compressed together through the linac. The final pulse can be cleaned up using the notch collimator.

¹ Compression of the electron beam works the same way as a laser beam. The beam is given an energy chirp along its length before being passed through a set of optics where the different energy components take different path lengths. The particles at the rear of the beam take a shorter path and catch up to the particles at the front. In a chicane the difference in path length comes from dispersing the beam. In the transport matrix formalism, a bunch compressor has a non-zero R_{56} .

5.2.2 Beam Parameters

The FACET-II team at SLAC has run numerous particle tracking simulations of the beam through the accelerator to develop machine configurations for various experiments. These simulations provide the full phase space distribution of the electron beam throughout the linac. The phase space of the nominal two bunch configuration at the beginning of the plasma source is shown in Fig. 5.9. The two bunch structure in these simulations was created by sending two laser pulses onto the cathode. The accelerator is highly tunable, and the bunch structure can be adjusted in a range around these parameters, although these simulations are near peak compression so the bunch length can only be made longer.

The drive and witness beams do not have exactly the same energy. Any chromatic aberrations in the focusing system will cause the two beams to be focused differently, especially when they are focused to very small spot sizes. The beam parameters shown in Fig. 5.9 are taken at the vacuum focus of the witness beam, which is called the IP and is located near the beginning of the density ramp of the plasma source. The drive beam size is much larger at this location because the drive beam waist in the horizontal x direction is 14.8 cm downstream and in the vertical y direction is 34.1 cm downstream. At the IP, the CS parameters of the witness beam in x and y , respectively, are $\alpha_{x,y} = 0.1, 0.7$ and $\beta_{x,y} = 5.0, 7.0$ cm. For the drive beam they are $\alpha_{x,y} = 1.7, 2.2$ and $\beta_{x,y} = 33, 33$ cm. Each beam can be modelled by the Gaussian distribution of Sec. 2.3.4 using the Gaussian fits in Fig. 5.9 combined with the CS parameters.

5.2.3 IP Area

The area around the electron beam focus is called the IP area. The IP area is where most of the experimental hardware, including the plasma source, is located along with several beam diagnostics. It is situated within the slightly larger region called the experimental area which includes everything in the accelerator downstream of the W-chicane. Fig. 5.10 is a schematic of the experimental area. In a PWFA experiment, the incoming beam is focused down to a small spot

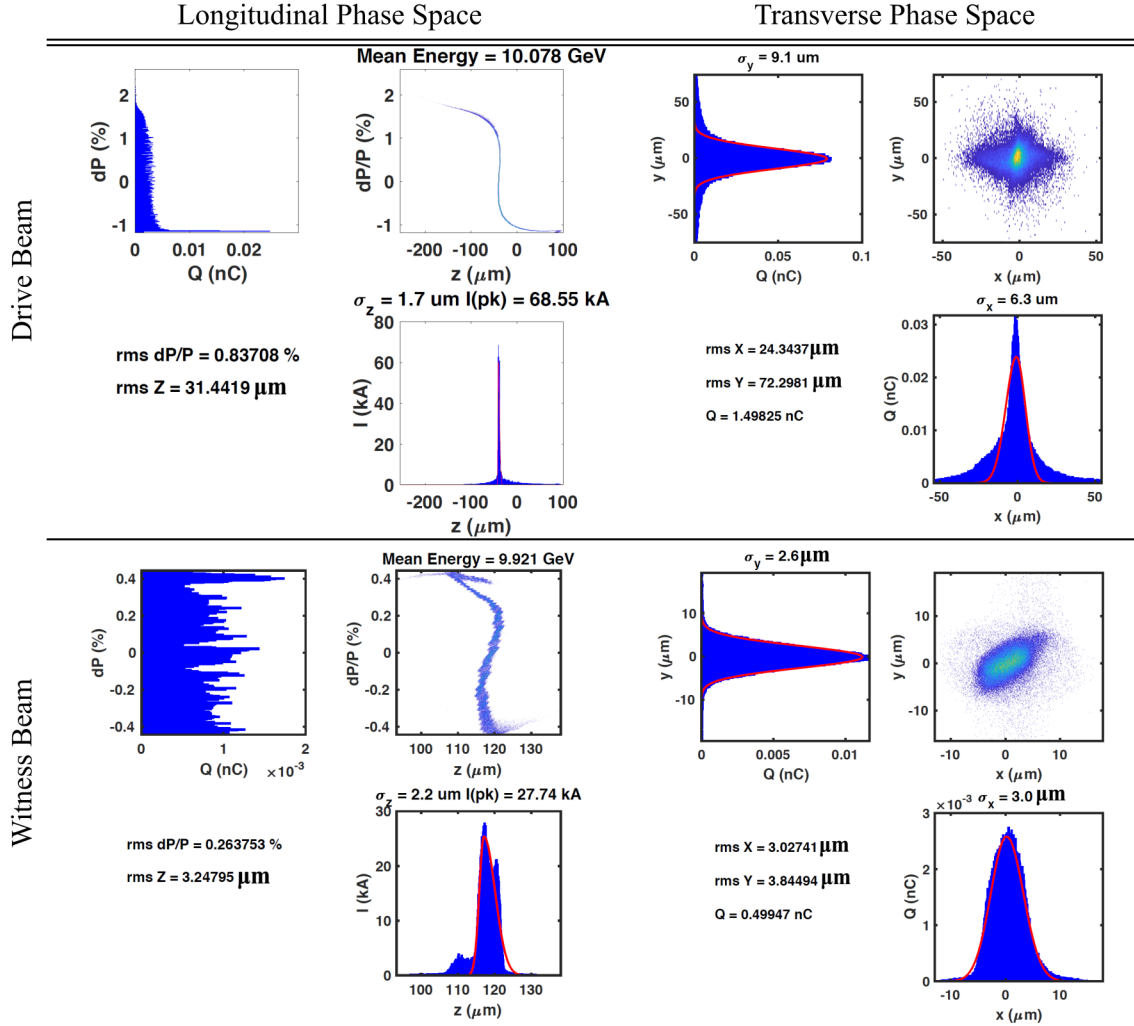


Figure 5.9: Beam parameters at the IP for the nominal two bunch configuration. The RMS values are the normal statistical definition applied to the full distribution while the σ values above each plot are the fit parameters for the Gaussian curve shown in red. Reproduced from [122, 63].

size by five families of quadrupoles that comprise the final focus. The beam then enters the plasma source where the plasma-based accelerator is located. Just before the beam arrives at the plasma, a laser pulse is sent through to pre-ionize the vapor source, comprised of either a lithium vapor oven or uniform gas pumped into the vacuum system. After exiting the plasma, the beam enters an imaging spectrometer which measures the energy distribution of the beam. Subsequently, the beam leaves vacuum and is stopped by the beam dump.

The IP area itself, stretching from the end of the final focusing optics to the start of the

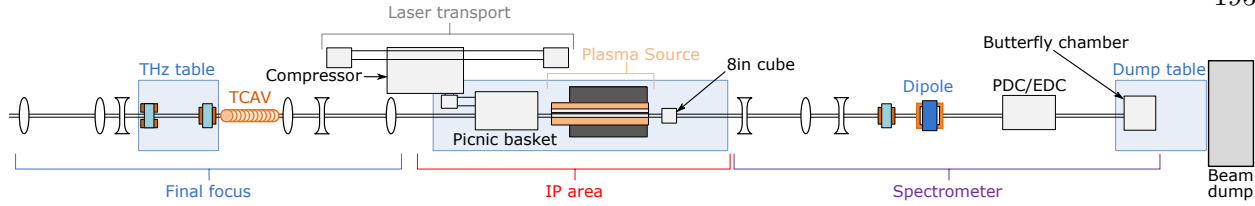


Figure 5.10: Schematic of the FACET-II experimental area. The electron beam enters from the left and is focused to the plasma source using the final focus quadrupoles. The laser enters the laser transport from a penetration in the ceiling of the tunnel before traveling to the upstream side of the compressor. After compression, the laser is focused and injected co-linear to the electron beam in the picnic basket. The laser is extracted from the 8 in cube and the electron beam continues into the spectrometer before being dumped.

spectrometer, is one of the most complex parts of the accelerator. It is designed to operate and switch between seven or more experiments remotely and is where the experimental laser system interacts with the electron beam. The accelerator tunnel is sealed up during operation because of the high levels of radiation, so the laser systems must then be aligned remotely. On top of this, LCLS-II sends its beam through a transport line on the roof of the tunnel. This means the tunnel can only be accessed when both LCLS-II and FACET-II are turned off. Because these outages are not frequent enough for the experimental schedule, the hardware in the IP area is designed to switch between experiments remotely using over 120 motorized degrees of freedom.

The layout of the IP area is summarized in Fig. 5.11. The electron beam enters from the left. A pair of beryllium windows—thin sheets of beryllium in the beam line—create an impermeable barrier isolating the IP area vacuum from the surrounding vacuum. The electron beam can pass through the beryllium windows with only a modest increase in emittance due to multiple scattering. The region in between the windows is back filled with gas during PWFA experiments. There are toroids at the upstream and downstream end of the IP area to measure the beam charge. Throughout the IP area are a series of optical transition radiation (OTR) screens that can be lowered into the beam and viewed with a camera to find the beam position and profile. In addition, there are a series of beam profile monitors that allow for high precision measurements of the beam size. The laser enters the tunnel through a penetration in the ceiling before travelling upstream to

enter the laser compressor. More details about the laser system are given in Sec. 5.3.

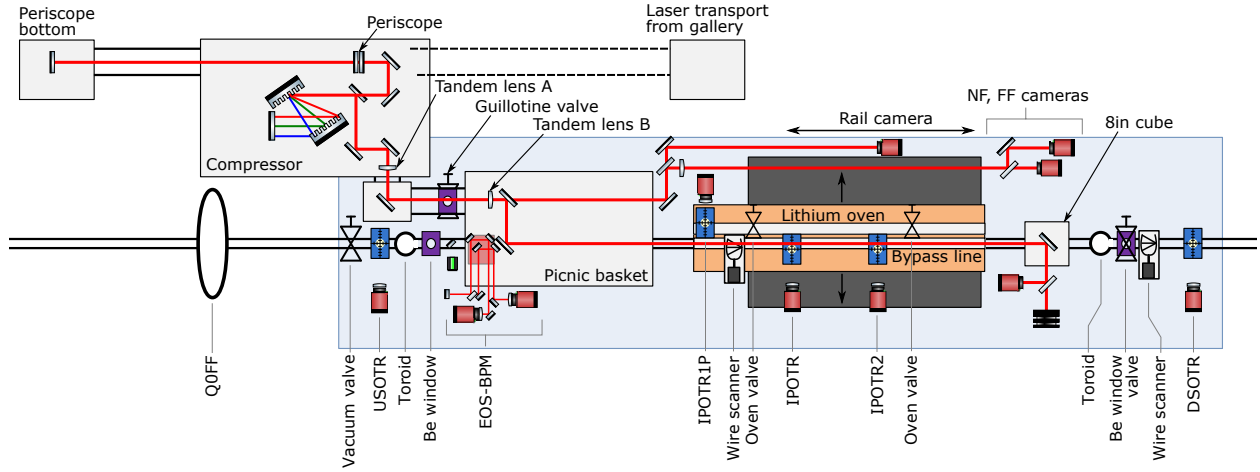


Figure 5.11: Schematic of the FACET-II IP area showing the two main plasma sources (lithium oven and bypass line), the laser path, and diagnostics for both the electron beam and laser.

The vacuum system between the two beryllium windows has two main chambers, the plasma sources, and an 8 in cube. The large vacuum chamber next to the electron beamline is the compressor chamber (Fig. 5.12(a)). It contains the laser compressor, a removable optic mount, and a beam block for the main laser. The compressor vacuum can be isolated from the rest of the IP area by a gate valve with a window, called the guillotine valve, that sits between the compressor and the second large vacuum chamber. This second chamber is called the picnic basket, Fig. 5.12(b), and has two functions. It is the laser integration chamber which houses the optics to align the main laser co-linearly with the electron beam. The picnic basket is also the main experimental chamber and contains hardware for most of the planned experiments at FACET-II. The chamber contains a couple things relevant to PWFA experiments. A motorized mount for ionization optics sits near the port where the laser enters the chamber, and the electro-optic sampling beam position monitor is located along the electron beam axis at the upstream side of the chamber. The other apparatuses in the chamber allow solid targets and gas jet targets to be moved into the path of the electron beam. The large apparatus in the center allows the electron beam to be collided with a tightly focused laser pulse. Additional optical setups exist to probe the interactions occurring in the various beam targets.

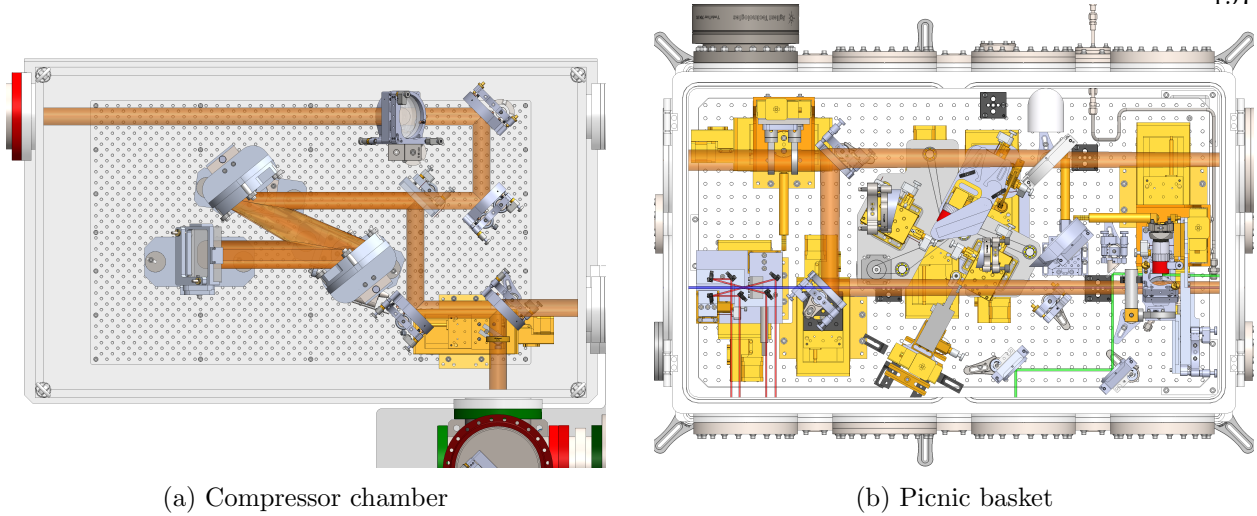


Figure 5.12: The vacuum chambers where the laser is compressed and focused. The laser is compressed in the compressor chamber (a) and passes through the first tandem optic. The picnic basket (b) is the main experimental chamber on the beamline and allows remote operation and switching between experiments. It houses the second tandem lens and the optics for injecting the laser along the electron beam.

Just downstream of the picnic basket is the plasma source for PWFA experiments. Two plasma sources are installed, a lithium vapor oven and the bypass line. The bypass line is so called because it provides a means for the electron beam to bypass the lithium oven. It can be used to house a laser-ionized gas plasma source in a flooded chamber configuration. The bypass line is filled to a designated static pressure and then the plasma source is created via laser ionization. The two sources exist on separate beam lines placed side by side on top of a lateral mover. They are joined to the rest of the beamline by two large bellows which allow the system to translate. The lithium oven is built by K. Marsh at UCLA, a detailed description can be found in Ref. [116]. In brief, the oven works by vaporizing solid lithium into a gas at high temperature. The oven starts filled with helium gas and solid lithium around the walls of the oven. As the oven heats up, the lithium is vaporized and the vapor pressure begins to displace some of the helium from the center of the oven. The lithium flows outward where it is cooled by the helium gas until it condenses on the walls of the oven and is wicked back towards the center by a mesh placed on the inside of the oven. Water jackets are placed on either side of the oven to cool the helium and ensure the lithium

condenses. The lithium vapor pressure is in equilibrium with the helium pressure, the helium pressure determines the bulk density of the plasma. The length and ramp shapes are determined by the temperature distribution, adjustable by changing the heater power.

5.2.4 Diagnostics

We briefly mentioned toroids, OTRs and wire scanners as beam diagnostic devices. There are several more complicated diagnostics installed on the beamline capable of looking at the longitudinal phase space of the beam. The first is an imaging spectrometer that uses three quadrupoles and a vertical dipole to image the beam onto a screen with vertical dispersion. The vertical direction on the screen then corresponds to the particle energy and the horizontal direction corresponds to the transverse size of the beam. The spectrometer only images in the horizontal plane, in the vertical plane it focuses the beam to a small beta function to maximize the energy resolution. The spectrometer can be set to image different locations in the IP area allowing energy resolved scans of the beam size after the plasma source. The right side of Fig. 5.10 is a diagram of the spectrometer setup.

Two screens are available for measuring the beam spectrum: a scintillating YAG screen or a polished titanium screen for OTR. Two cameras view the screen: DTOTR1 is a small field of view high resolution camera with about a 2 GeV field of view that can be moved to view any part of the spectrum, DTOTR2 is a large field of view camera capable of viewing the full spectrum from 0 to 10 GeV. A simulated spectrometer image is shown in Fig. 5.13(a). The screens are located in vacuum in the butterfly chamber at the end of the beamline.

