UNIVERSITY OF CALIFORNIA
Los Angeles

Advancing Ultrafast Nano Probes: Imaging and Diffraction

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Physics

by

Paul Elliot Denham

2024

ABSTRACT OF THE DISSERTATION

Advancing Ultrafast Nano Probes: Imaging and Diffraction

by

Paul Elliot Denham

Doctor of Philosophy in Physics

University of California, Los Angeles, 2024

Professor Pietro Musumeci, Chair

To enable more groundbreaking scientific discoveries, imaging, diffraction, and spectroscopy techniques must evolve beyond traditional static or low-speed imaging by dramatically improving temporal resolution. Being able to resolve processes at the ps and fs time scales would enable visualization of the dynamics of atomic motion at its fundamental timescales but requires the development of new electron sources and advanced optics. Among the techniques under active research, ultrafast electron diffraction (UED) is the most advanced, while ultrafast transmission electron microscopy (UTEM) and ultrafast electron energy loss spectroscopy (EELS) are still in earlier research and development stages. Concept designs for these instruments leverage high extraction field radiofrequency photoinjectors as these sources can produce dense electron bunches with the necessary brightness to enable high spatial and temporal resolution. This work presents several techniques developed in order to preserve the initial beam brightness from the cathode to the detector, thereby meeting the stringent requirements for capturing ultrafast dynamics and advancing ultrafast electron scattering instrumentation.

We focus on the opportunities offered by operating a photoinjector in the so-called cigar regime, where the beam's elongated aspect ratio in its own reference frame enables the requisite brightness for high spatiotemporal resolution in UTEM, and investigate the longitudinal phase space manipulation of these beams using RF fields. Pairing the Pegasus RF photoinjector source with a newly installed 3rd harmonic RF cavity, we show the 6D phase space of the electron beam can be shaped to achieve optimal bunching conditions or minimal energy spread, making Pegasus well-suited for high-fidelity UED or UTEM. An envelope equation-based approach is employed to derive analytical scal-

ing laws for RF-based pulse compression, revealing the dependencies on beam energy and charge. Our results indicate that relativistic energies are crucial for achieving sub-femtosecond pulse lengths with electron bunches containing 10^6 electrons. We further demonstrate experimentally that the 3rd harmonic removes non-linear effects of RF curvature, hence, shrinking energy spread by nearly two orders of magnitude to 10 parts per million and paving the way for ultrashort beams in the sub-femtosecond regime.

Next, we theoretically address the effects of space charge fields on imaging performance in single-shot time-resolved TEM. By employing a Green's function perturbation method, we derive analytical estimates of space charge-induced aberration coefficients and validate them through particle tracking simulations. Our findings provide critical insights into how space charge nonlinearity affects image formation and offer fundamental scaling laws for balancing temporal and spatial resolution in time-resolved TEM. These results provide an important framework for improving the performance of ultrafast electron scattering instruments, particularly in high-charge, single-shot modes.

Finally, we propose advancing UED to higher beam energies, potentially exceeding 10 MeV. Higher beam energies flatten the Ewald sphere, bringing higher-order Bragg reflections into the field of view, while also reducing space charge effects, allowing more charge to be loaded into the bunch and enhancing the intensity of Bragg orders. Additionally, higher energies provide greater penetration depths and improved temporal resolution, though they introduce challenges related to beam rigidity and focusing. We address these challenges by utilizing post-sample strong focusing permanent magnet quadrupole (PMQ) optics, which provide angular magnification to overcome the point spread of the detector. Our method employs a triplet of compact, high field gradient (> 500 T/m), small-gap (3.5 mm) Halbach PMQs. These PMQ lenses allow us to maintain high-quality diffraction patterns at higher energies. With this optical setup, we demonstrate a tunable camera length, achieving a $6\times$ improvement and reciprocal space resolution better than 0.1 Å^{-1} with an 8.2 MeV electron beam and a crystal Au sample. Future designs should consider larger aperture PMQs to capture more Bragg orders, as larger apertures also reduce the aberrations.

The dissertation of Paul Elliot Denham is approved.

Brian C. Regan

Anshul Kogar

James Rosenzweig

Pietro Musumeci, Committee Chair

University of California, Los Angeles

2024

To my family..

Table of Contents

List of Figures

List of Tables

Acknowledgments

Growing up, I had no clear idea of what I wanted to pursue in life. After high school, I spent my time working minimum wage jobs, training in Brazilian Jiu-Jitsu, and playing competitive computer games—until a severe leg injury changed everything. That injury gave me the unexpected time to focus on community college, where I unknowingly developed an obsession with physics, setting me on the path to becoming a physicist. Perhaps it was the Feynman Lectures in sensei Kevin Howell's study that first piqued my curiosity. Twelve years later, I now find myself at the culmination of a doctorate in physics—with a severe foot injury once again.

I want to acknowledge all the teachers I had in community college, Pablo Bert, Bruce Chaffee, and Lee Loveridge, who were my mentors at the beginning of this journey, who generously answered my questions and encouraged me to pursue a higher education in physics.

To my undergraduate research advisor, Smadar Naoz: thank you for your early encouragement to continue to graduate school and for the opportunity to research 4-body planetary systems in planetary dynamics. Your initial guidance was crucial to my future development as a dynamicist.

To all the PBPLers who have moved on or are about to: I want to especially acknowledge David Cesar and Jorge Giner for being such incredible mentors at the start of my time at PBPL. Andrew Urbanowicz and Chad Pennington, who were there with me from the beginning in undergrad, and Eric Cropp, for all the late-night sessions working on flat beams and skew quad transformations—through those long hours, I gradually learned how to operate Pegasus. Gerard Lawler, Walter Lynn, Krish Kabra, River Robles, Pratik Manwani, and Joshua Mann, thank you for your insight and camaraderie in the lab.

Andrew Fisher, Sophie Crisp, Alex Ody, Youna Park, Max Lenz, Nathan Burger, David Dang, Edmund Liu, Atharva Kulkarni, and Veronica Guo: Thank you all for your critical roles and for spending countless hours with me running the beamline, solving technical problems, and offering support through all the struggles in the lab. I'm incredibly grateful to have worked with you all.

Oliver Williams and David Schiller, you two hold a special place in my heart. Through-

out graduate school, you were the ones who truly appreciated my humor, and I love you both for that. You made this journey an absolute pleasure.

I want to acknowledge Gerard Andonian. Working with you, sharing ideas, and hanging out has been a blast. You have been like a second advisor, playing a major role in my growth and development. Always approachable and wise, you have been a guiding presence, and I am grateful for the opportunity to work with you.

I would also like to thank my advisor, Pietro, for your patience, guidance, and openness to discussion. Since my undergraduate years, your door has always been open, welcoming me in. It has been a genuine pleasure being mentored by you and being shaped into the physicist I am today. Witnessing your dedication to the lab has been a huge source of inspiration, and I am grateful for the opportunity to have worked with such excellence.

I can't forget the countless hours spent playing foosball with all of you. Those moments made PBPL such a great place to work and helped create an unforgettable atmosphere.

I want to acknowledge my family: my mom, who has always been there for me, enabling me to return to my education in my early twenties, and for the tremendous support you've given me throughout my life. This would not have been possible without you. Likewise, my grandmother always encouraged me to pursue education from childhood. As a first-generation college student, I owe so much to you both. My dad, stepdad Robert, brother Nolan, and longtime friend Kyle have all played significant roles in showing me the meaning of hard work.

Finally, I thank my wife, the love of my life, Jacqueline, who brought our son Daniel into the world at the beginning of this dissertation and supported me through the most challenging moments of this journey. I love you and am infinitely grateful that we are together.

Vita

2017	Bachelor of Science in Physics and Minor in Mathematics, University of California, Los Angeles, *summa cum laude*.
2020	M.S. (Physics), University of California, Los Angeles.
2020 – 2024	Graduate Student Researcher, Physics and Astronomy, UCLA.

Publications

P. Denham†, N. Burger D. Cesar, D. Dang, P. Musumeci, "Progress in Time-Resolved MeV Transmission Electron Microscopy at UCLA", Contribution to NAPAC2019 (2019).

P. Denham, F. Cropp, P. Musumeci, "Analysis of Skew Quadrupole Compensation in RF-Photoinjectors", e-Print: 2003.00049 [physics.acc-ph] (2020).

P. Denham, P. Musumeci, "Space Charge Aberrations in Single Shot Time-Resolved Transmission Electron Microscopy", Phys. Rev. Applied 15 (2021) 2, 024050 (2021).

P. Denham, P. Musumeci, "Analytical Scaling Laws for Radiofrequency-Based Pulse Compression in Ultrafast Electron Diffraction Beamlines", Instruments 7 (2023) 4, 49 (2021).

V. Guo, P. Denham, P. Musumeci, A. Ody, Y. Park, "4D Beam Tomography at the UCLA Pegasus Laboratory", JACoW IBIC2021 (2021) TUPP15 (2021).

P. Denham, P. Musumeci, "Pushing Spatial Resolution Limits In Single-Shot Time-Resolved Transmission Electron Microscopy at the UCLA Pegasus Laboratory", Contribution to IPAC 2021 (2021).

P. Denham, P. Musumeci, A. Ody, "Commissioning of an X-Band Cavity for Longitudinal Phase Space Linearization at UCLA PEGASUS Laboratory", JACoW IPAC2022 (2022) THPOST040 (2022).

P. Denham, P. Musumeci, A. Ody, "X-Band Harmonic Longitudinal Phase Space Linearization at the PEGASUS Photoinjector", JACoW NAPAC2022 (2022) TUPA77 (2022).

Y. Park, R. Agustsson, W.J. Berg, J. Byrd, T.J. Campese, D. Dang, P. Denham, et al., "Tapered helical undulator system for high efficiency energy extraction from a high brightness electron beam", Nuclear Instruments and Methods in Physics Research Section A, Vol. 1028, 2022.

A. Kulkarni, P. Denham, A. Kogar, P. Musumeci, "Dual Radiofrequency Cavity Based Monochromatization for High Resolution Electron Energy Loss Spectroscopy", JACoW NAPAC2022 (2022) TUXD6 (2022).

P. Denham, D. Gavryushkin, G. Andonian, N. Burger, N. Cook, et al., "Developments and characterization of a gas jet ionization imaging optical column", JACoW IPAC2023 (2023) THPL147 (2023).

P. Denham, Y. Yang, V. Guo, A. Fisher, X. Shen, T. Xu, R. J. England, R. K. Li, P. Musumeci, "High energy electron diffraction instrument with tunable camera length", Structural Dynamics, 25 March 2024.

CHAPTER 1

Introduction

1.1 Probing irreversible ultra-fast dynamics

Transmission Electron Microscopy (TEM) has historically offered the unique capability to capture static images of the nano and subatomic world [1]. However, it has fallen short in frame rate, which leaves a significant lack in our ability to track motion on these scales. For instance, during thermal excitation of a gold lattice, grain boundary motion occurs on a scale of several nanometers [2]; given the speed of sound, and the typical spatial scale of the crystal grains, the spatiotemporal scale approaches the nanometer/picosecond. Phenomena, such as phase transformations, defect migrations, and chemical reactions, occur on these scales and dictate the properties and behaviors of materials in various applications, branches of science ranging from microelectronics to biochemistry [3]. These phenomena benefit from being studied with pump-probe techniques. In the pump-probe scheme, a short laser pulse (the pump) excites the sample, and a delayed electron pulse (the probe) captures the resulting dynamics either in a direct image or encoded in a diffraction pattern, enabling the visualization of structural dynamics.

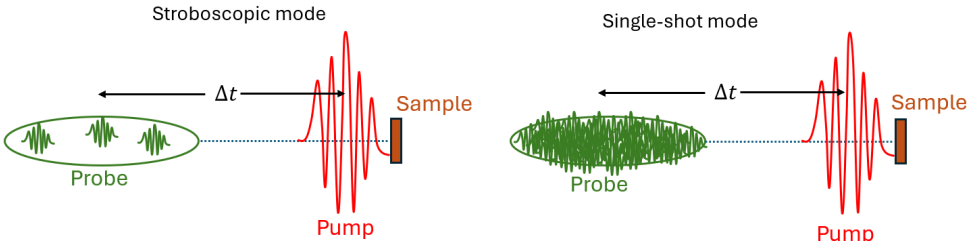


Figure 1.1: Illustration of laser pump-electron probe schemes, showing techniques arranged from left to right with increasing electron charge, designed to observe progressively irreversible phenomena. A controllable delay, Δt, between the pump and probe is intentionally introduced for temporally correlated studies.

There are primarily two modes of operation for pump-probe electron-based instrumentation:

Stroboscopic Mode: A pump-probe scheme with a variable repetition rate between 250 kHz and 25 MHz. The delayed probing electron pulse contains few electrons within a sub 100 fs envelope, minimizing space charge effects. This method provides excellent temporal and Angstrom-level spatial resolution but is limited to reversible processes where the sample recovers between excitations. To achieve high resolution, electron column currents are minimized, sometimes to single-electron emission levels, and the probing beam is synchronized with the pump laser delay. Despite requiring high repetition rates for reasonable acquisition times, this technique has enabled breakthroughs in visualizing electric and magnetic field dynamics, particularly in PINEM and magnetic vortex scenarios [4, 5, 6, 7, 8, 9, 10].

Single-Shot Mode: Captures an image in a single electron pulse. The pulse contains just enough electrons to form an image, typically more than 10^6, resulting in a peak current far exceeding that of conventional TEM. This high current leads to significant space charge effects, degrading resolution. To manage these effects, some groups use few-MeV electron sources to shorten the pulse, leveraging relativistic effects to suppress space charge [11, 12, 13, 14]. The main advantage of the single-shot method is its ability to study irreversible processes, unlike the stroboscopic approach.

The exposure times in conventional TEMs are many orders of magnitude beyond the ps scale because they typically operate with nA currents and a sufficient amount of electrons must be accumulated on the detector to form an image. So, there is a scientific interest to devise machines that can maintain sufficient spatial resolution while increasing the peak current for single-shot acquisition. To address this need, the Basic Energy Science Report on the Future of Electron Scattering recommended developing instruments capable of ultrafast electron diffraction (UED) with potentially sub-micron beam sizes, achieving 100 fs temporal resolution in single shot mode. Additionally, it

recommended single-shot real-space imaging (UTEM) with spatial-temporal resolution of 10 nm - 10 picoseconds (ps) [15].

The electron sources required for these instruments to work represent a thousand-fold improvement over current instruments. Single-shot microscopes require electron beam bunches with more than a million electrons, nanometer-normalized emittance, and a relative energy spread smaller than ten parts per million (ppm). The peak current of the electron source is several orders of magnitude higher than current electron microscopes. Moreover, the beam energy of these instruments is significantly higher, necessitating advanced electron optics capable of focusing the beam within a compact space. The consensus from the workshop on the Future of Electron Sources was that realizing these capabilities requires next-generation R&D in generating high-brightness beams, advanced cathodes, higher cathode extraction fields and beam energy, and advanced acceleration and beam manipulation schemes [16].