The spectrometer is only able to image a single energy at a time onto the DTOTR screen. The resulting image takes on the butterfly shape where the image on the screen is narrow at the imaged energy and widens from there. The variation can be quantified by the function $\sigma_x(\delta)$, which is easily extracted from the DTOTR image. The variation of the beam size with energy can be easily calculated using the transport matrix formalism. The transport matrix M of the beam from the exit of the plasma source to the spectrometer screen is found by using Eq. (2.124) to

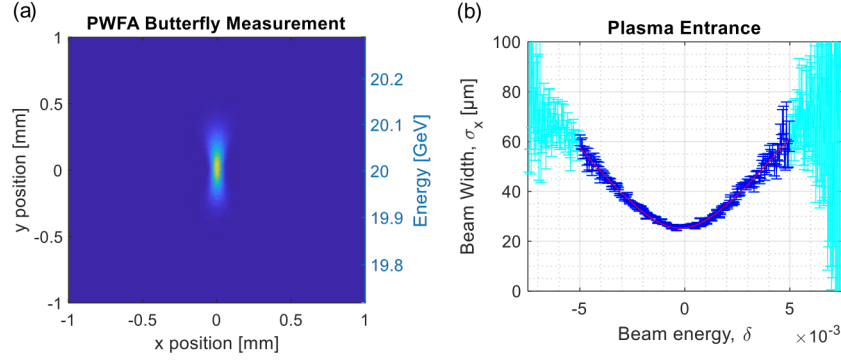


Figure 5.13: Imaging spectrometer data. (a) typical DTOTR1 image of the beam at the end of the imaging spectrometer. The quadrupoles are set up to image 20 GeV while other energy slices are out of focus creating the typical butterfly shape. (b) width of each energy slice. Fitting Eq. (5.1) to the data can be used to find the beam emittance assuming no correlations. Reproduced from Ref. [123].

combine the transport matrices of the individual elements given by Eq. (2.125). In general, the transport matrix is a function of energy $M(\delta)$. If we assume all the energy slices start with the same emittance and CS parameters at the object plane of the imaging spectrometer, then the final spot size of the beam is given by combining Eqs. (2.140) and (2.155):

$$\sigma_x^2(\delta) = \frac{\epsilon_n}{\gamma_b} \left[M_{11}^2(\delta)\beta_0 - 2M_{11}(\delta)M_{12}(\delta)\alpha_0 + M_{12}^2(\delta) \left(\frac{1 - \alpha_0^2}{\beta_0} \right) \right]. \quad (5.1)$$

The emittance is found by fitting Eq. (5.1) to the the measured $\sigma_x(\delta)$ with ϵ_n , β_0 , and α_0 as fitting parameters.

As we pointed out in Sec. 3.4.5, the energy slice parameters will vary across the slices. The functional form of the dependence cannot be extracted from the butterfly without more information. This can be overcome by doing a so-called quad scan, where the focus of the quadrupole is scanned across a range of energies. The scan results in a beam size curve for each energy slice as a function of quadrupole parameters. The emittance and CS parameters for each slice can then be found via fitting. The disadvantage of this technique is that it cannot be accomplished in a single shot. As such, it relies on the beam being stable shot-to-shot for the duration of the measurement, a difficult proposition for many PWFA experiments.

In the middle of the final focusing optics, just downstream of the THz table, is the transverse

deflecting cavity (TCAV). The TCAV is a travelling wave RF-cavity, much like the accelerating structures used in the linac but the electric field is oriented in the transverse direction rather than the longitudinal direction. The TCAV is normally operated with the beam placed near the zero crossing of the RF phase. The result is a transverse kick on the beam that depends linearly on ξ . The kick effectively maps the ξ dimension onto the x dimension (nonzero R_{15}). If the beam is then sent through a dipole, it will be dispersed in y according to energy, effectively mapping the δ dimension onto the y dimension (nonzero R_{36}). Measuring the transverse profile using DTOTR then provides a measurement of the longitudinal profile. The TCAV is located upstream of the IP area which means its use requires kicking the beam before it reaches the plasma source. Unfortunately, the transverse kick from the TCAV renders it unusable during PWFA experiments. As such, the TCAV is considered a destructive measurement: either it or the experiment can be used, but not both at the same time. The EOS-BPM described below provides a more limited measurement of the current profile non-destructively.

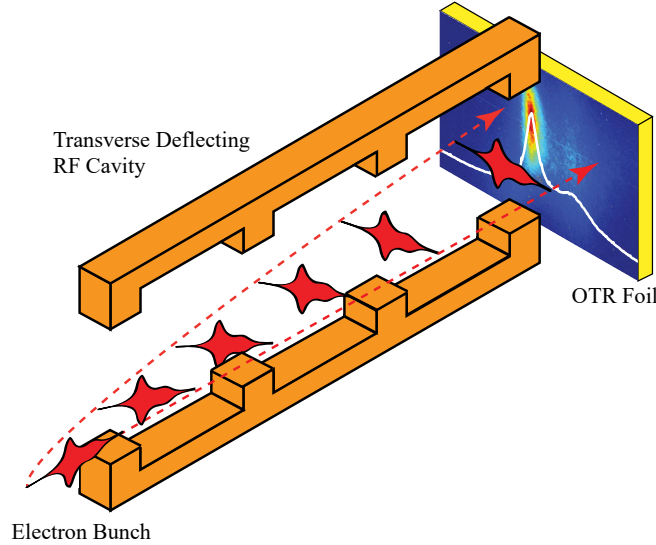


Figure 5.14: The transverse deflecting cavity can be used to measure the full longitudinal phase space. In the cavity, a RF pulse with a transverse electric field co-propagates with the electron beam. The beam sits at the zero crossing of the RF phase, so the front of the pulse is deflected transversely to the right and the rear of the pulse is deflected to the left. After sending the beam through a dipole, it is dispersed vertically in energy. The result is a beam where the longitudinal phase space has been mapped onto the transverse phase space, which can be measured with an OTR. Reproduced from Ref.[124].

The drive and witness bunches are separated by about $150\text{ }\mu\text{m}$ or about 500 fs . A typical beam position monitor (BPM) uses the signal difference between four electrodes spaced every 90° around the beamline to determine the location of the beam. Time scales in the hundreds of femtoseconds are three orders of magnitude lower the characteristic response time of the electronics so there is no way for a traditional BPM to distinguish between the drive and witness beam. Lasers, however, can routinely produce pulses on the order of 10 fs , which is more than fast enough to resolve the temporal structure of the dual electron bunches. Electro-optic sampling (EOS) is a technique for imprinting the temporal structure of the electron beam onto the laser pulse. The electro-optic sampling beam position monitor (EOS-BPM) places an EOS on each side of the beamline and looks at the signal difference, similar in concept to a traditional BPM. The EOS-BPM was developed out of our group at University of Colorado, more details can be found in Ref. [125].

EOS sampling works using the electro-optic effect in a non-linear medium. The transverse electric field from the electron beam induces birefringence in a nonlinear crystal—typically GaP or ZnTe—placed near the electron beam axis. The birefringence exists only during the electron beam’s passage. A horizontally polarized ultrashort pulse is sent through the crystal at an angle to probe the birefringence in an approach called spatial encoding, Fig. 5.15(a) is a diagram of the setup. The birefringence rotates the polarization of the laser when the electron beam is next to the crystal. By sending the laser in at an angle, the longitudinal beam structure is mapped onto the transverse polarization of the laser. Sending the laser through a vertical polarizer and onto a camera produces a signal like the one shown in Fig. 5.15(b). The spacing between the vertical stripes corresponds, with appropriate calibration, to the spacing between the electron bunches. The EOS-BPM apparatus installed at FACET-II is shown in Fig. 5.15(c) where two identical spatial encoding setups are installed on either side of the beamline. The difference between the signals from the two EOS arms can be used to independently measure the transverse offset of the witness beam and drive beam.

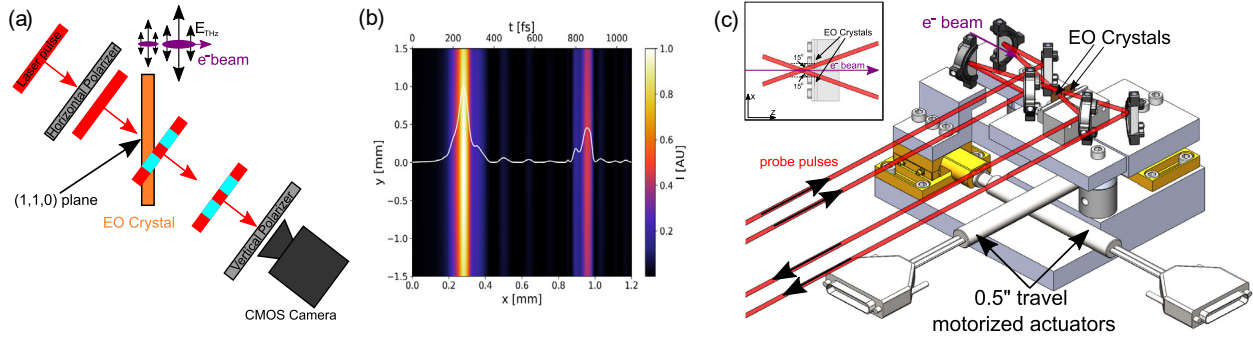


Figure 5.15: Electro-optic beam position monitor. Temporal information is encoded in the laser pulse through the spatial encoding scheme. (a) a horizontally polarized ultrashort laser (red) passes through an electro-optic crystal (orange) at an angle. The section of the laser within the crystal when the electric field of the electron beam (purple) passes will have its polarization rotated (blue). Using a camera with a vertical polarizer will produce an image of the longitudinal beam structure (b). A beam position monitor is created by placing a spatially encoded EOS on either side of the beam and looking at the signal difference (c). Reproduced from Ref. [125].

5.3 FACET-II Laser System

The experimental laser at FACET-II is very similar in layout to the system at CU, described in Sec. 5.1. Details of the original laser are presented in Green *et al.* [126]. The transport system after the main amplifier was completely rebuilt for FACET-II with more diagnostics and the addition of a deformable mirror. The laser is an 800 nm Ti:sapph based CPA system with a regenerative amplifier and two main amplifiers. The seed laser, amplifiers, and pump lasers are located at ground level in the FACET-II S20 laser room. After amplification, the laser is expanded onto a deformable mirror before it is launched into a transport line to the accelerator tunnel. The laser is transported out of the laser room, across the top of the Klystron gallery, down through a penetration to the accelerator housing and then up the tunnel to the compressor, where the first tandem lens can be inserted. From the compressor, the laser travels into the picnic basket where the second tandem lens can be inserted, and the laser is made co-linear with the electron beam. Finally, the laser is extracted from the electron beam in the 8in cube and dumped.

5.3.1 CPA System

The system is composed of an oscillator (Coherent Inc. Vitara-T), regenerative amplifier (regen, Coherent Inc. Legend Elite HE USP), multi-pass pre-amplifier (pre-amp), and multi-pass main amplifier (main-amp). The oscillator produces ultrashort pulses with 800 nm central wavelength and 60 nm FWHM bandwidth at 68 MHz. The pulse train from the oscillator is phase locked to the master RF signal from the SLAC linac to synchronize the laser to the accelerator. Measured RF to laser jitter is < 70 fs while jitter between the laser and electron beam is slightly worse due to variation in the electron beam arrival time. The pulses from the oscillator are stretched within the Legend Elite before entering the regenerative amplifier which amplifies the pulses to 1 mJ. Simultaneously, the regen acts as a pulse picker reducing the repetition rate to 120 Hz. A standalone pulse picker after the regen reduces the repetition rate down to either 1 or 10 Hz for full amplification.

The pulses are initially amplified in the four-pass preamplifier to an energy of 30 mJ. The pre-amp is pumped by a 130 mJ Nd:YAG laser (Quantel CFR200) frequency doubled to 532 nm. The pulses then enter the main amp where they are amplified to 1 J in four passes. The Ti:sapph crystal in the main amp is pumped from both sides by a pair (one laser incident on each side) of 1.8 J, 532 nm frequency doubled Nd:YAG lasers (Thales SAGA). After amplification the bandwidth is narrowed to around 24 nm FWHM. The beam exits the final amplifier with a diameter of around 10 mm. The optical layout is illustrated in Fig. 5.16. The entire laser system fits on two optical tables in the FACET-II S20 laser room.

The laser system has cameras and motorized mirrors allowing remote realignment to compensate for long term drift. The ability to align remotely allows adjustments to the laser to be done from the accelerator control room without having to have someone in the laser room as well. On top of this, the temperature stability in the FACET-II S20 laser room is less than ideal and the laser tends to drift depending on the outside temperature. The motorized system is able to compensate for this slow drift automatically.

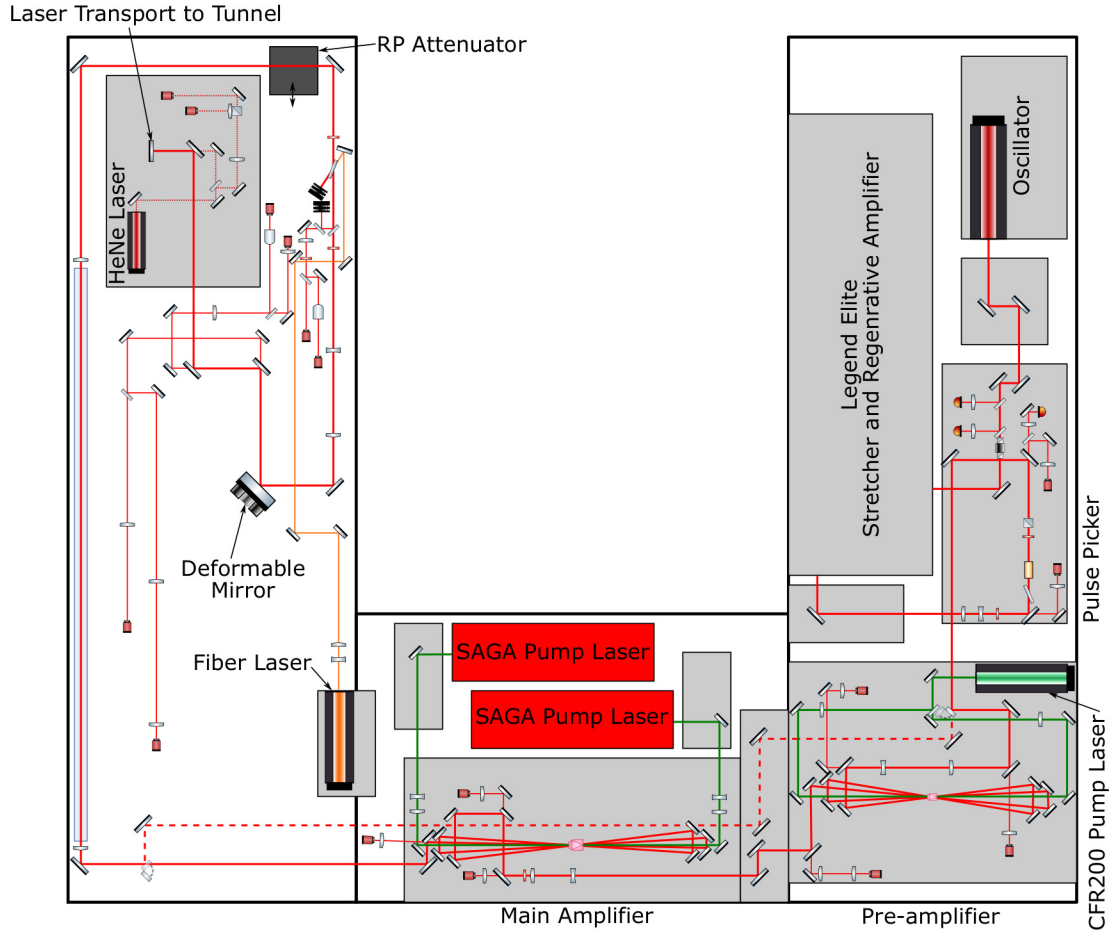


Figure 5.16: Overview of the experimental laser at FACET-II. The system starts with a commercial oscillator and amplifier before entering a sequence of two multi-pass amplifiers. The laser is imaged relayed to an energy attenuator after which the beam is sampled for diagnostic purposes. A deformable mirror corrects the wavefront of the pulse before the laser is sent into the 30.5 m vacuum transport to the accelerator.