Time-resolved imaging and diffraction techniques can greatly benefit from particle accelerator developments. For example, radio frequency-driven electron sources for free electron lasers have led to accelerating gradients of 100 MV/m, enabling rapid acceleration of short pulses without being degraded by space charge. Many facilities have adopted this RF source for UED and UTEM to deliver beams with variable charge with good quality while providing more control over the full beam phase space all in a compact footprint [17, 18, 19, 20, 21, 22].

Following this trend multiple beamlines have incorporated advanced accelerator beam manipulation technologies in time-resolved imaging applications. Examples include the use of high-accelerating-gradient RF cavities to reverse space-charge-induced temporal broadening [23, 24, 25, 26, 27]. RF deflecting cavities have been successfully used as ultrafast streak cameras [28], high-speed beam choppers—creating ultrashort, low-charge bunches with excellent beam quality at a high repetition rate by sweeping a continuous beam across an aperture [29]—and as high-resolution time-of-flight spectrometers [30]. Achromatic beam transport has made it possible to passively reverse space-charge-induced expansion while reducing the time-of-arrival jitter of the electron pulse at the sample [31, 32]. These demonstrations indicate a promising future for UED and UTEM probes as more electron scattering instruments adopt accelerator and beam physics tech-

niques.

In this thesis, the theory of 6D transport of relativistic electron beams is explored and applied to developing photoinjector-based UTEM and UED instruments. Advanced beam manipulation techniques are employed to optimally shape the 6D phase space of the electron beam, mitigating space charge effects and ensuring the desired spatiotemporal resolution is achieved. The operational mode focuses on utilizing a relatively low bunch charge to attain the low emittance required for probing nanometer-scale spatial resolution while maintaining just enough charge to form an image in a single shot without significant phase space dilution from non-linearities. The thesis reviews the theory of electron beam optics. Then it investigates longitudinal phase space shaping techniques to tailor the beam's energy spread and pulse duration for these applications. Then it explores transverse phase space shaping to mitigate the impact of transverse space charge effects on spatial resolution. Finally, the work delves into advanced permanent magnetic-based optics to enable higher energy UTEM and UED operation.

The rest of this chapter details the brightness requirements for functional UTEM and UED and concludes with a summary of the subsequent chapters that form the bulk of the dissertation.

1.2 Beam brightness requirements

1.2.1 Brightness definitions

A critical step in achieving spatiotemporal resolution required for successful UTEM operation is preserving the beam's phase space as it is transported from the source to the sample and on through the imaging transport. The figure of merit used to quantify phase space or beam quality is the volume occupied by the particles in phase space. A density distribution function can describe the beam 6d phase space. If the x-y-z phase space dynamics are decoupled and second-order moments of that distribution are used to describe the extent of the particles in phase space, then areas in respective sub phase spaces are called normalized RMS emittance. In each respective phase plane we can write:

$$\epsilon_{nqr} = \frac{1}{mc}\sqrt{\langle q^2\rangle\langle p_q^2\rangle - \langle qp_q\rangle^2}, \tag{1.1}$$

where q and p_q are conjugate variables, and q can be x,y, or z, m is the mass of the electron, c is the speed of light, x, y, z and p_x, p_y, p_z are the position and momentum. The angle brackets represent averages against the statistical distribution representing the ensemble of particles comprising the beam. If the x-y-z phase space dynamics are decoupled, then the product of the areas provide a measure of the 6D phase space volume $V_{6D} \sim \epsilon_{nxr}\epsilon_{nyr}\epsilon_{nzr}$. In the context of a beam phase space, the transverse momentum is typically approximated as paraxial and typically reparameterized in terms of the angle made by the electron trajectory in the horizontal direction, i.e., $p_x = p_c x'$ where p_c is the momentum of the beam, and x' is the angle made. Notions of phase space density, volume, and area are given greater depth in the second chapter. The smaller the emittance, the easier it becomes to keep the beam tightly collimated or bunched over longer distances and to be focused on smaller spots. Better transverse emittance also translates to sharper images in microscope setups. A smaller longitudinal emittance translates to a beam having less energy spread, which can be compressed into shorter pulses. A key fact of the rms emittance is that it remains constant when the equations of motion governing position and corresponding momentum are linear. However, it changes and commonly grows due to non-linearities, also called aberrations.

In terms of normalized RMS emittance, for an amount of charge Q, we can define the rms six-dimensional brightness as the inverse of phase space volume scaled by the charge in the beam

$$B_{6D,rms} = \frac{Q}{\epsilon_{nxr}\epsilon_{nyr}\epsilon_{nzr}}. \tag{1.2}$$

Four-dimensional brightness can be defined as:

$$B_{4D,rms} = \frac{Q}{\epsilon_{nxr}\epsilon_{nyr}}. \tag{1.3}$$

Since normalized rms emittance is constant in linear transport, so is brightness, which is determined at the source.

1.2.2 Phase space requirements for electron scattering instrumentation

We can estimate beam phase space quality requirements based on a desired performance. Contrast is essential for distinguishing features when scattering electrons off a sample. This is achieved by separating scattered and unscattered electrons either through beam

drift or using electron optics [1]. Ideally, the initial angular spread should be smaller than the first Bragg angle, which depends on the beam quality and required spot size.

For a thin crystalline sample with atomic periodic spacing a, the first-order Bragg reflection angles are proportional to $\theta_b = \frac{\lambda}{a}$, where λ is the electron wavelength. We need $\frac{\theta_b}{\sigma_\theta} = \frac{\lambda_c}{\epsilon_{nxr}} \frac{\sigma_x}{a} > 1$. For a fixed probe spot size of $1\mu m$, this requires the normalized rms emittance ϵ_{nxr} to be less than $\frac{\lambda_c \sigma_x}{a}$. With typical atomic spacings ranging from 1 to 5 angstroms, this results in a necessary normalized emittance between 5 nm and 20 nm.

Assuming the optical system is perfectly linear and the beam is monochromatic, the limiting aperture radius—and consequently the beam divergence—sets the diffraction limit on resolution. The Rayleigh criterion defines the resolution limit set by electron interference effects from the limiting aperture as $\mathcal{R}_d = 1.22\frac{\lambda}{\sigma_\theta}$ where λ is the electron wavelength and σ_θ is the maximum opening angle accepted by the limiting aperture.

Resolution is also significantly affected by energy spread and higher-order charged particle dynamics, major sources of aberrations. Chromatic aberration arises from non-monochromatic beams, where different energies have different focal planes. The transport coefficient C_c relates the offset from the linear imaging position to the energy design offset, typically similar in magnitude to the focal length. A bundle of rays from a point in the object plane, with a spread in angle and transverse momentum, forms a disk in the image plane with radius $\mathcal{R}_c = C_c \frac{\sigma_p}{p} \sigma_\theta$. Spherical aberration results from third-order contributions to the transport. The aberration coefficient C_s, similar in magnitude to the lens focal length, quantifies the displacement in the image plane of a ray originating from the object plane without lateral offset, based on its initial angle. A bundle of rays diverging from a point with an angle spread σ_θ will result in an error disk with radius $\mathcal{R}_s = C_s \sigma_\theta^3$.

The overall resolution, referring to the object plane, is determined by the quadrature sum of all competing aberrations:

$$d = \sqrt{\left(1.22\frac{\lambda}{\sigma_\theta}\right)^2 + \left(C_c \frac{\sigma_p}{p}\sigma_\theta\right)^2 + \left(C_s \sigma_\theta^3\right)^2} \tag{1.4}$$

For simplicity, we assume the resolution is well above the diffraction limit, as the de

[1] at the back focal plane of an objective lens, where the initial angle of an electron is directly related to its final position as $x = f\theta$

Broglie wavelength for MeV electrons is less than 1 pm. By neglecting the first term under the square root, we can estimate acceptable beam parameters using the spherical (C_s) and chromatic (C_c) aberration coefficients. These coefficients, as a first approximation, have values comparable to the lens focal distance, which can be as short as 2 cm for MeV electrons.