5.3.2 Deformable Mirror and Transport

The output of the main laser passes through a series of safety systems, diagnostics, and beam shaping devices before reaching the laser transport. A schematic of the laser system between the main amplifier and the entrance to the transport is shown in Fig. 5.17. After the main-amp, the laser enters a 4f image relay telescope to preserve the transverse profile. After the second telescope lens, there is the radiation protection (RP) attenuator, which must be inserted into the laser whenever people are working with the laser in the tunnel to prevent the unintentional generation

of ionizing radiation. The laser then passes through an energy attenuator composed of a waveplate and polarizer before being expanded from a 10 mm to a 40 mm diameter by a diverging-converging telescope. At this point the pulses are expanded to the size necessary for the compressor. The beam size can be adjusted using a motorized iris after the telescope.²

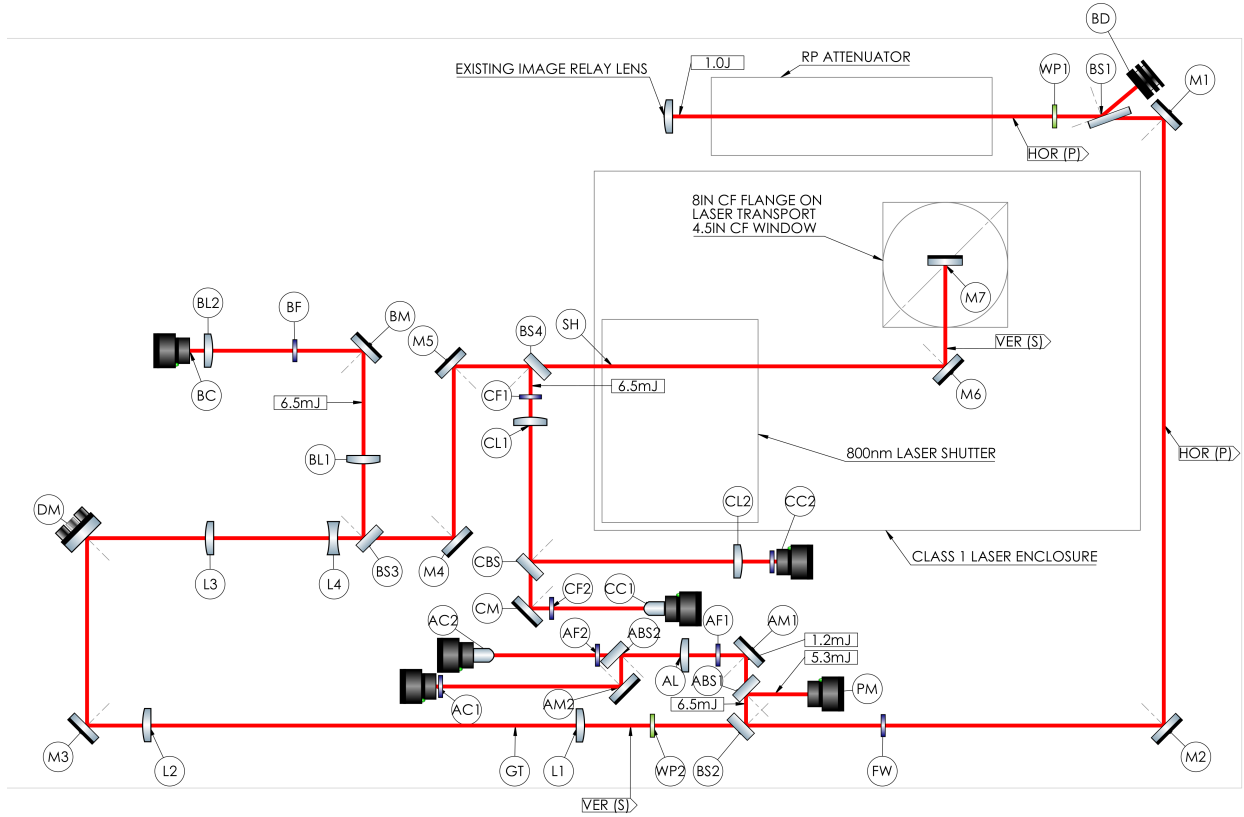


Figure 5.17: Schematic of the so called transport section of the laser system, i.e. everything in the laser room after the main amplifier. After the second multi-pass amplifier, the laser is imaged relayed to an energy attenuator after which the beam is sampled for diagnostic purposes. A deformable mirror corrects the wavefront of the pulse before the laser is sent into the 30.5 m vacuum transport to the accelerator. A HeNe laser is injected co-linear with the 800 nm beam to act as a reference for the alignment feedback system when the main laser is off.

The DM has to be placed after the telescope to ensure its entire surface is filled. In this case, it is placed directly after the telescope. Upstream of the converging-diverging telescope a beam sampler picks off 0.65% of the beam that is sent to a near field and far field imaging system to

² At full energy, the laser diameter needs to be large enough such that the intensity does not damage the compressor grating. Reducing the laser size with an iris removes energy from the beam but does not change the intensity.

monitor laser steering onto the DM. Most of the energy in the sampled beam (81.8%) is sent to an energy meter to measure the pulse energy. After the DM, the laser reflects from a pair of mirrors into an enclosure directly before the laser transport. A 4f imaging system images the surface of the DM onto a Shack-Hartmann wavefront sensor using the leakage light through the first mirror after the DM. The DM functions in closed loop with the wavefront sensor to produce a target wavefront in the laser.

Leakage light from the second mirror after the DM is sent into a near field and far field imaging system to monitor the alignment of the laser as it is launched into the transport. The transport itself is an approximately 30.5 m vacuum line with seven bends that is part of the same vacuum volume as the compressor in the tunnel. The transport rises vertically from the optics table in the laser room, turns (B0) and exits through the wall to the Klystron gallery, where it turns (B1) to travel upstream along the wall of the gallery to the z location of the penetration to the S20 experimental area. The line turns (B2) and crosses over the walkway in the gallery before turning (B3) downward and dropping 11 m into the tunnel. From there it turns (B4) upstream and travels along the tunnel wall before passing through a periscope (B5 and B6) to bring it to the entrance height of the compressor. Within the transport are two 4f image relay telescopes to image the beam from the DM to the entrance of the compressor. The transport must be at vacuum because the intensity at both of the telescope foci is high enough to ionize air; the resulting plasma would ruin the transverse profile of the beam.

The transport traverses three different temperature zones: it starts in the—nominally—temperature-controlled laser room, passes through the uncontrolled Klystron gallery, and finally enters the tunnel, which stays warm due to the $> 80^{\circ}\text{F}$ temperature set-point of the linac chilling water. The gallery temperature fluctuates significantly with the outside ambient temperature and the tunnel undergoes significant temperature cycles each time the linac is turned on and off. This is of course terrible for maintaining alignment. On top of this, the supports for the transport line are affixed to the structure of the gallery. In high winds, the structure flexes significantly causing misalignment of the transport.

To help maintain some level of alignment in the laser transport, an automated alignment system is installed to correct for drifts of the transport system. All the mirror mounts in the transport are motorized to allow for realignment while under vacuum. A diffuser is placed on the back of each mirror and viewed by a camera mounted to the outside of the transport line to measure the alignment at each bend box. The cameras in the transport system look at the position of the laser on each mirror so the alignment can only be corrected when the laser is on. When the laser system is turned off, the transport can shift far enough that the beam is no longer visible on the downstream mirrors. Recovering from this situation typically requires venting the transport to realign. To prevent this issue, a HeNe laser co-propagates with the IR laser and is left on at all times to provide a reference for maintaining alignment.

5.3.3 Laser System in the Tunnel

The laser enters the tunnel through a vertical penetration around the center of the IP area optical tables, see Fig. 5.11. The bend (B4) at the bottom of the penetration has a beamsplitter instead of a mirror to pick off a probe line from the main pulse. The transmitted probe exits the transport vacuum through a window and is sent to the optical table where it is compressed separate from the main pulse. The probe laser is used for EOS-BPM as well as several other experiments.

The main laser enters the compressor chamber from the transport. A motorized beam block can be inserted to allow operation of the probe laser independent of the main laser. The laser compressor, composed of a pair of diffraction gratings and a retro-reflector, follows the beam block and is able to compress the pulses down to 50 fs FWHM duration. The compressor efficiency was measured to be 72%. The overall transport efficiency from the main amplifier to the exit of the compressor chamber is around 45% with about half the losses in the compressor. The other large sources of loss are the polarizing beamsplitter in the energy attenuator and the beamsplitter for the probe beam. Leakage from the last mirror in the compressor is imaged with a near field and far field camera system as an alignment reference through the compressor. Just before the exit of the compressor, the mount for the first tandem optic is placed on a 100 mm translation stage oriented

transverse to the laser path. The stage allows the lens to be remotely removed and inserted into the beam line and allows alignment in the horizontal direction. Figure 5.12(a) shows the optical setup within the compressor chamber.

The beam exits the compressor, enters a conflat cross (called the funky cross) where it reflects off a mirror directing it downstream towards the picnic basket. Between the funky cross and the picnic basket is the guillotine valve that can be used to isolate the compressor from the experimental area when the lithium oven is operating or when the experimental area is operating as a flooded chamber. The laser then enters the picnic basket where it passes through the second tandem optic, mounted on a 3-axis translation stage. After the lens the beam reflects off a pair of mirrors to bring it co-linear with the electron beam. The second mirror is called the upstream holed mirror (USHM) and has a 5mm diameter hole in it for the electron beam to pass through. Leakage through the first mirror is sent to a camera on a 2 m rail in order to measure the longitudinal intensity profile of the laser after passing through the tandem optics. Some of this leakage light is picked off to send to a set of near and far field cameras.

The laser reflects from the upstream holed mirror into the plasma source: either the lithium oven or the bypass line filled with gas. After the plasma source, the laser is extracted from the beamline by the downstream holed mirror located in the 8in cube. The laser is sampled for a final near field and far field imaging system before being dumped. Fig. 5.11 shows a schematic of the main laser in the compressor chamber, picnic basket, 8in cube, and diagnostics on the IP area table.

Chapter 6

Experimental Results

The first step in demonstrating emittance preservation in a PWFA is to demonstrate the optical ionization of a suitable plasma source. Verifying that the laser ionization scheme works starts by verifying that the tandem optics technique, developed in Ch. 4, works as intended. An initial prototype, designed to ionize the lithium oven at FACET-II, was tested in our lab at the University of Colorado. In this chapter, we report on the results of these initial optical tests. The two tandem lenses are denoted by the letters A and B, with A being the first lens and B being the second lens. The designs are numbered by the purpose and revision number. The optics tested here are called LO8 for Lithium Oven revision 8. Lens and B are then LO8-A and LO8-B, respectively. For details on the lens design see Sec. 4.7.2. For details of the test setup see Sec. 5.1.4.

6.1 Intensity Profile Extraction

Testing the tandem lens optics entails measuring the full 3D intensity pattern downstream of the optic using a camera mounted on a 2.2m translation stage. The linear translation stage, however, is not perfectly straight. The laser position on the camera thus varies as a function of z . On top of this, laser pointing jitter causes the laser spot to move around from shot to shot. Because the jitter is angular, the offset of the laser spot due to jitter increases linearly with distance from the second tandem lens. Reconstructing the intensity distribution then requires transversely aligning the images of each individual laser shots to all the others. We use two different approaches to do the alignment. For the lens A measurements, the pointing jitter is a small percentage of the

beam size and is not important in comparison to the nonlinearities in the rail, which are magnified by the $M = 0.1$ imaging system. Aligning the images is then a simple matter of subtracting the z dependent offset introduced by the rail. Measuring the offset is done by taking images of the laser, without lens A installed, along the entire length of the translation stage. The location of the beam center on the image at each z position then corresponds to the offset of the image plane at that z position. In addition, because the pointing jitter is negligible, all the shots taken at a single z position can be directly averaged together.

For the higher resolution images of the Bessel focus, the offset due to pointing jitter is important. In this case, we numerically find the center of the beam in each image and use this to center all the images before averaging. The disadvantage of this approach is that defining the beam center is not always straightforward. The center of a Bessel function is well defined, so this approach works well in the bulk of the high intensity region. Where it breaks down is near the end where wavefront aberrations in the beam destroy the Bessel and produce strange intensity patterns. In these areas, the center of the beam is not so well defined. In the data, this appears as an apparent curvature in the intensity pattern near the end of the focus.

6.2 Characterization of Lens A

The tandem lens is designed to produce a target transverse phase and intensity in the laser pulse immediately after the second lens. Lens A shapes the intensity and lens B shapes the phase. The two lenses are separated by 1 m so the laser intensity 1 m downstream of lens A should match the target intensity. Unfortunately, the LO8 tandem lens was designed for the FACET-II laser which has a 40 mm diameter beam, rather than the 30 mm beam used for these tests. The optics were designed assuming a super-Gaussian intensity profile described by Eq. (4.54) with $m = 8$ and $w_0 = 24$ mm. The beam profile of the test laser, shown in Fig. 5.2(a), is not only too small, but not very super-Gaussian either.

Clearly, we cannot compare the experimental data to the target intensity the optics were designed to create because the incoming laser beam differs significantly from the intended profile.

Instead, we measured the intensity of the incoming beam and used this as the input to a Fourier optics simulation of the lens pair. The laser intensity at $z = 1$ m calculated in the simulation was then compared to the measured intensity to see if the optic was performing as expected. Both simulated and measured intensity are shown in Fig. 6.1. The measured and simulated intensity match reasonably well. The outer diameter of the measured intensity is wider than expected, but the overall profile shows good agreement. The discrepancies are most likely due to poor collimation in the incoming beam, discussed in more detail in Sec. 6.3.3.

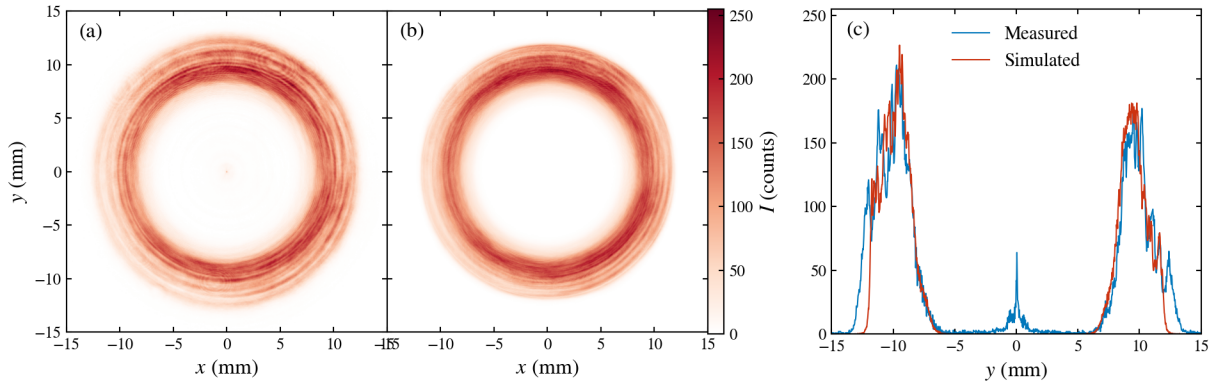


Figure 6.1: Intensity produced by the first tandem lens measured at the second tandem lens. (a) measured intensity at lens B ($z = 1$ m). (b) simulated intensity at lens B using the measured intensity at lens A ($z = 0$ m) as the initial conditions. (c) vertical line-out through the center of each image.

We also compared the full intensity profile to simulation. Images of the laser profile were taken in 5 cm steps starting 10 cm from the lens and ending 2.1 m from the lens. The resulting profile is compared to simulation in Fig. 6.2. Looking at the $y - z$ plane, the measured intensity is asymmetric compared to the simulations. In simulation, optic A acts almost like an annular lens, focusing the beam into a thin ring about 1.4 m from the lens. The measured data shows a ϕ dependent focal length for the ring focus. In a later test, the optic was removed from the beam path and the laser position was recorded at the location of optic A using a camera. The lens was then reinstalled and carefully aligned to the center of the incoming beam. After realigning, a sweep of intensity along z was taken and the focal length was more uniform as shown in Fig. 6.2(b). We

suspect that the alignment of the lens was to blame for the unexpected performance in the first test.

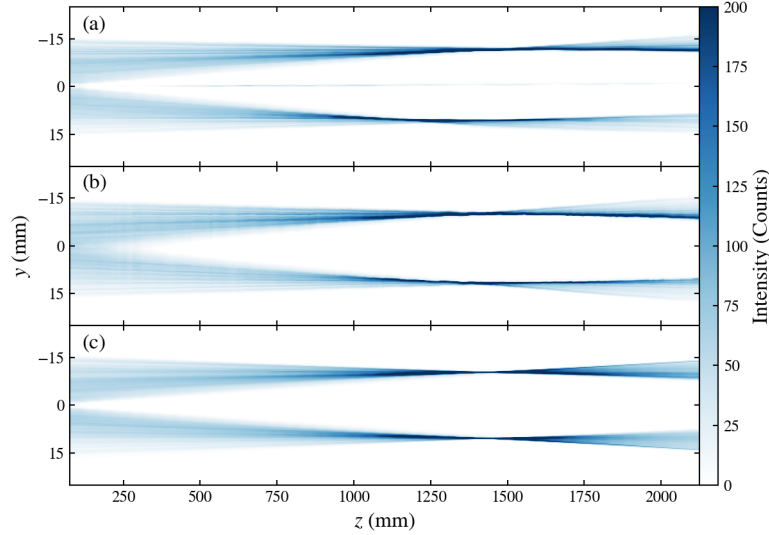


Figure 6.2: Intensity of the beam after passing through lens A. (a) measured intensity from 10 cm after the lens to 2.1 m after the lens. (b) the same as (a) but after realigning the lens. Notice that the upper and lower focus are now aligned in z . (c) simulation of the propagation after lens A.

For designing experiments, it is important to quantify the allowable misalignment of the optic. To this end, we measured and simulated the intensity of lens A when it was transversely offset from the laser beam. The results are shown in Fig. 6.3. As the lens is more and more offset, the ring becomes highly asymmetric. We would expect this to reduce the on-axis intensity as more of the laser light is sent into the $m = 2$ Bessel mode due to the breaking of azimuthal asymmetry. We revisit lens alignment tolerances in much more detail when we look at the output of the tandem lenses as a pair.

6.3 Characterization of Lens B

Lens B is difficult to characterize directly because the center of the optic is flat, it does not impart any phase on the laser. As a result, the light that passes through the center interferes with the light that is focused by the outside of the optic, making the resulting intensity patterns

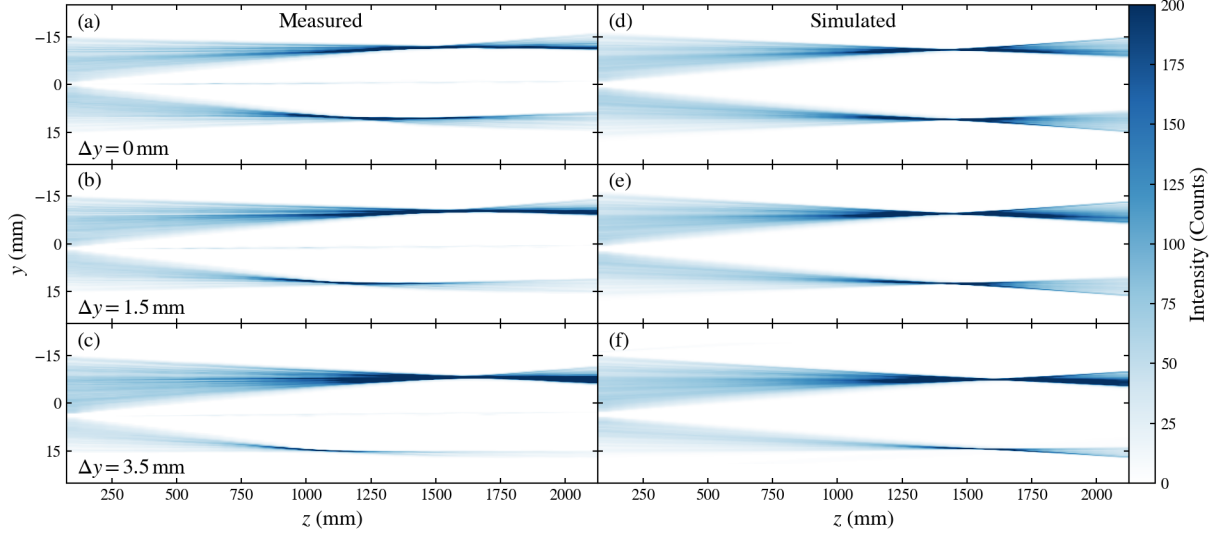


Figure 6.3: Laser intensity after lens A for different transverse offsets between the lens and the center of the beam. (a), (b), and (c) are the measured intensity at offsets of $\Delta y = 0$ mm, $\Delta y = 1.5$ mm, $\Delta y = 3.5$ mm, respectively. (d), (e), and (f) are the simulated intensities at the same offsets.

difficult to interpret. Instead, we characterize the lens by installing it with lens A and looking at the Bessel focus. We know from simulations that the alignment between the two lenses is sensitive on the $10\text{ }\mu\text{m}$ level and that a Bessel focus will not form at all if the alignment is worse than several $100\text{ }\mu\text{m}$. Conveniently, if the alignment is bad, the laser forms an interference pattern that looks like an arrow pointing in the direction the alignment needs to be adjusted. After aligning the lens, an image of the Bessel focus was taken 1.5 m downstream of lens B. The image is shown in Fig. 6.4. Fitting squared Bessel functions to the intensity results in a measured Bessel width of $R_b = (60.5\text{ }\mu\text{m}, 60.1\text{ }\mu\text{m})$ in (x, y) , which is within 1% of the design value of $R_b = 60\text{ }\mu\text{m}$. The width of the Bessel function did increase along the length of the plasma, reaching $R_b = (73.2\text{ }\mu\text{m}, 72.7\text{ }\mu\text{m})$ 2.5 m downstream of the lens.

The Bessel focus is sensitive to aberrations in the laser wavefront, particularly anything that breaks azimuthal symmetry. One convenient way of quantifying aberrations in the incoming laser is to decompose the wavefront error into Zernike polynomials. Typically, the error will be dominated by the lower order polynomials. The magnitude of the wavefront error for astigmatism and trefoil,

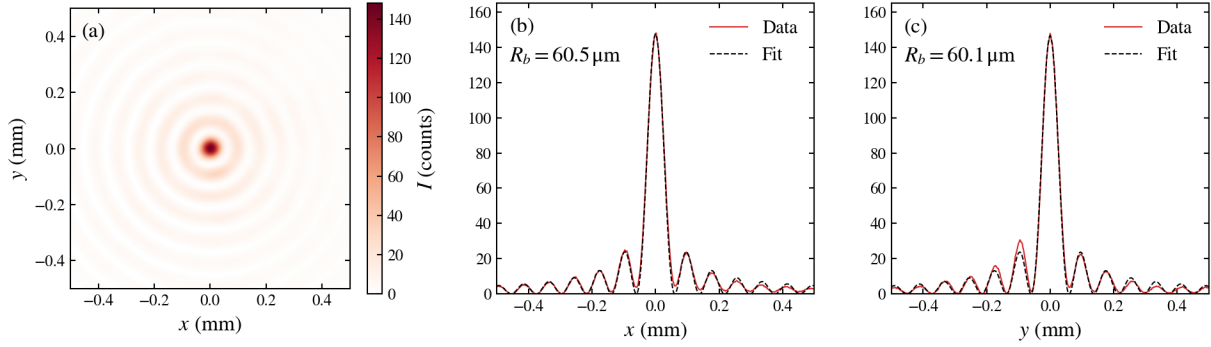


Figure 6.4: Intensity of the Bessel focus produced by the tandem lens. (a) shows an image of the focus. (b) and (c) are lineouts through the center of the focus in the x and y directions, respectively. The Bessel width is found by fitting a squared Bessel function to each curve. The optics are designed to produce a Bessel with $R_b = 60 \mu\text{m}$.

two of the more common aberrations, grows with r . Remember that large r corresponds to the end of the Bessel focus, so the wavefront errors will show up most noticeably near the end of the depth of focus. We can take advantage of this to correct the laser wavefront. Placing a camera at the end of the Bessel will show a multi-lobed pattern like those in Fig. 6.5. The number of lobes provides information about which aberration is present: two lobes indicate astigmatism or coma, while three lobes indicates trefoil. The wavefront errors can be corrected by adding the appropriate Zernike polynomial to the laser wavefront using the deformable mirror until the patterns disappear.

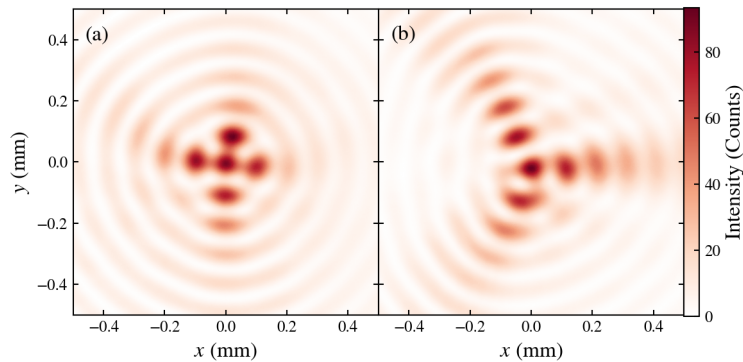


Figure 6.5: Bessel beam intensity with wavefront errors. Both plots are taken at a z location near the end of the focus, where the aberrations are most severe. (a) shows the aberration due to astigmatism, (b) due to trefoil.

The defining feature of the tandem optics is their ability to produce a prescribed intensity profile along the optical axis. Due to the difference between the laser profile used to test the optics and the laser profile the optics were designed to work with, the optics are not going to produce the target longitudinal intensity pattern. Nevertheless, we can compare to a simulation of the optics when illuminated by the test laser to see if the optics are behaving as expected. This comparison is shown in Fig. 6.6. The Bessel focus is the correct distance from the lens but is about twice as long as it should be. In the next sections, we measure the profile while intentionally misaligning the optics and adding aberrations into the beam. The goal is to determine if the discrepancy results from setup issues, aberrations in the incoming beam, or issues with the lenses themselves.

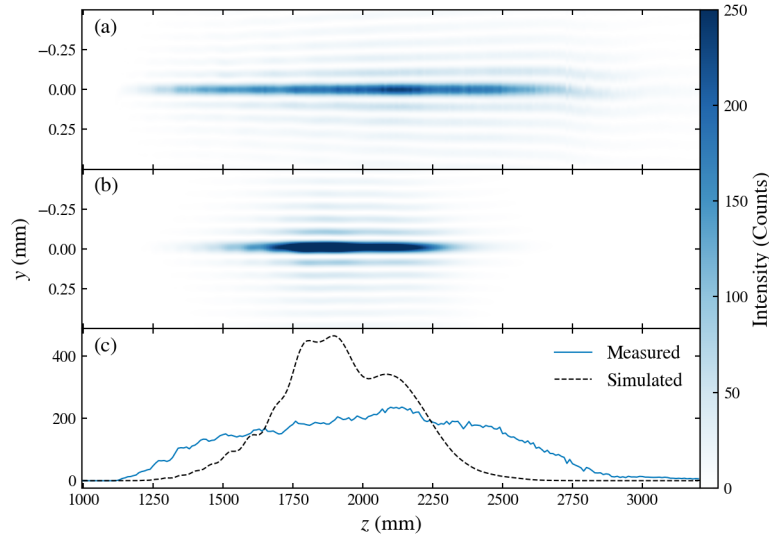


Figure 6.6: The longitudinal intensity profile produced by the tandem optics. (a) shows the measured intensity in the $y - z$ plane. (b) shows the simulated intensity using the measured transverse profile of the laser as input. (c) shows the discrepancy in the on-axis intensity.

6.3.1 Alignment Sensitivity of the Tandem Lens

Simulations predict that the tandem lenses are sensitive to the transverse alignment of lens B. The center of lens A, combined with the laser vector, sets the nominal beam path. The center of lens B needs to be aligned to this beam path with high tolerance. The allowed error that maintains

the longitudinal intensity shaping is expected to be on the order of $10\text{ }\mu\text{m}$. To measure the impact of position errors, we offset the lens between $2.5\text{ }\mu\text{m}$ and $50\text{ }\mu\text{m}$ and recorded the longitudinal intensity profile. The results are shown in Fig. 6.7(a) and (b). The tandem lens is less sensitive to transverse offsets than expected, although this is most likely due to how the tolerance was measured in simulation: the simulation results were not centered prior to finding the on-axis intensity.

The profile is expected to be sensitive to the spacing between the two lenses on length scales of several millimeters. If the lenses are too close, the length of the profile shortens and the intensity increases; if they are too far apart, the length of the profile contracts and the intensity decreases. The results of offsetting the lens from -10 mm to 5 mm are shown in Fig. 6.7(c) and (d) and the behavior qualitatively matches expectations. We do not believe an error in the spacing between the lenses can explain the discrepancies in the longitudinal profile. The lenses are mechanically mounted the correct distance from each other with a tolerance well within the range of adjustment we tested, yet the adjustment was not sufficient to correct the profile.

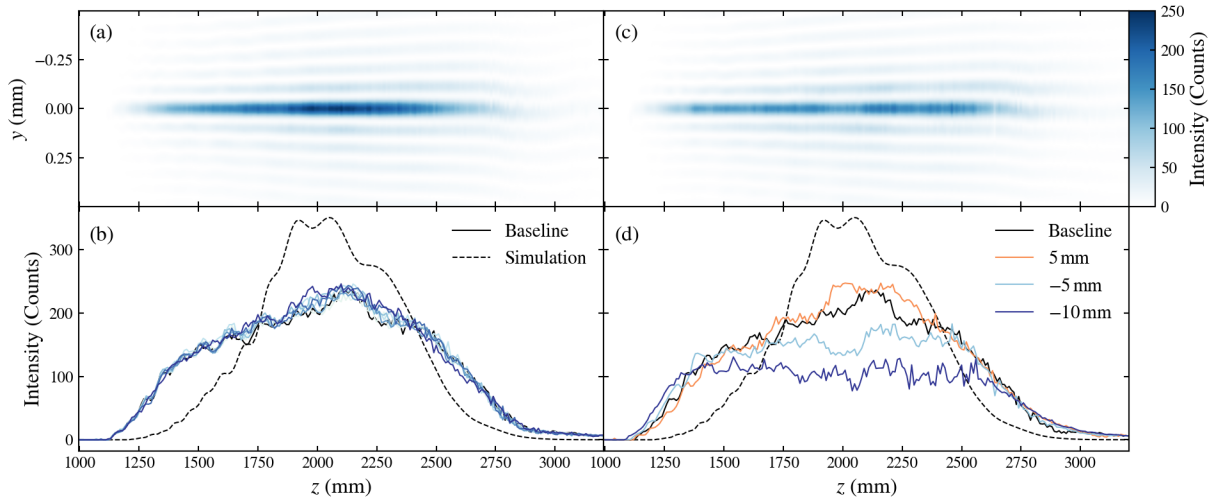


Figure 6.7: Sensitivity of the longitudinal intensity profile to transverse and longitudinal offsets of lens B with respect to lens A. (a) shows the longitudinal intensity for a transverse offset of $50\text{ }\mu\text{m}$ and (b) shows how the offset impacts the longitudinal profile. The tandem lens is less sensitive to transverse offsets between the lenses than originally anticipated. (c) shows the intensity when the distance between the two lenses is changed by -5 mm . (d) shows the on-axis intensity for different lens separations in increments of 5 mm .

6.3.2 Phase Sensitivity of the Tandem Lens

As discussed in Sec. 6.3, the Bessel beam is susceptible to wavefront errors in the incoming laser. Any non-azimuthally symmetric features in the beam will scatter light into higher order Bessel modes, which will not contribute to the on-axis intensity. Because our laser system has a deformable mirror, we can selectively add aberrations to the laser and see how it impacts the focal region. We chose the four most common Zernike polynomials that we typically see in laser systems: defocus, coma, astigmatism, and trefoil. For each aberration, we added it into the laser system in increasing amounts and measured the longitudinal intensity profile. The results for defocus and coma are shown in Fig. 6.8. All the aberrations we consider are symmetric except for defocus, which is asymmetric for positive and negative values.

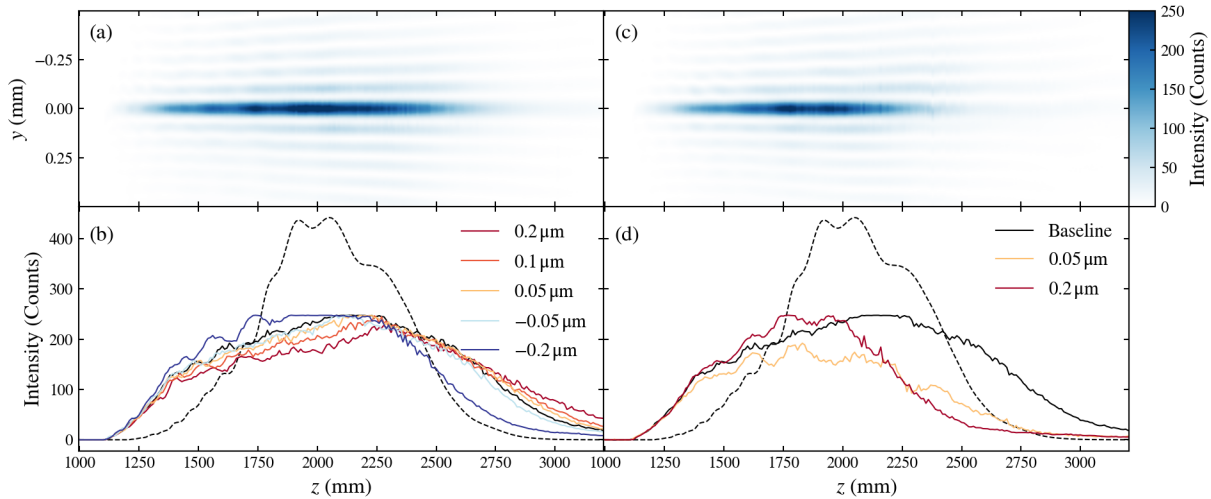


Figure 6.8: Laser intensity when the incoming laser has nonzero focus or coma. (a) intensity of the beam with P-V defocus of $-0.2 \mu\text{m}$. (b) on axis intensity for different amounts of focus aberration. The camera saturates at 255 counts, cutting off the large negative defocus case. (c) intensity with P-V coma of $0.2 \mu\text{m}$ of. (d) on axis intensity for different P-V amplitudes of coma.

Fig. 6.9 shows the change in the longitudinal intensity profile when adding astigmatism or trefoil errors into the incoming beam. The wavefront aberration data was taken on a different day than the alignment data. The reference curves for each day were taken immediately before taking

misalignment or aberrations data. This is why the reference curve for the aberration datasets is saturated in places, the laser power was slightly higher that day. Neither coma, astigmatism, nor trefoil are that interesting; the intensity is impacted most significantly at the downstream side of the focus where the variation in phase is the largest. The most difficult part of tuning the system is distinguishing between astigmatism and coma, which have very similar effects on the beam.

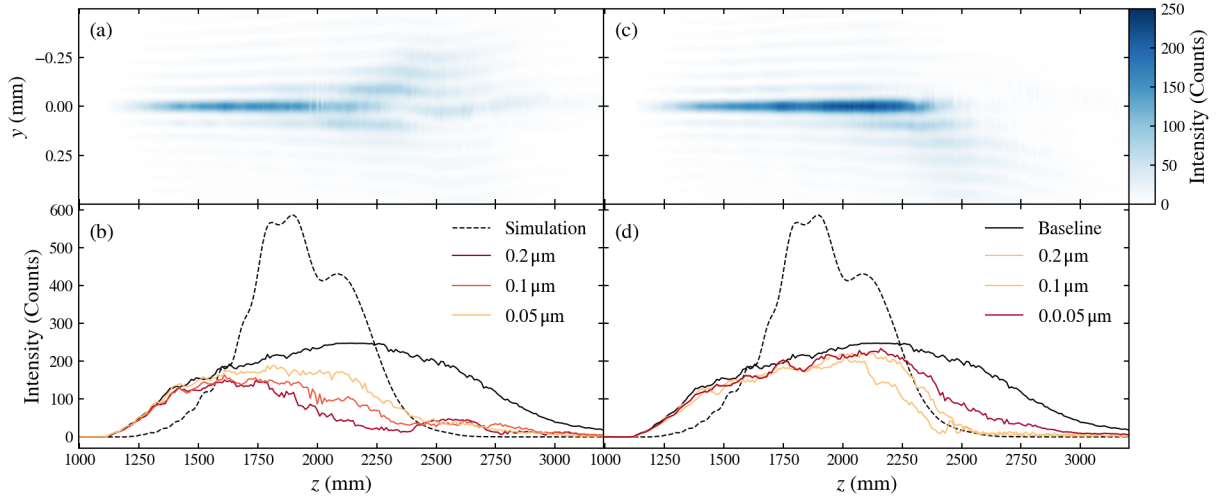


Figure 6.9: Laser intensity when the incoming laser has nonzero astigmatism or trefoil. (a) intensity of the beam with P-V astigmatism of $0.1 \mu\text{m}$. (b) on axis intensity for different P-V amplitudes of astigmatism. (c) intensity with P-V trefoil of $0.2 \mu\text{m}$ of. (d) on axis intensity for different amounts of trefoil.