From the above equation for d, achieving a spatial resolution of 10 nm requires a relative energy spread lower than 10^{-4} and a collection semi-angle of less than 5 mrad. To enhance contrast, it is preferable to have a smaller divergence, ideally at most comparable to the Bragg angle, which is typically around 1 mrad for MeV electrons.

1.2.3 Single shot dose requirements

1.2.3.1 Imaging

In order to form an image of a sample, we hit the sample with the electron beam. Some fraction of the beam scatters and some transmits. Suppose we could perfectly intercept the scattered electrons and keep only the ones transmitted. If the transport is imaging the sample plane, then we will be able to see the features that ultimately causes the scattering because there will be a loss of intensity into the corresponding pixels at the detector.

For now, we won't concern ourselves with the intricacies of the scattering process, such as if the scattered and transmitted fractions can be isolated from one another spatially. For reference on the details consider [33].In this, we assume we can successfully isolate and block the scattered electrons. So we can think of this process as if a mask is applied to the image of the direct beam. The mask would apply a contrast map to each pixel so that, as a whole, an image is formed. We look at a pixel, having area A, and compare before and after the mask is applied. Before the mask is applied, the original pixel value is directly proportional to the intensity of the direct beam, (for simplicity, we assume the direct beam has a uniform circular transverse profile). Then once the mask is applied (the collimator intercepts the scattered beam), that pixel loses some intensity and takes on a new, lower value. We can contrast the difference between these events with the

contrast ratio:

$$C = \frac{I_b - I_s}{I_b} \tag{1.5}$$

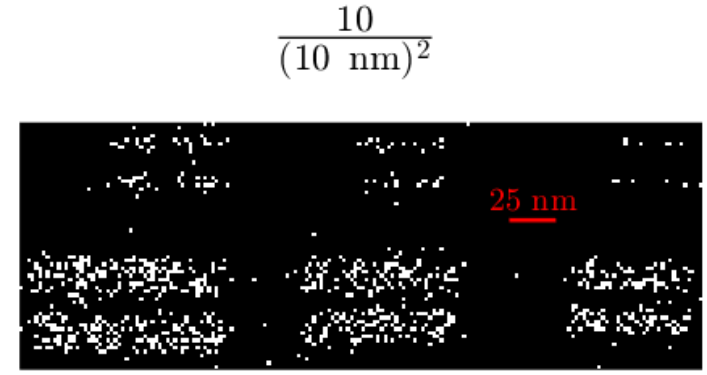
where I_b is the intensity of the pixel without the mask, i.e., when we are not blocking the scattered beam, and I_s is the intensity of that pixel when we do block the scattered beam. That is, if I_s is zero, all particles that were supposed to map to this pixel were scattered and successfully blocked, corresponding with perfect contrast or $C = 1$. On the other hand, if nothing was scattered, then $I_s = I_b$ and $C = 0$, so there is no contrast. Analogously, we can apply this same definition to contrast adjacent pixels.

If we consider the signal to be the difference, $\Delta I = I_b - I_s = CI_b$, then $N_s = A\Delta I = CAI_b$ gives the amount of blocked electrons. Meanwhile, the background noise level is taken when none of the electrons are blocked, $\sigma = \sqrt{AI_b}$, so the signal-to-noise ratio is $SNR = C\sqrt{AI_b}$. Thus, for a desired $SNR = 5$, and $C = 0.5$, we need $I_b = \frac{100}{A}$, or 100 electrons per pixel area. If the pixel area, is exactly one-to-one with a resolution area unit, with dimensions d^2, where d is the instrument's resolution, and we are trying to obtain $10nm$ resolution. We would need $1e/(nm^2)$ to distinguish the image visually. So if the beam fills a field of view that is a $1\mu m$ circle at the sample plane, the total charge needed to distinguish features is roughly 3×10^6 electrons, or about 0.5pC.

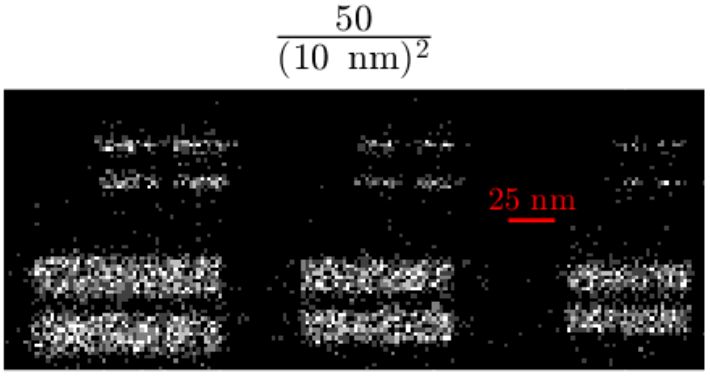
To simulate the effects of low-dose imaging, we employ a test target consisting of horizontal line pairs, the largest being 100 nm by 20 nm, decreasing in scale until the final set of bars is 10 nm by 2 nm. A beam having 3 mrad divergence is imaged with a magnification of 5 by a 5 cm focal length lens. The image is converted back to the object plane. Here the pixel size is 10nm. The results are shown in Fig. 1.6, for different doses: (a) $10e/(10\text{ nm})^2$, (b) $50e/(10\text{ nm})^2$, and (c) $250e/(10\text{ nm})^2$.

1.2.3.2 Diffraction

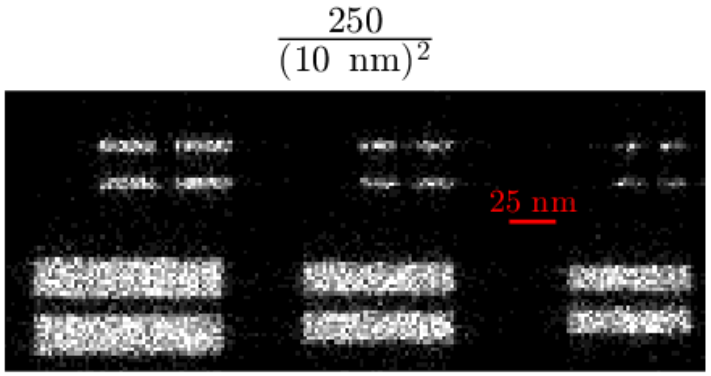
To estimate the charge required for producing diffraction patterns in a single shot, we start with a simple model of forming a diffraction pattern image to visualize desirable SNR in accordance with the Rose Criteria. We let an initial Gaussian distribution represent an image of the direct beam, which is what would be viewed on a screen when there is no sample inserted; $I(x, y) = I_0 \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$.



(a)



(b)



(c)

Figure 1.2: Simulated images of a sample target for different doses: (a) $10e/(10\ \text{nm})^2$, (b) $50e/(10\ \text{nm})^2$, and (c) $250e/(10\ \text{nm})^2$.

When a thin crystalline sample is inserted, the beam is scattered and produces an intensity distribution with a characteristic diffraction pattern, the resulting intensity distribution on the screen is $\widetilde{I}(x,y) = I_1\left(\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) + \sum_{k=1}^{M} F_k \exp\left(-\frac{(x-x_k)^2+(y-y_k)^2}{2\sigma^2}\right)\right)$, where I_1 represents the amplitude or peak value of the primary beam after diffraction, F_k denotes the fraction of I_1 corresponding to the intensity of Bragg peak k such that $I_k = F_k I_1$, x_k and y_k represent the mean or center of the Gaussian for Bragg peak k relative to the primary, respectively, M is the number of Bragg peaks, and σ is the root mean square (RMS) width of each Gaussian. Assuming Bragg peaks mirror the shape of the original beam, radially offset from the direct beam. Also, the direct beam intensity diminishes and redistributes into the Bragg peaks to conserve the total charge. Effectively, upon inserting the sample, a mask is applied to the original image. Conservation of charge implies that $I_0 = I_1(1 + \sum_{i=1}^{M} F_i)$, where $\sum_{i=1}^{M} F_i > 0$[2]. Each Gaussian can be normalized to the total number of electrons within it by using $N_k = (2\pi\sigma^2)I_k$, where N_k represents the total electron count on the screen in a particular Bragg peak. Since, $I_k = F_k I_1$ and I_1 is related to I_0, we have $N_k = F_k N_0/(1 + \sum_{i=1}^{m} F_i)$. We can invert the relation, $\sqrt{I_k A} = \text{SNR}$, to obtain the required input beam charge for a desired signal-to-noise ratio:

$$Q_0 = 2\pi(\text{SNR} \cdot \sigma(\text{pixels}))^2 \left(\frac{1 + \sum_{i=1}^{M} F_i}{F_k}\right) \text{(electrons)} \tag{1.6}$$

where σ is in units of pixels. This expresses the total bunch charge required to resolve a particular Bragg peak in a single shot. One interesting observation from this expression is that less charge is required as long as the spot size can be smaller, which makes sense because more electrons are packed into a pixel. So, a beam of higher quality having less charge can be sufficient for single-shot imaging. The nuclear cross-section and expected positions of the Bragg peaks can be used to theoretically determine $\sum_{i=1}^{M} F_i$ and F_k for a particular material. For simplicity, assume $\sum_{i=1}^{M} F_i = 0.25$, and $F_k = 0.01$, i.e., an intensity that is 100 times smaller than the transmitted beam, then to have a $SNR = 5$ from a 5 pixel spot, we would need , $Q_0 = 5 \times 10^5$e, nearly a 100fC, assuming the detector accurately counts single events.

In Fig. 1.3, we present 300x300 pixel diffraction pattern images that explore the impact of varying electron counts and Gaussian standard deviations on the diffraction

[2]if it were 0, then there is no scattering.

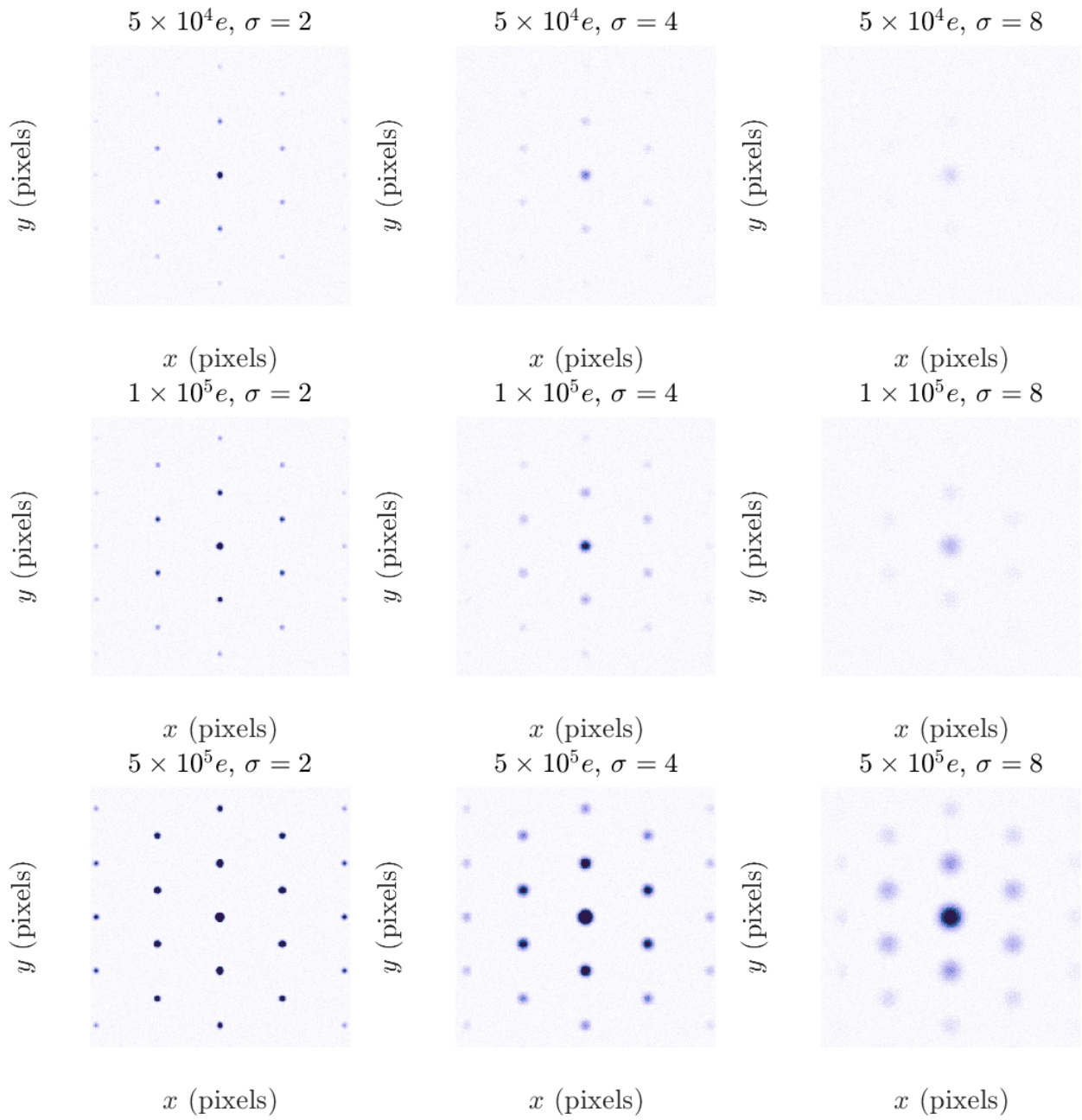


Figure 1.3: 300x300 pixel diffraction pattern images generated with varying total number of electrons and Gaussian standard deviation (σ). Each row represents a different total number of electrons (5×10^4, 1×10^5, 5×10^5), while each column represents a different Gaussian standard deviation ($\sigma = 2, 4, 8$). The images are overlaid with background Poisson noise, demonstrating the influence of these parameters on the diffraction pattern.

patterns. Each row corresponds to a different total number of electrons: 5×10^4, 1×10^5, and 5×10^5 electrons. Each column represents a different Gaussian standard deviation (σ): 2, 4, and 8 pixels. Pixel values range from 0 to 255. The images are overlaid with background Poisson noise to simulate realistic experimental conditions. The variations in electron count demonstrate how increasing the number of electrons enhances the intensity and definition of the diffraction spots. Meanwhile, the variations in σ illustrate the effect of beam divergence or focusing, where larger σ values result in broader and more diffuse diffraction spots and the possibility of using less charge.

1.2.4 Beam quality from photoemission

Based on the discussion in sections 1.2.3 and 1.2.4, we now understand the beam parameters—such as energy spread, emittance, and beam charge—required for electron scattering instrumentation in both imaging and diffraction modes. According to the Liouville theorem, the phase space density cannot be altered in Hamiltonian transport, making it crucial to start with the highest possible phase space density to meet these requirements. The photoemission process determines the initial beam quality.