Defocus is the most interesting aberration because it tends to length the focal depth and reduce the intensity if the amplitude is negative. It does the opposite if the amplitude is positive. This makes it a potentially useful tool for actively adjusting the length of the plasma in a PWFA experiment. Normally changing the plasma length would require changing the optic, but instead a simple adjustment using a deformable mirror can be used to scan through a range of plasma lengths. From Fig. 6.8, we can see that the measured intensity profile begins to approach the profile expected from simulation when the defocus is negative. We will discuss this in the next section.

6.3.3 Laser Collimation

The longitudinal intensity profiles measured throughout this chapter are not consistent with the expectations from simulation. Several different features in the data point towards a diverging beam as the cause: the wider annulus seen from lens A, the variation in the width of the Bessel in z , and the improvement in profile when adding negative defocus. If we take a measurement of the beam propagation without any lens installed, we can see the divergence in the beam directly. To check if this could be the source of the discrepancy between our measurements and simulations, we ran simulations with additional defocusing phase added to the laser. The best match between the simulation and the data is shown in Fig. 6.10. The remaining discrepancy could easily be a combination of misalignment, particularly lens spacing error, and other aberrations in the beam. Unfortunately, we were not able to confirm this hypothesis by adjusting the laser collimation during these tests.

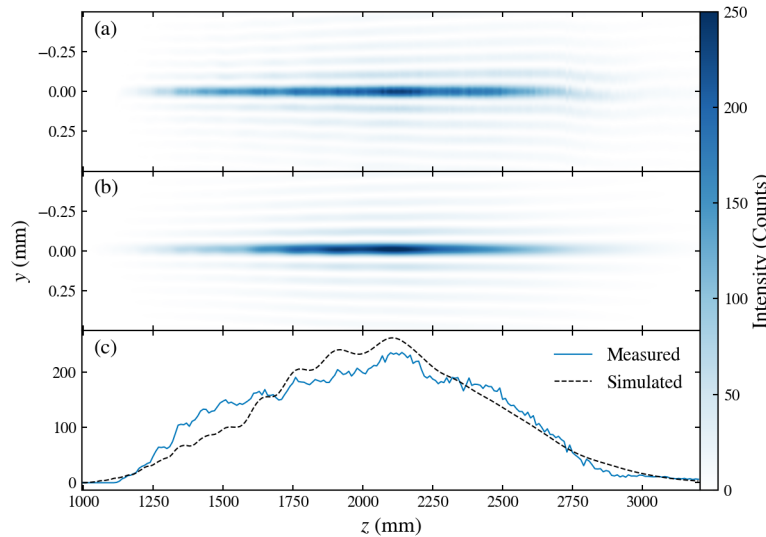


Figure 6.10: Poor collimation in the incoming beam likely explains the discrepancy between experiment and simulation. (a) shows the measured intensity produced by the tandem lens on the $y - z$ plane. (b) shows a simulation of the tandem lens where the incoming beam has $2.5\ \mu\text{m}$ of P-V defocus added. (c) shows the agreement between the on-axis intensity of the simulation and the data.

6.3.4 Frequency Effects

In Sec. 4.5.2 we noted that one of the disadvantage of diffractive optics is the highly chromatic nature of the focusing. The focal length of a diffractive optic is proportional to λ_0/λ , where λ_0 is the design wavelength [113]. Short pulse laser must have large bandwidths, leading to variation in focal position with frequency and a corresponding increase in pulse length. The tandem lenses are relatively immune to this effect because the length of the focus is so long. Our lens has a focus that extends from about 1.5 m to about 2.5 m. A laser with a 5% bandwidth would focus the low frequency components from 1.46 m to 2.44 m while the high frequency components would be focused from 1.54 m to 2.56 m, so only the first and last couple percent of the focus would have an increased pulse length. To demonstrate this, we measured the laser intensity through 10 nm wide band-pass filters centered at 770 nm and 790 nm. The intensity traces are shown in Fig. 6.11. As expected, the focus of the higher frequency components extends farther at the end of the focus than the low frequency components, but only the very ends of the focal region are impacted.

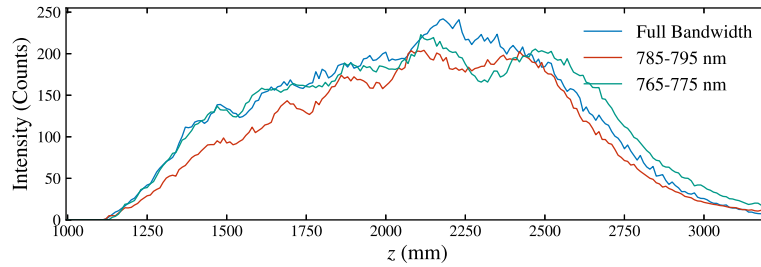


Figure 6.11: On axis intensity produced by the tandem lens measured through different band pass filters. The long wavelength components (red) have shorter focal lengths than the short wavelength components (green), the effect is most noticeable at the end of the focus. The overlap is good enough that there is only pulse lengthening at the very front and rear of the focus.

Chapter 7

Conclusion

In this dissertation, we have made progress toward achieving the goal of emittance preservation in a plasma wakefield accelerator. We have worked out a detailed theory for the beam dynamics that causes emittance growth within a plasma wakefield accelerator. As part of this theory, we identified three regimes of plasma ramps: adiabatic, non-adiabatic, and perturbative. We derived analytic solutions for the beam evolution in both the adiabatic and perturbative regimes, assuming the beam energy is constant in the ramp. These analytic theories together are sufficient to calculate the necessary length of the ramp, the ideal location of the beam's vacuum waist, and the emittance growth within the ramp. Most importantly, we showed that many different ramp shapes can match the beam into the bulk plasma source. The profile of the ramp is not so important for matching, as long as the length is correct for the given ramp shape. However, the length of the ramp can play a role in determining how much energy spread and energy loss is imparted on the beam while traversing the ramp. We showed that it is the integral of the plasma density along the ramp that determines the deleterious perturbation to the witness beam.

We developed a full analytic framework that includes the effects of energy change, wake loading, and adiabatic variations in plasma density. Using this framework, we analytically calculated the evolution of the emittance when the witness beam is mismatched and/or transversely offset from the center of the plasma wake. This allowed us to derive simple closed form saturation lengths for the projected and slice emittances for cases where the witness beam has either an initial mismatch, an initial offset, or both. These analytic formulas are a useful tool that will allow experimentalists

to quickly estimate the tolerances on beam parameters required to achieve emittance preservation. Such numerical explorations of the possible parameter phase space are otherwise unachievable given the time and expense of PIC simulations. We also considered the detailed evolution of the longitudinal and energy slices of the beam for the first time. In particular, we showed that the energy slices undergo irreversible emittance growth within the plasma due to energy slice mixing as a result of non-uniform wake loading. Of particular importance for experiments, we discovered that, in the case of mismatch or offset, a correlation between Courant-Snyder parameters and energy will develop in the beam. The standard downstream emittance diagnostics, particularly the butterfly technique, rely on the assumption that all the energy components have the same parameters at the exit of the plasma. The correlations will thus impact the numbers retrieved from these diagnostics and have implications for how the diagnostic techniques must be altered and interpreted to account for this effect.

The theoretical work provides guidance for designing a plasma ramp—minimize the integrated density and keep the shape as adiabatic as possible. Defining the full plasma source requires determining the density, length, and width of the uniform plasma where acceleration occurs. Optimizing the bulk density of the plasma source is the first step in the design process. Using PIC simulations, we found the optimal combination of beam parameters and plasma density that maximized the transformer ratio and minimized the final energy spread within the constraints of FACET-II. The length of the plasma follows straightforwardly from the average decelerating field experienced by the drive beam in the simulations. For the width, we presented a simple argument for how wide the plasma must be to avoid impacting the nonlinear wake dynamics and verified it with PIC simulations.

With requirements for the plasma source in hand, we developed an optical technique for producing suitable plasmas. Our technique combines the ideas of on-axis intensity shaping, tandem lenses, and binary diffractive optics to shape a high-power laser into a controlled meter scale focus. We then implemented a version of the split step algorithm that combined Fourier propagation with laser ionization. Using this algorithm, we calculated how much the index of refraction from the

plasma formed by the front of the pulse impacts the propagation of the rear of the pulse. Using this information, we were able to define the optimal width of the focus to create the widest possible plasma using our technique. The optical setup was demonstrated experimentally to show that the optics work as intended. We also measured the alignment sensitivity and tolerable wavefront error of the optics, allowing us to set requirements on the quality of the incoming laser.

7.1 Future Work

The work on preserving emittance presented in this dissertation is far from complete. The goal is to demonstrate emittance preservation experimentally at FACET-II. There are several smaller steps required to reach this goal. The next step in plasma source development is to ionize a plasma with the tandem optics and verify that the plasma density matches expectation. Measuring the plasma source is a challenge in and of itself. At densities of $3.5 \times 10^{16} \text{ cm}^{-3}$, the index of refraction is small. Orthogonally directed, low energy laser probes will not experience a sufficiently large phase delay from propagation across the plasma filament, rendering techniques such as shadowgraphic phase reconstruction impossible in this configuration. Instead, a probe laser must be sent through the plasma at a shallow oblique angle such that it passes through a sufficiently long segment of plasma to pick up a detectable phase shift that can be observed as an intensity modulation further downstream.

After verifying the ionization optics produce the expected plasma, the optics need to be installed, aligned, and characterized in the FACET-II IP area. Initial experiments with the electron beam should focus on demonstrating a reduction in head erosion using only the drive beam. Once this has been achieved, and there is confidence that the plasma source is working as intended, two bunch beams should be introduced. It will be necessary to spoil the emittance of the beams, that is, render it artificially larger than what is delivered out of the linac, in order to prevent beam ionization of the helium buffer gas on either end of the lithium oven. Proof of principle experiments with spoiled emittance will provide enough confidence in the technique to move onto a plasma source with a shaped gas density, possibly an elongated gas jet, to demonstrate emittance

preservation at the micron-radian level.

If emittance preservation can be demonstrated, then all the pieces for a functional plasma wakefield accelerator stage will have been achieved experimentally. It is then a matter of achieving all of them at once. Once a plasma accelerator produces beams with low energy spread, high efficiency, large energy gain (full depletion of the drive beam), and emittance preservation then a PWFA could be considered ready for use in applications such as light sources and colliders.

Bibliography

- [1] J. W. Wang and G. A. Loew. Field Emission and RF Breakdown in High-Gradient Room-Temperature Linac Structures. Technical Report SLAC-PUB-7684, Stanford Linear Accelerator Center, Stanford Univ., CA (United States), 1997.
- [2] J.N. Galayda. The LCLS-II: A High Power Upgrade to the LCLS. In *Proceedings of the 9th International Particle Accelerator Conference (IPAC'18)*, volume 9, pages 18–23, Geneva, Switzerland, 2018. JACoW Publishing.
- [3] T. Tajima and J. M. Dawson. Laser Electron Accelerator. *Physical Review Letters*, 43(4):267, 1979.
- [4] R. D. Ruth, A. W. Chao, P. L. Morton, and P. B. Wilson. A Plasma Wake Field Accelerator. *Particle Accelerators*, 17(6):171, 1985.
- [5] P. Chen, J. M. Dawson, R. W. Huff, and T. Katsouleas. Acceleration of Electrons by the Interaction of a Bunched Electron Beam with a Plasma. *Physical Review Letters*, 54(7):693, 1985.
- [6] I. Blumenfeld, C. E. Clayton, F.-J. Decker, M. J. Hogan, C. Huang, R. Ischebeck, R. Iverson, C. Joshi, T. Katsouleas, N. Kirby, W. Lu, K. A. Marsh, W. B. Mori, P. Muggli, E. Oz, R. H. Siemann, D. Walz, and M. Zhou. Energy doubling of 42 GeV electrons in a metre-scale plasma wakefield accelerator. *Nature*, 445(7129):741, 2007.
- [7] M. J. Hogan, T. O. Raubenheimer, A. Seryi, P. Muggli, T. Katsouleas, C. Huang, W. Lu, W. An, K. A. Marsh, W. B. Mori, C. E. Clayton, and Joshi C. Plasma wakefield acceleration experiments at FACET. *New Journal of Physics*, 12(5):055030, 2010.
- [8] M. Litos, E. Adli, W. An, C. I. Clarke, C. E. Clayton, S. Corde, J. P. Delahaye, R. J. England, A. S. Fisher, J. Frederico, S. Gessner, S. Z. Green, M. J. Hogan, C. Joshi, W. Lu, K. A. Marsh, W. B. Mori, P. Muggli, N. Vafaei-Najafabadi, D. Walz, G. White, Z. Wu, V. Yakimenko, and G. Yocky. High-efficiency acceleration of an electron beam in a plasma wakefield accelerator. *Nature*, 515(7525):92, 2014.
- [9] M. Litos, E. Adli, J. M. Allen, W. An, C. I. Clarke, S. Corde, C. E. Clayton, J. Frederico, S. J. Gessner, S. Z. Green, M. J. Hogan, C. Joshi, W. Lu, K. A. Marsh, W. B. Mori, M. Schmeltz, N. Vafaei-Najafabadi, and V. Yakimenko. 9 GeV energy gain in a beam-driven plasma wakefield accelerator. *Plasma Physics and Controlled Fusion*, 58(3):034017, 2016.

- [10] W. P. Leemans, B. Nagler, A. J. Gonsalves, Cs. Tóth, K. Nakamura, C. G. R. Geddes, E. Esarey, C. B. Schroeder, and S. M. Hooker. GeV electron beams from a centimetre-scale accelerator. *Nature Physics*, 2(10):696, 2006.
- [11] W. P. Leemans, A. J. Gonsalves, H.-S. Mao, K. Nakamura, C. Benedetti, C. B. Schroeder, Cs. Tóth, J. Daniels, D. E. Mittelberger, S. S. Bulanov, J.-L. Vay, C. G. R. Geddes, and E. Esarey. Multi-GeV Electron Beams from Capillary-Discharge-Guided Subpetawatt Laser Pulses in the Self-Trapping Regime. *Physical Review Letters*, 113(24):245002, 2014.
- [12] N. Barov and J. B. Rosenzweig. Propagation of short electron pulses in underdense plasmas. *Physical Review E*, 49(5):4407, 1994.
- [13] X. L. Xu, J. F. Hua, F. Li, C. J. Zhang, L. X. Yan, Y. C. Du, W. H. Huang, H. B. Chen, C. X. Tang, W. Lu, P. Yu, W. An, C. Joshi, and W. B. Mori. Phase-Space Dynamics of Ionization Injection in Plasma-Based Accelerators. *Physical Review Letters*, 112(3):035003, 2014.
- [14] P. Michel, C. B. Schroeder, B. A. Shadwick, E. Esarey, and W. P. Leemans. Radiative damping and electron beam dynamics in plasma-based accelerators. *Physical Review E*, 74(2):026501, 2006.
- [15] T. Mehrling, J. Grebenyuk, F. S. Tsung, K. Floettmann, and J. Osterhoff. Transverse emittance growth in staged laser-wakefield acceleration. *Physical Review Special Topics - Accelerators and Beams*, 15(11):111303, 2012.
- [16] K. Floettmann. Adiabatic matching section for plasma accelerated beams. *Physical Review Special Topics - Accelerators and Beams*, 17(5):054402, 2014.
- [17] I. Dornmair, K. Floettmann, and A. R. Maier. Emittance conservation by tailored focusing profiles in a plasma accelerator. *Physical Review Special Topics - Accelerators and Beams*, 18(4):041302, 2015.
- [18] X. L. Xu, J. F. Hua, Y. P. Wu, C. J. Zhang, F. Li, Y. Wan, C.-H. Pai, W. Lu, W. An, P. Yu, M. J. Hogan, C. Joshi, and W. B. Mori. Physics of Phase Space Matching for Staging Plasma and Traditional Accelerator Components Using Longitudinally Tailored Plasma Profiles. *Physical Review Letters*, 116(12):124801, 2016.
- [19] A. Aschikhin, T. J. Mehrling, A. Martinez de la Ossa, and J. Osterhoff. Analytical model for the uncorrelated emittance evolution of externally injected beams in plasma-based accelerators. *Nuclear Instruments and Methods in Physics Research, Section A*, 909:414, 2018.
- [20] Y. Zhao, W. An, X. Xu, F. Li, L. Hildebrand, M. J. Hogan, V. Yakimenko, C. Joshi, and W. B. Mori. Emittance preservation through density ramp matching sections in a plasma wakefield accelerator. *Physical Review Accelerators and Beams*, 23(1):011302, 2020.
- [21] R. Assmann and K. Yokoya. Transverse beam dynamics in plasma-based linacs. *Nuclear Instruments and Methods in Physics Research, Section A*, 410(3):544, 1998.
- [22] C. A. Lindstrøm, E. Adli, J. Pflugstner, E. Marin, and D. Schulte. Transverse Tolerances of a Multi-Stage Plasma Wakefield Accelerator. In *Proceedings of the 7th International Particle Accelerator Conference (IPAC'16)*, volume 7, pages 2561–2564, Geneva, Switzerland, 2016. JACoW Publishing.