The source size is determined by how tightly the drive laser can be focused. To address the reduced average current, single-shot UTEM and UED instruments generate pulses with many electrons per bunch from flat photocathode surfaces, where a smaller initial angular spread of emitted electrons increases beam brightness [34]. The precise timing control of laser-driven photoemission is also essential for adequate shaping of the electron beam temporal profile. Spicer's theory indicates that the duration of photoemission from metal cathodes is extremely brief, typically just a few femtoseconds, as documented in [35]. This duration is significantly shorter than the incident laser, whose optics can stretch to tens of picoseconds. Consequently, it is generally assumed that the initially emitted electron distribution mirrors the temporal profile of the incoming laser beam.

Assuming no correlation between the electron's emission angle and position, the transverse emittance upon photoemission is given by:

$$\epsilon_{x,th} = \sigma_{x_0} \sqrt{\frac{MTE}{mc^2}} \tag{1.7}$$

where MTE is the mean transverse energy, and mc^2 is the rest mass energy of the

electron. In typical photoemission systems, the MTE is expressed as $(\hbar\omega - W)/3$, where ω is the laser frequency, W is the material's work function, and $\hbar$ is the reduced Planck constant [36]. The initial spot size σ_{x_0} refers to the size of the laser beam on the cathode surface. To minimize the thermal contribution to emittance, reducing the laser spot size on the cathode is crucial. A best-case estimate of the initial emittance can be made by knowing the work function and the photon energy.

The presence of an applied electric field reduces the work function through the Schottky effect, effectively lowering the work function due to the image charge field's influence on the total potential by an amount $\Delta W = \sqrt{e^3 E_{\text{acc}}/4\pi\epsilon_0}$, where E_{acc} represents the applied electric field. The accelerating field of a state-of-the-art photo injector can, in principle, reach 100 MV/m at the cathode surface [37]. This accelerating field lowers the work function by about 0.37 eV and raises the excess energy to around 0.4 eV. If we could tune the laser wavelength, we could, in principle, lower the effective work function, but doing so negatively reduces the QE, and thus requires more laser energy to achieve the same amount of charge [38]. A more practical approach is to make the laser spot size on the cathode as small as possible. but there are two issues in making the spot small. A practical one is how far the last lens focusing the laser onto the cathode can be positioned outside the vacuum. A more fundamental one is the saturation current limitation, which is the amount of charge that can be extracted from very small spot sizes.

1.2.5 Saturation current

When trying to minimize the initial beam emittance by shrinking the laser spot size, we encounter a limit due to space charge effects [39]. In 1D, the classic Child-Langmuir equation predicts the limiting current density $j_{\text{CL,1D}} = \frac{4\epsilon_0}{9}\sqrt{\frac{2e}{m}}\frac{V_{acc}^{3/2}}{d^2}$ where V_{acc} is the total applied voltage across the gap, d is the gap length, ϵ_0 is the vacuum dielectric permittivity, m is the electron mass, and e is the electron charge.

This can be generalized to a realistic photoemission case, considering the variation of the electric field during the photoemission process; the on-axis electric field for a disk of radius R[3] is given by $E_z(s) = \frac{\sigma}{2\epsilon_0}\left(1 - |s|/\sqrt{s^2 + R^2}\right)$, where ζ is the on-axis coordinate measured from the center of the disk. This field decreases significantly with increasing

[3] representing a slice of the beam

distance from the disk, implying that particles beyond a distance $s \approx R$ contribute less to the fields at the cathode.

To modify the C-L equation to the diode geometry, we substitute the traditional gap length d with the effective length $z_{\text{eff}} \sim R$, and an effective gap voltage $V_{acc} \sim E_{acc} z_{eff} = E_{acc} R$, where E_{acc} is the constant gun accelerating field assumed in this derivation. This assumption is valid under static accelerating fields, such as DC guns, or when the laser pulse duration is significantly shorter than the radio frequency (rf) period. The two-dimensional space-charge limiting current for the effective diode can be calculated by incorporating these modifications into Child's law and multiplying by the beam's cross-sectional area at the cathode πR^2:

$$I_{\text{sat,2D}} = C_c I_0 \sqrt{\frac{2}{9}} \left(\frac{e E_{acc} R}{m c^2} \right)^{\frac{3}{2}}, \tag{1.8}$$

where $I_0 = \frac{4\pi\epsilon_0 m c^3}{e} \approx 17$ kA denotes the characteristic Alfven current, and C_c is an order-of-unity constant added for calibration. This formula indicates how much charge can be effectively extracted within a pulse of a specified length for a given spot size on the cathode before virtual cathode instabilities occur [40]. For example, inspecting this expression's scalings, it can be found that a 10 ps current pulse with emittance below 10 nm and pC levels of charge can be achieved without saturation, provided the extraction fields exceed 15 MV/m.

Alternatively, if the laser pulse length is short so that the initial photoemitted beam has a pancake-like aspect ratio, then the maximum charge that can be extracted from the photocathode, assuming a thin disk, is $Q_{sat} = \epsilon_0 E_{acc} \pi R^2$. In this regime, the electron pulse undergoes a strong longitudinal space charge expansion [41]. Extensive simulations and measurements are required to predict its pulse length, which is strongly charge-dependent. For more control over the electron pulse duration, it may be preferable to initialize the beam in the cigar mode of operation.

1.2.6 Relativistic energies for electron scattering instrumentation

Once the beam has been photoemitted, it is critically important to preserve its phase space density throughout the transport. We will see that non-linearities in external fields and in the beam self-fields are the main contributors to emittance growth. While the non-

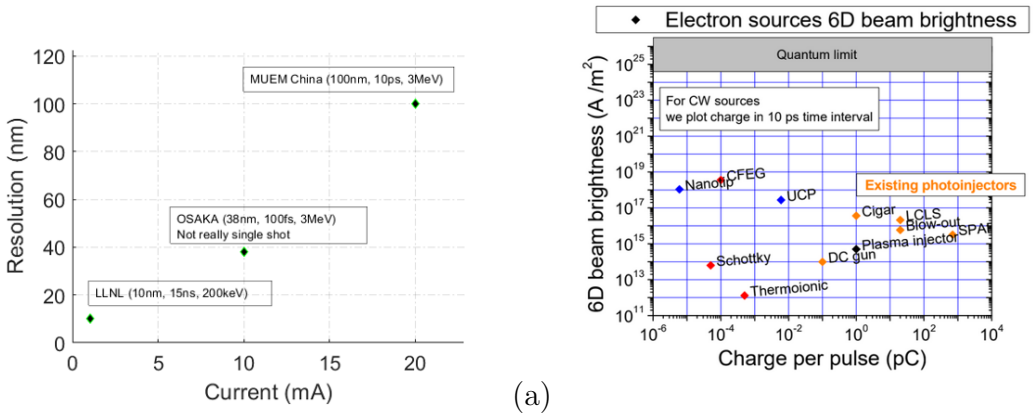


Figure 1.4: (a) Resolution versus current for different instruments. Insets label the beam energy and pulse duration used in each instrument. The data points represent: MUEM China (100 nm resolution, 10 ps pulse duration, 3 MeV energy), OSAKA (38 nm resolution, 100 fs pulse duration, 3 MeV energy), and LLNL (10 nm resolution, 15 ns pulse duration, 200 keV energy). (b) Comparison of 6D beam brightness versus charge per pulse for various electron sources.

linearities can be controlled by properly designing the electron optics, since a high signal-to-noise ratio is needed to improve the quality of the data, there is no simple solutions to the space charge effects. Figure 1.4(a) shows the state-of-the-art resolution versus current, with insets labeling beam energy and pulse duration used in the instrument. In single-shot TEM, if electrons did not repel each other, sub-nanometer resolution could be attained. However, due to the current levels in use, the resolution degrades to tens or even hundreds of nanometers. This illustrates that space charge aberrations far exceed lens aberrations, prompting the question: how can we efficiently manage space charge effects in magnification stages?

Relativistic effects suppress space charge forces.In fact, by Gauss's law, the electric field will be proportional to the total charge density of the beam. An inverse square dependence on the relativistic gamma factor arises because the beam's collective magnetic field, $\mathbf{B}_{sc}$, is related to the collective electric field, $\mathbf{E}_{sc}$ according to $\mathbf{B}_{sc} = \mathbf{v} \times \mathbf{E}_{sc}/c^2$, where $\mathbf{v}$ is the velocity of the beam, perpendicular to $\mathbf{E}_{sc}$. Calculating the transverse Lorentz force, it can be seen that the electric and magnetic fields cancel at higher beam energy, where explicitly $F_{sc} \propto \frac{Q}{\gamma^2 \sigma_x \sigma_y \sigma_z}$. The scalings on the transverse force reveal that as we compress the beams longitudinal dimension σ_z, for a fixed amount of charge Q,

the space charge forces increase. Still, the inverse quadratic scaling with beam energy strongly favors increasing beam energy to bring the magnitude of the force to acceptable levels for imaging.

RF photoinjectors are effective at generating high energy, pulsed, bright beams with low emittance, addressing both space charge and brightness requirements with their high accelerating gradients at the cathode surface ($E_0 > 50\,\mathrm{MV/m}$). They allow for variable pulse lengths ranging from fs to ps durations, as well as variable charge ranging from $10^5 e$ to $10^{10} e$. Adjusting the laser striking the cathode can tailor the initial electron pulse distribution.

Consider for example a single-shot microscope application, with a current of approximately 25 mA, an illumination radius of about 1 µm, an angular divergence of around 3 mrad, a beam gamma factor of 10, and a relative energy spread of 10^{-4}, this yields a 6D brightness of approximately $10^{17}\,\mathrm{A/m^2}$. Fig. 1.4(b) shows how cigar beams generated by recent photoinjectors nearly meet this requirement. The requirement is an estimate, and tradeoffs can be made between current and beam quality. Ongoing developments in enhanced photocathode materials [42] and optimized RF geometries are making significant strides. These advancements are aimed at improving both extraction emittance and brightness preservation.

In order to motivate this thesis, we consider the design of a single-shot UTEM based on an RF photoinjector source in [43]. This dissertation work began with the goal of experimentally demonstrating the capability of that concept source. The scheme employs a photoinjector that operates in the cigar regime, producing 10 ps long shots with an emittance of 20 nm. An X-band frequency RF cavity suppresses the relative energy spread of the beam to below 10 parts per million. This configuration ensures a high-quality electron beam suitable for high-resolution imaging.

After the sample, the design employs a permanent magnet-based quadrupole objective lens with an effective focal length of 1.5 cm. The optical system is based on quadrupole lenses, which provide strong focusing capabilities. However, the resolution from such systems is not easily tractable analytically and requires detailed simulations to accurately predict performance. These simulations, spanning from start to end, have demonstrated that the machine can achieve single-shot imaging with a spatial resolution of 10 nm.

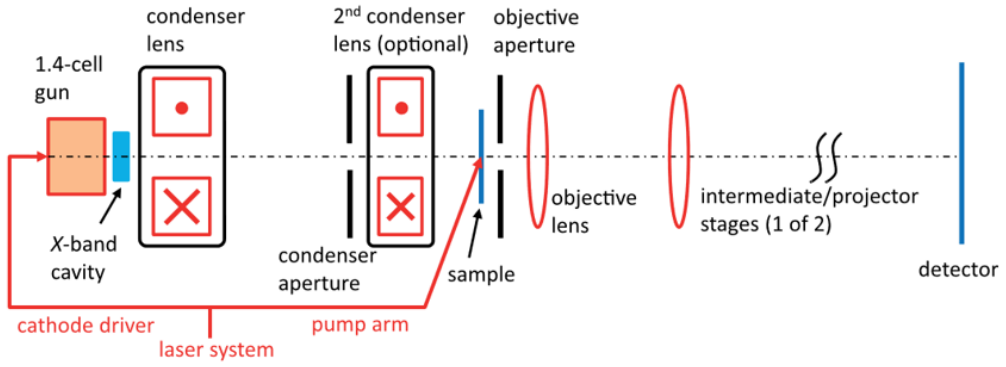


Figure 1.5: Conceptual design of a Single Shot 10 ps UTEM. Figure taken from Ref. [43]

However, despite the promising results, the initial study was primarily a design report. It highlighted the system's capabilities but did not provide a generalized framework for other research groups to easily determine the spatial resolution their systems could achieve. This gap underscores the need for further work to develop analytical models and methodologies that can be universally applied, enabling broader adoption and adaptation of single-shot TEM technology for various research applications. A framework for analytically understanding space charge aberrations in imaging is a core element of a later chapter of this thesis.

Although solenoids can provide necessary focusing at few MeV energies, they are not compact or economical because of their poor inverse square scaling with beam energy. An economic alternative is alternating gradient focusing using quadrupole magnets, whose focusing capability scales inversely with the beam energy. Recent advancements in the laboratory have led to the development of permanent magnet quadrupole (PMQ) optics analogous to those in [43]. These efforts resulted in the demonstration of single-shot picosecond time-resolved MeV electron with PMQ optics that result in a magnification factor over 30.

MeV-based time-resolved electron microscopy has a lot of potential. This method can achieve 10 nm scale resolution with 10 ps temporal resolution when combined with proper phase space shaping. To realize this potential, it will be essential to develop advanced engineering solutions and machine learning-driven feedback control systems to manage the increased complexity of these instruments. High-energy UTEM systems with permanent magnet-based focusing elements, if properly aligned and optimized, could pave

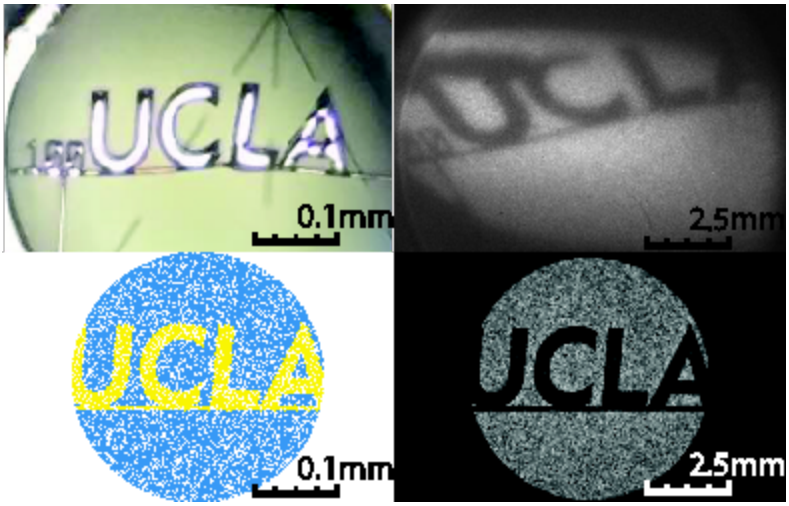


Figure 1.6: Optical and electron images obtained using PMQ optics for magnification. The top images show a nanofabricated UCLA target, while the bottom images display simulated distributions at the target and image plane. The simulation color coding indicates the division between scattered and unscattered particles. Figure taken from Ref. [45]

the way for breakthroughs in ultrafast dynamics and material science. Using machine learning-driven feedback control systems could be particularly beneficial in managing these complexities [44].