- [23] M. Thévenet, R. Lehe, C. B. Schroeder, C. Benedetti, J.-L. Vay, E. Esarey, and W. P. Leemans. Emittance growth due to misalignment in multistage laser-plasma accelerators. *Physical Review Accelerators and Beams*, 22(5):051302, 2019.
- [24] T. O. Raubenheimer. Estimates of emittance dilution and stability in high-energy linear accelerators. *Physical Review Special Topics - Accelerators and Beams*, 3(12):121002, 2000.
- [25] W. Baumjohann and R. A. Treumann. *Basic Space Plasma Physics*. Imperial College Press, London, 1997.
- [26] F. F. Chen. *Introduction to Plasma Physics and Controlled Fusion: Volume 1: Plasma Physics*. Springer US, 2nd edition, 2013.
- [27] P. M. Bellan. *Fundamentals of Plasma Physics*. Cambridge University Press, 2008.
- [28] R. Keinigs and M. E. Jones. Two-dimensional dynamics of the plasma wakefield accelerator. *Physics of Fluids*, 30(1):252, 1987.
- [29] I. Blumenfeld. *Scaling of the Longitudinal Electric Fields and Transformer Ratio in a Non-Linear Plasma Wakefield Accelerator*. PhD thesis, Stanford University, 2009.
- [30] S. J. Gessner. *Demonstration of the Hollow Channel Plasma Wakefield Accelerator*. PhD thesis, Stanford University, 2016.
- [31] C. A. Lindstrøm. *Emittance growth and preservation in a plasma-based linear collider*. PhD thesis, University of Oslo, 2019.
- [32] J. D. Jackson. *Classical Electrodynamics*. Wiley, Hoboken, 3rd edition, 1998.
- [33] M. J. Hogan, R. Assmann, F.-J. Decker, R. Iverson, P. Raimondi, S. Rokni, R. H. Siemann, D. Walz, D. Whittum, B. Blue, C. E. Clayton, E. Dodd, R. Hemker, C. Joshi, K. A. Marsh, W. B. Mori, S. Wang, T. Katsouleas, S. Lee, P. Muggli, P. Catravas, S. Chattopadhyay, E. Esarey, and W. P. Leemans. E-157: A 1.4-m-long plasma wake field acceleration experiment using a 30 GeV electron beam from the Stanford Linear Accelerator Center Linac. *Physics of Plasmas*, 7(5):2241, 2000.
- [34] M. J. Hogan, C. D. Barnes, C. E. Clayton, F. J. Decker, S. Deng, P. Emma, C. Huang, R. H. Iverson, D. K. Johnson, C. Joshi, T. Katsouleas, P. Krejcik, W. Lu, K. A. Marsh, W. B. Mori, P. Muggli, C. L. O’Connell, E. Oz, R. H. Siemann, and D. Walz. Multi-GeV Energy Gain in a Plasma-Wakefield Accelerator. *Physical Review Letters*, 95(5):054802, 2005.
- [35] T. Katsouleas, S. Wilks, P. Chen, J. M. Dawson, and J. J. Su. Beam Loading in Plasma Accelerators. *Particle Accelerators*, 22:81, 1987.
- [36] K. L. F. Bane, P. Chen, and P. B. Wilson. On collinear wake field acceleration. Technical Report SLAC-PUB-3662, Stanford Linear Accelerator Center, Stanford Univ., CA (United States), 1985.
- [37] T. Katsouleas. Physical mechanisms in the plasma wake-field accelerator. *Physical Review A*, 33(3):2056, 1986.
- [38] P. Chen, J. J. Su, J. M. Dawson, K. L. F. Bane, and P. B. Wilson. Energy Transfer in the Plasma Wake-Field Accelerator. *Physical Review Letters*, 56(12):1252, 1986.

- [39] K. L. F Bane, P. B. Wilson, and T. Weiland. Wake fields and wake field acceleration. In *AIP Conference Proceedings*, volume 127, pages 875–928, New York, 1985. AIP.
- [40] J. Albritton and P. Koch. Cold plasma wavebreaking: Production of energetic electrons. *Physics of Fluids*, 18(9):1136, 1975.
- [41] W. Lu, C. Huang, M. Zhou, W. B. Mori, and T. Katsouleas. Nonlinear Theory for Relativistic Plasma Wakefields in the Blowout Regime. *Physical Review Letters*, 96(16):165002, 2006.
- [42] W. Lu, C. Huang, M. Zhou, M. Tzoufras, F. S. Tsung, W. B. Mori, and T. Katsouleas. A nonlinear theory for multidimensional relativistic plasma wave wakefields. *Physics of Plasmas*, 13(5):056709, 2006.
- [43] A. A. Golovanov, I. Yu. Kostyukov, L. Reichwein, J. Thomas, and A. Pukhov. Excitation of strongly nonlinear plasma wakefield by electron bunches. *Plasma Physics and Controlled Fusion*, 63(8):085004, 2021.
- [44] J. M. Dawson. Nonlinear Electron Oscillations in a Cold Plasma. *Physical Review*, 113(2):383, 1959.
- [45] S. A. Yi, V. Khudik, C. Siemon, and G. Shvets. Analytic model of electromagnetic fields around a plasma bubble in the blow-out regime. *Physics of Plasmas*, 20(1):013108, 2013.
- [46] T. N. Dalichaouch, X. L. Xu, A. Tableman, F. Li, F. S. Tsung, and W. B. Mori. A multi-sheath model for highly nonlinear plasma wakefields. *Physics of Plasmas*, 28(6):063103, 2021.
- [47] J. Thomas, I. Yu. Kostyukov, J. Pronold, A. Golovanov, and A. Pukhov. Non-linear theory of a cavitated plasma wake in a plasma channel for special applications and control. *Physics of Plasmas*, 23(5):053108, 2016.
- [48] A. A. Golovanov, I. Yu. Kostyukov, A M Pukhov, and J Thomas. Generalised model of a sheath of a plasma bubble excited by a short laser pulse or by a relativistic electron bunch in transversely inhomogeneous plasma. *Quantum Electronics*, 46(4):295, 2016.
- [49] M. Tzoufras, W. Lu, F. S. Tsung, C. Huang, W. B. Mori, T. Katsouleas, J. Vieira, R. A. Fonseca, and L. O. Silva. Beam Loading in the Nonlinear Regime of Plasma-Based Acceleration. *Physical Review Letters*, 101(14):145002, 2008.
- [50] S.Y. Lee. *Accelerator Physics*. World Scientific Publishing Co. Pte. Ltd., Hackensack, 3rd edition, 2012.
- [51] E. D. Courant and H. S. Snyder. Theory of the alternating-gradient synchrotron. *Annals of Physics*, 3(1):1, 1958.
- [52] K. Floettmann. Some basic features of the beam emittance. *Physical Review Special Topics - Accelerators and Beams*, 6(3):034202, 2003.
- [53] J. Buon. Beam phase space and emittance. Technical Report LAL-RT-90-15-REV, CERN, Geneva (Switzerland), 1992.
- [54] R. L. Mills and A. M. Sessler. Liouville’s theorem and phase-space cooling. Technical Report LBL-34667, Lawrence Berkeley Lab., CA (United States), 1993.

- [55] A. Sampath. *Strong-field QED and collisional effects in electron beam-plasma interaction*. PhD thesis, Ruperto-Carola-University of Heidelberg, 2020.
- [56] C. Nieter and J. R. Cary. VORPAL: a versatile plasma simulation code. *Journal of Computational Physics*, 196(2):448, 2004.
- [57] R. Ariniello, C. E. Doss, K. Hunt-Stone, J. R. Cary, and M. D. Litos. Transverse beam dynamics in a plasma density ramp. *Physical Review Accelerators and Beams*, 22(4):041304, 2019.
- [58] R. Ariniello, C. E. Doss, V. Lee, C. Hansel, J. R. Cary, and M. D. Litos. Chromatic dynamics of an electron beam in a plasma based accelerator. *Physical Review Research*, 2022 submitted, arXiv:2111.02332.
- [59] P. Chen, K. Oide, A. M. Sessler, and S. S. Yu. Plasma-based adiabatic focuser. *Physical Review Letters*, 64(11):1231, 1990.
- [60] R. L. Williams and T. Katsouleas. Numerical studies on the ramped density plasma lens. In *AIP Conference Proceedings*, volume 279, pages 565–578, New York, 1992. AIP.
- [61] C. A. Lindstrøm and E. Adli. Design of general apochromatic drift-quadrupole beam lines. *Physical Review Accelerators and Beams*, 19(7):071002, 2016.
- [62] C. Joshi, E. Adli, W. An, C. E. Clayton, S. Corde, S. Gessner, M. J. Hogan, M. Litos, W. Lu, K. A. Marsh, W. B. Mori, N. Vafaei-Najafabadi, B. O’shea, X. Xu, G. White, and V. Yakimenko. Plasma wakefield acceleration experiments at FACET II. *Plasma Physics and Controlled Fusion*, 60(3):034001, 2018.
- [63] V. Yakimenko, L. Alsberg, E. Bong, G. Bouchard, C. Clarke, C. Emma, S. Green, C. Hast, M. J. Hogan, J. Seabury, N. Lipkowitz, B. O’Shea, D. Storey, G. White, and G. Yocky. FACET-II facility for advanced accelerator experimental tests. *Physical Review Accelerators and Beams*, 22(10):101301, 2019.
- [64] A. Aschikhin, C. Behrens, S. Bohlen, J. Dale, N. Delbos, L. di Lucchio, E. Elsen, J.-H. Erbe, M. Felber, B. Foster, L. Goldberg, J. Grebenyuk, J.-N. Gruse, B. Hidding, Z. Hu, S. Karstensen, A. Knetsch, O. Kononenko, V. Libov, K. Ludwig, A. R. Maier, A. Martinez de la Ossa, T. Mehrling, C. A. J. Palmer, F. Pannek, L. Schaper, H. Schlarb, B. Schmidt, S. Schreiber, J.-P. Schwinkendorf, H. Steel, M. Streeter, G. Tauscher, V. Wacker, S. Weichert, S. Wunderlich, J. Zemella, and J. Osterhoff. The FLASHForward facility at DESY. *Nuclear Instruments and Methods in Physics Research, Section A*, 806:175, 2016.
- [65] C. A. Lindstrøm, J. M. Garland, S. Schröder, L. Boulton, G. Boyle, J. Chappell, R. D’Arcy, P. Gonzalez, A. Knetsch, V. Libov, G. Loisch, A. Martinez de la Ossa, P. Niknejadi, K. Pöder, L. Schaper, B. Schmidt, B. Sheeran, S. Wesch, J. Wood, and J. Osterhoff. Energy-Spread Preservation and High Efficiency in a Plasma-Wakefield Accelerator. *Physical Review Letters*, 126(1):014801, 2021.
- [66] Z. Huang, Y. Ding, and C. B. Schroeder. Compact X-ray Free-Electron Laser from a Laser-Plasma Accelerator Using a Transverse-Gradient Undulator. *Physical Review Letters*, 109(20):204801, 2012.

- [67] T. I. Smith, J. M. J. Madey, L. R. Elias, and D. A. G. Deacon. Reducing the sensitivity of a free-electron laser to electron energy energy. *Journal of Applied Physics*, 50(7):4580, 1979.
- [68] P. Baxevanis, Y. Ding, Z. Huang, and R. Ruth. 3D theory of a high-gain free-electron laser based on a transverse gradient undulator. *Physical Review Special Topics - Accelerators and Beams*, 17(2):020701, 2014.
- [69] M. D. Litos, R. Ariniello, C. E. Doss, K. Hunt-Stone, and J. R. Cary. Beam emittance preservation using Gaussian density ramps in a beam-driven plasma wakefield accelerator. *Philosophical Transactions of the Royal Society A*, 377(2151):20180181, 2019.
- [70] M. Zhou, C. E. Clayton, C. Huang, C. Joshi, W. Lu, K. A. Marsh, W. B. Mori, T. Katsouleas, P. Muggli, E. Oz, M. Berry, I. Blumenfeld, F. J. Decker, M. J. Hogan, R. Ischebeck, R. Iverson, N. Kirby, R. Siemann, and D. Walz. Beam head erosion in self-ionized plasma wakefield accelerators. In *Proceedings of the 2007 IEEE Particle Accelerator Conference (PAC)*, pages 3064–3066. IEEE, 2007.
- [71] W. An, M. Zhou, N. Vafaei-Najafabadi, K. A. Marsh, C. E. Clayton, C. Joshi, W. B. Mori, W. Lu, E. Adli, S. Corde, M. Litos, S. Li, S. Gessner, J. Frederico, M. J. Hogan, D. Walz, J. England, J. P. Delahaye, and P. Muggli. Strategies for mitigating the ionization-induced beam head erosion problem in an electron-beam-driven plasma wakefield accelerator. *Physical Review Special Topics - Accelerators and Beams*, 16(10):101301, 2013.
- [72] D. H. Whittum, W. M. Sharp, S. S. Yu, M. Lampe, and G. Joyce. Electron-hose instability in the ion-focused regime. *Physical Review Letters*, 67(8):991, 1991.
- [73] M. Lampe, G. Joyce, S. P. Slinker, and D. H. Whittum. Electron-hose instability of a relativistic electron beam in an ion-focusing channel. *Physics of Fluids B*, 5(6):1888, 1993.
- [74] A. A. Geraci and D. H. Whittum. Transverse dynamics of a relativistic electron beam in an underdense plasma channel. *Physics of Plasmas*, 7(8):3431, 2000.
- [75] S. Deng, C. D. Barnes, C. E. Clayton, C. O’Connell, F. J. Decker, R. A. Fonseca, C. Huang, M. J. Hogan, R. Iverson, D. K. Johnson, C. Joshi, T. Katsouleas, P. Krejcik, W. Lu, W. B. Mori, P. Muggli, E. Oz, F. Tsung, D. Walz, and M. Zhou. Hose Instability and Wake Generation by an Intense Electron Beam in a Self-Ionized Gas. *Physical Review Letters*, 96(4):045001, 2006.
- [76] C. Huang, W. Lu, M. Zhou, C. E. Clayton, C. Joshi, W. B. Mori, P. Muggli, S. Deng, E. Oz, T. Katsouleas, M. J. Hogan, I. Blumenfeld, F. J. Decker, R. Ischebeck, R. H. Iverson, N. A. Kirby, and D. Walz. Hosing Instability in the Blow-Out Regime for Plasma-Wakefield Acceleration. *Physical Review Letters*, 99(25):255001, 2007.
- [77] T. J. Mehrling, C. Benedetti, C. B. Schroeder, A. Martinez de la Ossa, J. Osterhoff, E. Esarey, and W. P. Leemans. Accurate modeling of the hose instability in plasma wakefield accelerators. *Physics of Plasmas*, 25(5):056703, 2018.
- [78] T. J. Mehrling, R. A. Fonseca, A. Martinez de la Ossa, and J. Vieira. Mitigation of the Hose Instability in Plasma-Wakefield Accelerators. *Physical Review Letters*, 118(17):174801, 2017.

- [79] E. Adli, S. Corde, R. J. England, J. Frederico, S. J. Gessner, S. Li, Z. Wu, M. J. Hogan, M. D. Litos, W. An, C.E. Clayton, K. A. Marsh, W. Lu, W. Mori, C. Joshi, N. Vafaei-Najafabadi, and P. Muggli. First results from the electron hose instability studies in facet. In *Proceedings of the 3rd International Particle Accelerator Conference (IPAC'12)*, volume 3, pages 43–45. IEEE, 2012.
- [80] T. J. Mehrling, C. Benedetti, C. B. Schroeder, E. Esarey, and W. P. Leemans. Suppression of Beam Hosing in Plasma Accelerators with Ion Motion. *Physical Review Letters*, 121(26):264802, 2018.
- [81] S. Lee and T. Katsouleas. Wakefield accelerators in the blowout regime with mobile ions. In *AIP Conference Proceedings*, volume 472, pages 524–533, New York, 1999. AIP.
- [82] J. B. Rosenzweig, A. M. Cook, A. Scott, M. C. Thompson, and R. B. Yoder. Effects of Ion Motion in Intense Beam-Driven Plasma Wakefield Accelerators. *Physical Review Letters*, 95(19):195002, 2005.
- [83] R. Gholizadeh, T. Katsouleas, P. Muggli, and W. Mori. Analysis of Ion Motion and Scattering in the Extreme Regime of High Intensity Electron Beams in Plasma Wakefield Accelerators. In *AIP Conference Proceedings*, volume 877, pages 504–510, New York, 2006. AIP.
- [84] C. Benedetti, C. B. Schroeder, E. Esarey, and W. P. Leemans. Emittance preservation in plasma-based accelerators with ion motion. *Physical Review Accelerators and Beams*, 20(11):111301, 2017.
- [85] W. An, W. Lu, C. Huang, X. Xu, M. J. Hogan, C. Joshi, and W. B. Mori. Ion Motion Induced Emittance Growth of Matched Electron Beams in Plasma Wakefields. *Physical Review Letters*, 118(24):244801, 2017.
- [86] R. Gholizadeh, T. Katsouleas, P. Muggli, C. Huang, and W. Mori. Preservation of beam emittance in the presence of ion motion in future high-energy plasma-wakefield-based colliders. *Physical Review Letters*, 104(15):155001, 2010.
- [87] N. B. Delone and V. P. Krainov. *Multiphoton Processes in Atoms*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1994.
- [88] C. I. Blaga. *Atoms and Molecules in Strong Midinfrared Laser Fields*. PhD thesis, Stony Brook University, 2009.
- [89] L. V. Keldysh. Ionization in the field of a strong electromagnetic wave. *Soviet Physics JETP*, 20(5):1307, 1965.
- [90] F. A. Ilkov, J. E. Decker, and S. L. Chin. Ionization of atoms in the tunnelling regime with experimental evidence using Hg atoms. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 25(19):4005, 1992.
- [91] M. V. Ammosov, N. B. Delone, and V. P. Krainov. Tunnel ionization of complex atoms and of atomic ions in an alternating electromagnetic field. *Soviet Physics JETP*, 64(6):1191, 1986.
- [92] B. M. Smirnov and M. I. Chibisov. The Breaking Up of Atomic Particles by an Electric Field and by Electron Collisions. *Soviet Physics JETP*, 22(3):585, 1966.

- [93] A. M. Perelomov, V. S. Popov, and M. V. Terent'ev. Ionization of atoms in an alternating electrical field. *Soviet Physics JETP*, 23(5):924, 1966.
- [94] V. S. Popov. On the history of developing the theory of tunneling ionization in atoms and ions. *Physics-Uspekhi*, 42(7):733, 1999.
- [95] A. Couairon and A. Mysyrowicz. Femtosecond filamentation in transparent media. *Physics Reports*, 441(2-4):47, 2007.
- [96] D. L. Bruhwiler, D. A. Dimitrov, J. R. Gary, E. Esarey, W. Leemans, and R. E. Giacone. Particle-in-cell simulations of tunneling ionization effects in plasma-based accelerators. *Physics of Plasmas*, 10(5):2022, 2003.
- [97] A. M. Perelomov and Vladimir S. Popov. Ionization of atoms in an alternating electric field. III. *Soviet Physics JETP*, 25(2):336, 1967.
- [98] V. S. Popov. Tunnel and multiphoton ionization of atoms and ions in a strong laser field (Keldysh theory). *Physics-Uspekhi*, 47(9):855, 2004.
- [99] S. V. Popruzhenko, V. D. Mur, V. S. Popov, and D. Bauer. Strong Field Ionization Rate for Arbitrary Laser Frequencies. *Physical Review Letters*, 101(19):193003, 2008.
- [100] Y. H. Lai, J. Xu, U. B. Szafruga, B. K. Talbert, X. Gong, K. Zhang, H. Fuest, M. F. Kling, C. I. Blaga, P. Agostini, and L. F. DiMauro. Experimental investigation of strong-field-ionization theories for laser fields from visible to midinfrared frequencies. *Physical Review A*, 96(6):063417, 2017.
- [101] Joel Venzke. personal communication.
- [102] S. Hooker and C. Webb. *Laser Physics*. Oxford Master Series in Physics. OUP Oxford, 2010.
- [103] J. W. Goodman. *Introduction to Fourier Optics*. Roberts & Comapny Publishers, Greenwood Village, 3rd edition, 2005.
- [104] J. A. Ratcliffe. Some aspects of diffraction theory and their application to the ionosphere. *Reports on Progress in Physics*, 19(1):188, 1956.
- [105] R. W. Boyd. The Nonlinear Optical Susceptibility. In *Nonlinear Optics*. Academic Press, Boston, 2008.
- [106] C. E. Max, J. Arons, and A. B. Langdon. Self-modulation and self-focusing of electromagnetic waves in plasmas. *Physical Review Letters*, 33(4):209, 1974.
- [107] P. Sprangle, C.-M. Tang, and E. Esarey. Relativistic self-focusing of short-pulse radiation beams in plasmas. *IEEE Transactions on Plasma Science*, 15(2):145, 1987.
- [108] T. Cizmar and K. Dholakia. Tunable Bessel light modes: engineering the axial propagation. *Optics Express*, 17(18):15558, 2009.
- [109] I. Ouadghiri-Idrissi, R. Giust, L. Froehly, M. Jacquot, L. Furfaro, J. M. Dudley, and F. Courvoisier. Arbitrary shaping of on-axis amplitude of femtosecond Bessel beams with a single phase-only spatial light modulator. *Optics Express*, 24(11):11495, 2016.

- [110] M. Honkanen and J. Turunen. Tandem systems for efficient generation of uniform-axial-intensity Bessel fields. *Optics Communications*, 154(5-6):368, 1998.
- [111] J. Sochacki, A. Kołodziejczyk, Z. Jaroszewicz, and S. Bará. Nonparaxial design of generalized axicons. *Applied Optics*, 31(25):5326, 1992.
- [112] G. J. Swanson. Binary optics technology: theoretical limits on the diffraction efficiency of multilevel diffractive optical elements. Technical report, Massachusetts Institute of Technology, Lexington Lincoln Lab. MA (United States), 1991.
- [113] D. A. Buralli, G. M. Morris, and J. R. Rogers. Optical performance of holographic kinoforms. *Applied Optics*, 28(5):976, 1989.
- [114] M. Turner, A. J. Gonsalves, S. S. Bulanov, C. Benedetti, N. A. Bobrova, V. A. Gasilov, P. V. Sasorov, G. Korn, K. Nakamura, J. Van Tilborg, C. G. Geddes, C. B. Schroeder, and E. Esarey. Radial density profile and stability of capillary discharge plasma waveguides of lengths up to 40 cm. *High Power Laser Science and Engineering*, 9:e17, 2021.
- [115] L. Goldberg. *Spectroscopic Electron Density Determination of Plasma Targets for Plasma Wakefield Acceleration*. PhD thesis, Univeristy of Hamburg, 2018.
- [116] N. Vafaei-Najafabadi, J. L. Shaw, K. A. Marsh, C. Joshi, and M. J. Hogan. Meter scale plasma source for plasma wakefield experiments. In *AIP Conference Proceedings*, volume 1507, pages 650–655, New York, 2012. AIP.
- [117] B. R. Munson, D. F. Young, T. H. Okiishi, and W. W. Huebsch. *Fundamentals of Fluid Mechanics*. Wiley, Hoboken, 6th edition, 2009.
- [118] K. Schmid and L. Veisz. Supersonic gas jets for laser-plasma experiments. *Review of Scientific Instruments*, 83(5):053304, 2012.
- [119] G. P. Sutton and O. Biblarz. *Rocket Propulsion Elements*. John Wiley & Sons, New York, 2001.
- [120] Technical Design Report for the FACET-II Project at SLAC National Accelerator Laboratory. Technical Report SLAC-R-1072, SLAC National Accelerator Laboratory, CA (United States), 2016.
- [121] C. Emma, A. Edelen, M. J. Hogan, B. O’Shea, G. White, and V. Yakimenko. Machine learning-based longitudinal phase space prediction of particle accelerators. *Physical Review Accelerators and Beams*, 21(11):112802, 2018.
- [122] Glen White. FACET-II Start-to-End Simulations of Electron Beam, October 2019. FACET-II Science Workshop 2019.
- [123] Doug Storey. Spectrometer diagnostics, October 2019. FACET-II Science Workshop 2019.
- [124] A. Scheinker and S. Gessner. Adaptive method for electron bunch profile prediction. *Physical Review Special Topics - Accelerators and Beams*, 18(10):102801, 2015.
- [125] K. Hunt-Stone, R. Ariniello, C. Doss, V. Lee, and M. Litos. Electro-optic sampling beam position monitor for relativistic electron beams. *Nuclear Instruments and Methods in Physics Research Section A*, 999:165210, 2021.

- [126] S. Z. Green, E. Adli, C. I. Clarke, S. Corde, S. A. Edstrom, A. S. Fisher, J. Frederico, J. C. Frisch, S. Gessner, S. Gilevich, P. Hering, M. J. Hogan, R. K. Jobe, M. Litos, J. E. May, D. R. Walz, V. Yakimenko, C. E. Clayton, C. Joshi, K. A. Marsh, N. Vafaei-Najafabadi, and P. Muggli. Laser ionized preformed plasma at FACET. *Plasma Physics and Controlled Fusion*, 56(8):084011, 2014.
- [127] H. Goldstein, C. P. Poole, and J. Safko. *Classical Mechanics*. Pearson, 2013.
- [128] K. Mishima, M. Hayashi, J. Yi, S. H. Lin, H. L. Selzle, and E. W. Schlag. Generalization of Keldysh’s theory. *Physical Review A*, 66(3):033401, 2002.

Appendix A

Green's Function for a Linear Plasma Wave

The equation of motion for a fluid element in a linear plasma wave is

$$\frac{\partial^2 \delta n}{\partial \xi^2} + \frac{\omega_p^2}{v_b^2} \delta n = -\frac{\omega_p^2}{v_b^2} n_b. \quad (\text{A.1})$$

We can use a Green's function approach to solve the problem for an arbitrary beam density at a given x, y position. The Green's function for Eq. (A.1) is

$$G(\xi, \xi') = -\frac{\omega_p}{v_b} \sin \left[\frac{\omega_p}{v_b} (\xi - \xi') \right] \Theta(\xi' - \xi), \quad (\text{A.2})$$

where $\Theta(-\xi)$ is the Heavyside step function. The Green's function can be verified by checking that

$$\frac{\partial^2 G(\xi, \xi')}{\partial \xi^2} + \frac{\omega_p^2}{v_b^2} G(\xi, \xi') = -\frac{\omega_p^2}{v_b^2} \delta(\xi - \xi'). \quad (\text{A.3})$$

We start with the left hand side:

$$\begin{aligned} & \frac{\partial^2 G(\xi, \xi')}{\partial \xi^2} + \frac{\omega_p^2}{v_b^2} G(\xi, \xi') \\ &= -\frac{\omega_p}{v_b} \frac{\partial^2}{\partial \xi^2} \sin \left[\frac{\omega_p}{v_b} (\xi - \xi') \right] \Theta(\xi' - \xi) - \frac{\omega_p^3}{v_b^3} \sin \left[\frac{\omega_p}{v_b} (\xi - \xi') \right] \Theta(\xi' - \xi) \\ &= -\frac{\omega_p}{v_b} \left\{ -\frac{\omega_p^2}{v_b^2} \sin \left[\frac{\omega_p}{v_b} (\xi - \xi') \right] \Theta(\xi' - \xi) + \sin \left[\frac{\omega_p}{v_b} (\xi - \xi') \right] \delta'(\xi - \xi') \right\} \\ &\quad - \frac{\omega_p^3}{v_b^3} \sin \left[\frac{\omega_p}{v_b} (\xi - \xi') \right] \Theta(\xi' - \xi) \\ &= \frac{\omega_p}{v_b} \sin \left[\frac{\omega_p}{v_b} (\xi - \xi') \right] \delta'(\xi - \xi') \\ &= -\frac{\omega_p^2}{v_b^2} \delta(\xi - \xi'). \end{aligned}$$

Appendix B

Adiabatic Solution for the Beta Function

The adiabatic solution for the beta function is given by using the theory of adiabatic invariants to find the evolution of the action-angle coordinates. These coordinates are used to derive the transport matrix for a single beam particle, from which the evolution of β is found using Eq. (2.140). The single particle Hamiltonian is given by

$$H = \frac{1}{2}x'^2 + \frac{1}{2}\eta x^2. \quad (\text{B.1})$$

The action-angle coordinates for a constant η are given by the generating function

$$F_1(x, \theta_0) = -\frac{x^2}{2\beta_m} \tan \theta_0. \quad (\text{B.2})$$

In an adiabatic plasma, the lowest order expansion in α_m gives the evolution as (see Sec. 3.2.1 for the derivation)

$$\theta_0^{(0)} = \int_0^s \frac{d\xi}{\beta_m(\xi)} + \theta_0(0), \quad J_0^{(0)} = J_0(0). \quad (\text{B.3})$$

The action-angle coordinates are related to x and x' by [127]

$$x' = \frac{\partial F_1}{\partial x} = -\frac{x}{\beta_m} \tan \theta_0, \quad J = -\frac{\partial F_1}{\partial \theta_0} = \frac{x^2}{2\beta_m} \sec^2 \theta_0. \quad (\text{B.4})$$

Solving the two equations for J gives

$$J = \frac{1}{2}\beta_m(\eta x^2 + x'^2). \quad (\text{B.5})$$

Similarly, we can write x and x' in terms of the action-angle coordinates

$$x = \sqrt{2J_0\beta_m} \cos \theta_0, \quad x' = -\sqrt{\frac{2J_0}{\beta_m}} \sin \theta_0. \quad (\text{B.6})$$

The solution for x and x' in the adiabatic case are trivially given by

$$x(s) = \sqrt{2J_0^{(0)}\beta_m} \cos [\theta_0^{(0)}(s)], \quad x'(s) = -\sqrt{\frac{2J_0^{(0)}}{\beta_m}} \sin [\theta_0^{(0)}(s)]. \quad (\text{B.7})$$

We can find the transport matrix for a single particle by finding particular solutions for $x_1(0) = 0$ and $x'_1(0) = x'_0$, and $x_2(0) = x_0$ and $x'_2(0) = 0$. Each solution gives two elements of the transport matrix. In the first case, $\theta_0(0) = -\pi/2$ and $J_0(0) = \frac{\beta_{m0}}{2}y_0^2$; the solutions for x_1 and x'_1 are given by inserting the initial conditions into Eq. (B.7):

$$x_1(s) = x'_0 \sqrt{\beta_m \beta_{m0}} \cos [\theta_0^{(0)}(s)], \quad x'_1(s) = -x'_0 \sqrt{\frac{\beta_{m0}}{\beta_m}} \sin [\theta_0^{(0)}(s)]. \quad (\text{B.8})$$

In the second case $\theta_0(0) = 0$, $J_0(0) = x_0^2 \frac{1}{2\beta_{m0}}$, and the solutions are

$$x_2(s) = x_0 \sqrt{\frac{\beta_m}{\beta_{m0}}} \cos [\theta_0^{(0)}(s)], \quad x'_2(s) = -x_0 \sqrt{\frac{1}{\beta_{m0}\beta_m}} \sin [\theta_0^{(0)}(s)]. \quad (\text{B.9})$$

The two sets of solutions have different values of $\theta_0(0)$, and therefore, $\theta_0^{(0)}$ is not defined the same way for both sets. We can fix this by rewriting everything in terms of the betatron phase advance θ defined as

$$\theta = \int_0^s \frac{d\xi}{\beta_m(\xi)}. \quad (\text{B.10})$$

The four solution then become

$$\begin{aligned} x_1(s) &= x'_0 \sqrt{\beta_m \beta_{m0}} \sin \theta, & x'_1(s) &= x'_0 \sqrt{\frac{\beta_{m0}}{\beta_m}} \cos \theta \\ x_2(s) &= x_0 \sqrt{\frac{\beta_m}{\beta_{m0}}} \cos \theta, & x'_2(s) &= -x_0 \sqrt{\frac{1}{\beta_{m0}\beta_m}} \sin \theta \end{aligned} \quad (\text{B.11})$$

From these four solution we can read off the elements of the transport matrix

$$M = \begin{pmatrix} \sqrt{\frac{\beta_m}{\beta_{m0}}} \cos \theta & \sqrt{\beta_m \beta_{m0}} \sin \theta \\ -\sqrt{\frac{1}{\beta_{m0}\beta_m}} \sin \theta & \sqrt{\frac{\beta_{m0}}{\beta_m}} \cos \theta \end{pmatrix}, \quad (\text{B.12})$$

which can be written as a squeeze then a rotation then a squeeze in the opposite direction

$$M = \begin{pmatrix} \sqrt{\beta_m} & 0 \\ 0 & \sqrt{\frac{1}{\beta_m}} \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \sqrt{\frac{1}{\beta_{m0}}} & 0 \\ 0 & \sqrt{\beta_{m0}} \end{pmatrix} \quad (\text{B.13})$$

Using these we can write down the solution for the beta function using Eq. (2.140):

$$\beta = M_{11}^2 \beta_0 - 2M_{11}M_{12}\alpha_0 + M_{12}^2 \gamma_0. \quad (\text{B.14})$$

Inserting the elements of the transport matrix, we get that, in the adiabatic limit, the beta function evolves according to

$$\beta = \frac{\beta_m}{2} \left(\frac{\beta_0}{\beta_{m0}} - \beta_{m0}\gamma_0 \right) \cos 2\theta - \beta_m \alpha_0 \sin 2\theta + \frac{\beta_m}{2} \left(\frac{\beta_0}{\beta_{m0}} + \beta_{m0}\gamma_0 \right). \quad (\text{B.15})$$

Appendix C

Perturbative Solution for the Beta Function

In the regime where $\alpha_m \ll 1$, the plasma should act only as a perturbation to the beam, resulting in a small deviation from free space propagation. As such, the evolution can be looked at using classical time-dependent perturbation theory. Imagine we have a Hamiltonian with a known solution, denoted $H_0(q, p, t)$, and that the solution has been obtained using Hamilton's principle function to generate a canonical transformation to a new Hamiltonian, which is zero. The transformed variables denoted by A and B are constant. The perturbed Hamiltonian is given by

$$H(q, p, t) = H_0(q, p, t) + \Delta H(q, p, t). \quad (\text{C.1})$$

The transformed Hamiltonian is given by

$$K(A, B, t) = H_0 + \Delta H + \frac{\partial S}{\partial t} = \Delta H(A, B, t). \quad (\text{C.2})$$

The equations of motion for A and B are given by

$$\dot{A} = -\frac{\partial \Delta H(A, B, t)}{\partial B}, \quad \dot{B} = \frac{\partial \Delta H(A, B, t)}{\partial A}. \quad (\text{C.3})$$

Now enters the perturbation theory, plug the constant values of A and B into the right hand sides of the equations above to get the first order correction, plug this in to get the second, and so on.

For transverse motion in a plasma-based accelerator, the variables (p, q) are replaced by (y, y') . The unperturbed Hamiltonian is simply that of free space propagation

$$H_0 = \frac{1}{2}y'^2, \quad (\text{C.4})$$

and the perturbing Hamiltonian is the focusing force given by

$$\Delta H = \frac{1}{2}\eta y^2. \quad (\text{C.5})$$

The Hamilton-Jacobi equation is easy to solve for this system (the solution is in Ref. [127]) and the transformed coordinates are given by

$$A = y'_0, \quad B = y - As. \quad (\text{C.6})$$

The inverse transformation is $y = As + B$. Both coordinates are constants of the motion. The perturbed Hamiltonian written in terms of these coordinates is given by

$$\Delta H(A, B, s) = \frac{1}{2}\eta(s)(As + B)^2. \quad (\text{C.7})$$

The equations of motion for the canonical coordinates are

$$\dot{A} = -\eta(s)(As + B), \quad \dot{B} = \eta(s)s(As + B). \quad (\text{C.8})$$

The zeroth order solution is the constant solution of the unperturbed Hamiltonian $A_0 = y'_0$ and $B_0 = y_0$. The first order solution is found by inserting these into the right hand side of the differential equations

$$\dot{A}_1 = -\eta(s)(A_0s + B_0), \quad \dot{B}_1 = \eta(s)(A_0s^2 + B_0s). \quad (\text{C.9})$$

Which reduces down to calculating three integrals over the plasma density of the form

$$I_n(s) = \int_0^s x^n \eta(x) dx. \quad (\text{C.10})$$

The first order solution is given by

$$A_1 = A_0 - A_0 I_1(s) - B_0 I_0(s), \quad B_1 = B_0 + A_0 I_2(s) + B_0 I_1(s). \quad (\text{C.11})$$

Transforming back to the original coordinates gives

$$y = y_0(1 - I_0s + I_1) + y'_0((1 - I_1)s + I_2), \quad y' = -y_0 I_0 + y'_0(1 - I_1), \quad (\text{C.12})$$

which allows us to define the transport matrix as

$$M = \begin{pmatrix} 1 - I_0 s + I_1 & s - I_1 s + I_2 \\ -I_0 & 1 - I_1 \end{pmatrix}. \quad (\text{C.13})$$

The beta function evolution is given by Eq. (2.140)

$$\beta = \beta_0(1 - I_0 s + I_1)^2 - 2\alpha_0(1 - I_0 s + I_1)(s - I_1 s + I_2) + \gamma_0(s - I_1 s + I_2)^2. \quad (\text{C.14})$$

Appendix D

Particle Tracking Algorithm

This appendix is a direct reproductions of Appendix A from Ariniello *et al.* [58] © 2022 American Physical Society.

The particle tracking simulations shown in this dissertation all use a simple particle tracking code. Each particle is propagated by numerically solving Eq. (2.121). The values of $\gamma_b(s)$ and $\gamma'_b(s)$ are known from the acceleration model used for each simulation. The focusing force, parameterized by $k_\beta(s)$, is easily found from $\gamma_b(s)$ and the plasma density profile $n_e(s)$. Both $k_\beta(s)$ and $\gamma_b(s)$ are functions of the particles initial ξ , and either δ or τ . The particle transverse positions are updated from step s_i to s_{i+1} using the transport matrix formalism

$$\begin{pmatrix} x_{i+1} \\ x'_{i+1} \end{pmatrix} = M(s_{i+1}|s_i) \begin{pmatrix} x_i \\ x'_i \end{pmatrix}, \quad (\text{D.1})$$

where we use the standard transport matrix

$$M(s_{i+1}|s_i) = \begin{pmatrix} \cos(k_\beta \Delta s) & \frac{1}{k_\beta} \sin(k_\beta \Delta s) \\ -\theta k_\beta \sin(k_\beta \Delta s) & \theta \cos(k_\beta \Delta s) \end{pmatrix}. \quad (\text{D.2})$$

Here, $\Delta s = s_{i+1} - s_i$ is the step size, $k_\beta = k_\beta(s_i + \Delta s/2)$ is the betatron wavenumber evaluated at the half-step, and $\theta = 1 - \frac{\gamma'_b(s_i + \Delta s/2) \Delta s}{\gamma_b(s_i + \Delta s/2)}$ describes the adiabatic dampening. The step size Δs is much less than $1/k_\beta$.

The particles are initialized in action-angle variable space (J and ψ) using the distribution

$$\rho(J) = \frac{1}{\epsilon} e^{-J/\epsilon}, \quad (\text{D.3})$$

$$\rho(\psi) = \frac{1}{2\pi} \text{rect} \left(\frac{\psi}{2\pi} \right). \quad (\text{D.4})$$

The particle's initial position in real space is calculated from J and ψ using

$$x_0 = \sqrt{2J\beta_0} \cos \psi + \Delta x, \quad (\text{D.5})$$

$$x'_0 = -\sqrt{\frac{2J}{\beta_0}} (\sin \psi + \alpha_0 \cos \psi). \quad (\text{D.6})$$

The longitudinal positions are initialized based on the distribution of interest.

Appendix E

Beam Evolution in a Plasma Source with Adiabatic Ramps

This appendix is a direct reproductions of Appendix B from Ariniello *et al.* [58] © 2022 American Physical Society.

Finding the beam evolution in an adiabatic plasma source with ramps requires assuming the energy change is small in the ramps but including energy change in the bulk plasma. As a result, the analytic expressions are piece-wise with different expressions for the ramps and the bulk. The evolution is fully described by γ_b , $\bar{\phi}$, ω_ξ , and ω_δ . The expression for γ_b and $\bar{\gamma}_b$ are not piece-wise and are given in Sec. 3.4.6. The expressions for $\bar{\phi}$, ω_ξ , and ω_δ are

$$\bar{\phi} = \begin{cases} \frac{\omega_{pu}}{c\sqrt{2}\gamma_{b0}} \left[G_1 - \frac{E_0 e}{2m_e c^2 \gamma_{b0}} (2D_2 - D_1) \right] & s \leq l \\ \bar{\phi}|_l + \frac{\omega_{pu}}{c\sqrt{2}} \frac{2m_e c^2}{E_0 e} (\sqrt{\gamma_b} - \sqrt{\gamma_{bl}}) & l < s \leq l + L \\ \bar{\phi}|_{l+L} + \frac{\omega_{pu}}{c\sqrt{2}\gamma_{bL}} \left[G_1^* - \frac{E_0 e}{2m_e c^2 \gamma_{bL}} (2D_2^* - D_1^*) \right] & l + L < s \end{cases}$$

$$\omega_\xi = \begin{cases} \frac{\omega_{pu}}{c\sqrt{2}\gamma_{b0}} \frac{E_1 e}{2m_e c^2 \gamma_{b0} \Delta\xi} D_2 & s \leq l \\ \omega_\xi|_l - \frac{\omega_{pu}}{c\sqrt{2}} \frac{E_1}{E_0 \Delta\xi} \left[\frac{1}{\sqrt{\gamma_b}} G_2 + \frac{2m_e c^2}{E_0 e} (\sqrt{\gamma_{bl}} - \sqrt{\gamma_b}) - \frac{1}{\sqrt{\gamma_{bl}}} G_2(l) \right] & l < s \leq l + L \\ \omega_\xi|_{l+L} + \frac{\omega_{pu}}{c\sqrt{2}\gamma_{bL}} \frac{E_1 e}{2m_e c^2 \gamma_{bL} \Delta\xi} [D_2^* + G_1^* G_2(L)] & l + L < s \end{cases}$$

$$\omega_\delta = \begin{cases} \frac{1}{2} \frac{\omega_{pu}}{c\sqrt{2}\gamma_{b0}} G_1 & s \leq l \\ \omega_\delta|_l - \frac{\omega_{pu}}{c\sqrt{2}} \frac{m_e c^2 \gamma_{b0}}{E_0 e} \left(\frac{1}{\sqrt{\gamma_b}} - \frac{1}{\sqrt{\gamma_{bl}}} \right) & l < s \leq l + L \\ \omega_\delta|_{l+L} + \frac{1}{2} \frac{\omega_{pu}}{c\sqrt{2}\gamma_{bL}} \frac{\gamma_{b0}}{\gamma_{bL}} G_1^* & l + L < s, \end{cases}$$

where G_1 , G_2 , D_1 , and D_2 are integrals over the plasma density defined in Sec. 3.4.6. The starred quantities are the integrals evaluated over only the exit ramp:

$$G_1^* = G_1(s) - G_1(l + L)$$

$$G_2^* = G_2(s) - G_2(l + L)$$

$$D_1^* = \int_{l+L}^s ds' \sqrt{\eta(s')} G_1^*(s')$$

$$D_2^* = \int_{l+L}^s ds' \sqrt{\eta(s')} G_2^*(s').$$

Combining these expressions with Eqs. (3.65), (3.57) and (3.58) gives H_1 , H_2 , ψ_1 , and ψ_2 . These can be inserted into Eq. (3.59) to find the emittance growth and Eqs. (3.50)-(3.54) to find the evolution of the beam moments.

Appendix F

Versions of PPT

The PPT model is referenced throughout the literature and unfortunately the equations used are not all the same. In this section I will compare the equations and try to rectify discrepancies between them. The original PPT model with the Coulomb correction is given by [93, 97]

$$w_{lm}(E, \omega) = |C_{n^*l^*}|^2 \left(\frac{6}{\pi}\right)^{1/2} f(l, m) \xi_i \left(\frac{2E_0}{E}\right)^{2n^* - |m| - 3/2} \left(\sqrt{1 + \gamma^2}\right)^{|m| + 3/2} \times A_m(\omega, \gamma) \exp\left(-\frac{2E_0}{3E} g(\gamma)\right). \quad (\text{F.1})$$

Lai *et al.* [100] use the PPT theory to calculate ionization rates and compare them to experimental data. Particularly for linearly polarized light, the model they use has good agreement with the experimental data. The model they use is given by

$$w_{lm}(E, \omega) = |C_{n^*l^*}|^2 f(l, m) \xi_i \left(\frac{2E_0}{E}\right)^{2n^* - |m| - 1} \left(\sqrt{1 + \gamma^2}\right)^{|m| + 3/2} A_m(\omega, \gamma) \exp\left(-\frac{2E_0}{3E} g(\gamma)\right), \quad (\text{F.2})$$

which differs from the PPT model by a factor of

$$\left(\frac{3E}{\pi E_0}\right)^{1/2}, \quad (\text{F.3})$$

which is the term in the ADK model that accounts for averaging over a single cycle of the laser field, which is consistent with the claims in the paper that state the $\sqrt{3E/\pi E_0}$ factor results from cycle averaging. They also have a difference in the function A_m . In the original PPT paper, A_m was given by

$$A_m(\omega, \gamma) = \frac{4}{\sqrt{3\pi}} \frac{1}{|m|!} \frac{\gamma^2}{1 + \gamma^2} \sum_{n \geq v}^{\infty} \exp\{-\alpha(n - v)\} w_m\left(\sqrt{\beta(n - v)}\right). \quad (\text{F.4})$$

The version in Ref. [100] is

$$A_m(\omega, \gamma) = \frac{4\gamma^2}{\sqrt{3\pi} |m|! (1 + \gamma^2)} \sum_{q \geq q_{min}}^{\infty} \exp\{-\alpha(q - v)\} w_m \left(\sqrt{\beta(q - q_{min})} \right), \quad (\text{F.5})$$

where $q_{min} = (\xi_i + U_p)/\omega$, which is the minimum number of photons necessary to reach the effective ionization threshold. U_p is the pondermotive potential given in atomic units by

$$U_p = \frac{E^2}{4\omega^2}. \quad (\text{F.6})$$

It ends up that

$$U_p = \frac{\xi_i}{2\gamma^2}. \quad (\text{F.7})$$

Therefore, $q_{min} = v$ and the version matches the original PPT formula exactly.

Another version of the formula is presented by Popruzhenko *et al.* [99] in which the ionization rate for an s state ($l=m=0$) is given by

$$w_{lm}(E, \omega) = |C_{n^*l^*}|^2 \frac{2}{\pi} \xi_i \left(\frac{\omega}{\xi_i} \right)^{3/2} \sqrt{\beta} \sum_{n \geq v}^{\infty} \exp\{-\alpha(n - v)\} \mathcal{F} \left(\sqrt{\beta(n - v)} \right) \exp \left(-\frac{2E_0}{3E} g(\gamma) \right), \quad (\text{F.8})$$

where $\mathcal{F}$ is the Dawson integral given by

$$\mathcal{F}(x) = e^{-x^2} \int_0^x e^{y^2} dy. \quad (\text{F.9})$$

It is easy to see that $\mathcal{F}(x) = w_0(x)$. We can pull out the function A_0 which simplifies the expression for the ionization rate to

$$w_{lm}(E, \omega) = |C_{n^*l^*}|^2 \left(\frac{6}{\pi} \right)^{1/2} \xi_i \left(\frac{\omega}{\xi_i \beta} \right)^{3/2} \sqrt{2} A_0(\omega, \gamma) \exp \left(-\frac{2E_0}{3E} g(\gamma) \right). \quad (\text{F.10})$$

Using the definition of β gives

$$\left(\frac{\omega}{\xi_i \beta} \right)^{3/2} = \left(\frac{\omega \sqrt{1 + \gamma^2}}{2\gamma \xi_i} \right)^{3/2}. \quad (\text{F.11})$$

Inserting the definition of γ for the γ in the denominator gives

$$\left(\frac{\omega \sqrt{1 + \gamma^2}}{2\gamma \xi_i} \right)^{3/2} = 2^{3/2} \left(\frac{E \sqrt{1 + \gamma^2}}{2E_0} \right)^{3/2}. \quad (\text{F.12})$$

Inserting back into the original equation produces the original PPT theory evaluated at $l = m = 0$ without the Coulomb term and with an additional factor of 4:

$$w_{lm}(E, \omega) = 4 |C_{n^*l^*}|^2 \left(\frac{6}{\pi}\right)^{1/2} \xi_i \left(\frac{E\sqrt{1+\gamma^2}}{2E_0}\right)^{3/2} A_0(\omega, \gamma) \exp\left(-\frac{2E_0}{3E}g(\gamma)\right), \quad (\text{F.13})$$

The Coulomb term is explicitly removed in Ref. [99] thus its absence is expected. The corrected Coulomb term is calculated in this paper so they present the formula without the Coulomb correction. Most likely the factor of 4 comes from a difference in the definition of the coefficient $|C_{n^*l^*}|^2$. In previous works, Ref. [98], the authors used the Hartree expression for $|C_{n^*l^*}|^2$ which they defined as

$$|C_{n^*l^*}|^2 = \frac{2^{2n^*-2}}{n^*\Gamma(n^*+l+1)\Gamma(n^*-l)}, \quad (\text{F.14})$$

which is smaller by a factor of 4 from the coefficient used in the ADK theory and that used by Ref. [100]. Using this definition of $|C_{n^*l^*}|^2$ their form of the theory matches the original version.

An additional version of the PPT theory comes from Couairon and Mysyrowicz [95]:

$$w_{lm}(E, \omega) = \frac{\omega_{a.u.}}{2\xi_H} |C_{n^*l^*}|^2 \left(\frac{6}{\pi}\right)^{1/2} f(l, m) \xi_i \left(\frac{2E_0}{E\sqrt{1+\gamma^2}}\right)^{2n^*-|m|-3/2} A_m(\omega, \gamma) \exp\left(-\frac{2E_0}{3E}g(\gamma)\right). \quad (\text{F.15})$$

This differs from the original formula by the factor of

$$\sqrt{1+\gamma^2}^{-2n^*} \quad (\text{F.16})$$

and the $\omega_{a.u.}/2U_H$ factor in front. The definition of v also has an additional $\hbar$ in it from converting the photon energy from atomic units, where it has the same units as frequency, to eV

$$v = \frac{\xi_i}{\hbar\omega} \left(1 + \frac{1}{2\gamma^2}\right). \quad (\text{F.17})$$

The definition of E_0 is

$$E_0 = E_H \left(\frac{\xi_i}{\xi_H}\right)^{3/2}, \quad (\text{F.18})$$

and the definition of n^* is given by

$$n^* = Z \sqrt{\frac{\xi_H}{\xi_i}}. \quad (\text{F.19})$$

Everything with a H subscript refers to the equivalent quantity for atomic hydrogen. These quantities are given by

$$E_H[\text{V}/\text{m}] = \frac{\alpha^3 m_e^2 c^3}{\hbar e}, \quad (\text{F.20})$$

$$\xi_H[\text{eV}] = \frac{\alpha^2 m_e c^2}{2e}, \quad (\text{F.21})$$

$$\omega_{a.u.} = \frac{e E_H}{\sqrt{2} m_e \xi_H} = \frac{\alpha^2 m_e c^2}{\hbar}. \quad (\text{F.22})$$

The overall factor in SI is given by

$$\frac{\omega_{a.u.}}{2\xi_H} = \frac{\alpha^2 m_e c^2}{\hbar} \frac{e}{\alpha^2 m_e c^2} = \frac{e}{\hbar}, \quad (\text{F.23})$$

which exactly matches the conversion constant from Sec. 4.1.3. Plugging in all the hydrogen terms for E_0 gives

$$E_0 = 2^{3/2} \frac{\sqrt{m_e e}}{\hbar} \xi_i^{3/2}, \quad (\text{F.24})$$

which again matches the result Sec. 4.1.3. The unit conversion for n^* also matches Sec. 4.1.3:

$$n^* = \alpha c \sqrt{\frac{m_e}{2e}} \frac{Z}{\sqrt{\xi_i}}. \quad (\text{F.25})$$

The only unresolved term in this version of the formula is the $\sqrt{1 + \gamma^2}^{-2n^*}$ which comes from Ref. [90], where they claims it is part of the Coulomb correction from Ref. [97]. However, it is clearly not part of the Coulomb correction in [97], and it does not appear in any other papers; therefore, it is most likely an error in this reference.

The last piece that is unclear is a statement by Ref. [95] that the PPT ionization theory underestimates the ionization rates by a factor of 4. They reference Mishima *et al.* [128] as the source of this claim. However, Ref. [128] is fairly unintelligible so the validity of this claim is suspect.