1.3 Dissertation organization

The organization of this dissertation is inspired by the MeV TEM concept. Subsequent chapters establish an analytical framework that addresses the limitations inherent in these electron scattering instruments. This framework is then applied to further develop ultrafast nanoprobes.

Chapter 2: explores both the linearized theory charged particle beams and the impact of non-linearities on emittance, focusing on the root-mean-square (rms) beam emittance and its preservation. The chapter concludes by applying these concepts to a concept for a 200 keV Dynamic Transmission Electron Microscope (DTEM) source.

Chapter 3: the first half details longitudinal beam dynamics in radio frequency (RF) photoinjectors and cavities. Concepts of bunch compression and energy spread minimiza-

tion are detailed, and a 30 keV EELS beamline, leveraging 2 RF cavities for energy spread minimization, is presented. The second half is dedicated to the commissioning of the harmonic linearizer cavity and experimental demonstration of energy spread minimization on the Pegasus beamline, along with a simulation-based design study of distortion-corrected velocity bunching using the harmonic linearizer.

Chapter 4: presents techniques for calculating space charge aberrations and mitigating the effects space charge in UTEM. It examines the influence of space charge fields on single-shot TEM capabilities, derives analytical estimates using Green's function perturbation method, and validates these estimates with particle tracking simulations, revealing fundamental scaling laws and trade-offs between temporal and spatial resolution.

Chapter 5: investigates Ultrafast Electron Diffraction (UED), emphasizing how MeV electron beams, particularly several MeV, enhance temporal resolution by mitigating space charge effects in real-time atomic-scale structural analysis. It addresses technical challenges of increased beam energy with a compact magnetic, optical system using Halbach quadrupole lenses, demonstrated with an 8.2 MeV electron beam on an Au sample, achieving a significant reduction in camera length and improved reciprocal space resolution through tunable magnification.

CHAPTER 2

Transport of charged particles beams

2.1 Introduction

With a central theme of developing techniques for preserving beam quality and transporting charged particles, this chapter delves into deriving the key dynamical equations and establishing a framework and intuition for low-current transport.

A review of Liouville's theorem is provided to examine concepts such as phase space density and phase space area—the smallness of which determines beam quality. We elucidate how distributions of particles evolve under Hamiltonian flow.

To develop the linearized dynamical equations, linear equations of motion, including longitudinal acceleration, are generally solved. This discussion brings us back to concepts of phase space area, revealing an equivalence between a single particle's phase trajectory and the evolution of a phase ellipse whose boundary initially contains that particle's initial conditions. This fact allows us to establish a first-order transport matrix formalism to evaluate how statistical measures of the beam distribution evolve.

We then explore how space charge forces naturally tend to disrupt the linearity of the system through collective forces. To maintain a linear self-field, the beam density must be uniform—a challenging condition to meet throughout each section of transport. This analysis also serves to demonstrate the general scaling of collective space charge fields and the main field components for a beam with a specific aspect ratio. A review of the Kapchinsky-Vladimirsky (KV) distribution, the only known distribution that maintains uniformity throughout linear transport, is provided to showcase an ideal self-consistent model for imaging. The KV distribution is a highly idealized model though so in general, we want to minimize or at least control the space charge effects of realistic beams. We generally provide examples of charge distributions, their associated fields, and their

evolution when non-linear evolution can be neglected, such as in low current scenarios where external focusing primarily dictates the trajectories. This highlights that even if the beam density is uniform at one position, it may not be at other critical positions in the transport, potentially causing substantial aberrations in the beam transport.

The use of Green's functions to analyze the effect of higher-order Taylor map coefficients is examined. This perturbative technique utility is exemplified through calculations of spherical aberration coefficients for the optimization of an electrostatic DTEM source. We demonstrate how to calculate emittance growth from higher-order Taylor coefficients in the case of spherical aberration, ultimately showing how emittance growth in imaging transport determines resolution.

2.2 Hamiltonian dynamics and Liouville's theorem

After photoemission, the dynamics of charged particles in the accelerator follow the classical relativistic Hamiltonian. The equations of motion can be derived from a relativistic lagrangian [46], which is a scalar function of the conjugate positions and velocities, and by a Legendre transformation, is related to a Hamiltonian, another scalar function parameterizing total particle energy with position and momentum. For a particle of charge q and mass m, The Hamiltonian is given by:

$$H = \sqrt{(\mathbf{P} - q\mathbf{A})^2 c^2 + m^2 c^4} + q\phi \tag{2.1}$$

where $\mathbf{P}$ is the canonical momentum related to the mechanical momentum $\mathbf{p} = \mathbf{P} - q\mathbf{A}$, q is the charge of the particle, $\mathbf{A}$ is the vector potential of the magnetic field, ϕ is the scalar potential of the electric field, c is the speed of light in a vacuum.Momentum and velocity obey Hamilton's equations, i.e.:

$$\dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{p}} \tag{2.2}$$

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{r}} \tag{2.3}$$

According to Maxwell's equations, the fields are related to their potentials by the following relations:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \tag{2.4}$$

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{2.5}$$

the mechanical momentum to satisfy the Lorentz force law :

$$\dot{\mathbf{p}} = e\mathbf{E} + e\dot{\mathbf{r}} \times \mathbf{B} \tag{2.6}$$

where $\mathbf{p}$ is the mechanical momentum and $\mathbf{r}$ is the time derivative of the position, $\mathbf{E}$ and $\mathbf{B}$, are the electric and magnetic fields respectively.

The Hamiltonian formalism's utility becomes apparent when describing how an ensemble of particles will evolve while subject to the Lorentz force. We represent the particle density within phase space as a smooth function, $f(x, p_x, y, p_y, z, p_z; t)$; the distribution function. The quantity $dQ = f dV_{6d} = f d^3r d^3p$ represents the charge in a small volume element $dV_{6d} = d^3r d^3p$ of phase space. The charge density $\rho(\mathbf{r})$ is calculated by integrating f over all momentum.

The motion of particles in phase space corresponds to an evolution of a vector field in a phase space continuum, where each particle is associated with a phase space velocity field vector $\mathbf{u} = (\dot{\mathbf{r}}, \dot{\mathbf{p}})$. This 6D vector field maps initial conditions to a later position/momentum vector, i.e., $\mathbf{r} = \mathbf{r}(\mathbf{r}_0, \mathbf{p}_0, t)$ and $\mathbf{p} = \mathbf{p}(\mathbf{r}_0, \mathbf{p}_0, t)$, and is considered a Hamiltonian flow. As time progresses, this evolves, causing an ensemble of initial phase points to flow into a new configuration. Notably, the Hamiltonian dynamics result in an incompressible flow so that the volume containing the particles remains fixed, and the particle density remains constant along the streamlines of the phase space velocity field.

Phase space flow must adhere to a 6D continuity equation when particles are not created or destroyed, with a 6D velocity field along which particles stream; explicitly the continuity equation is given by:

$$\nabla \cdot (f\mathbf{u}) + \frac{\partial f}{\partial t} = 0. \tag{2.7}$$

The divergence term can be expanded as $\nabla \cdot (f\mathbf{u}) = f\nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla f$.

Using Hamilton's equations, the divergence of the velocity field is:

$$\nabla \cdot \mathbf{u} = \frac{\partial}{\partial \mathbf{r}} \cdot \frac{\partial H}{\partial \mathbf{p}} - \frac{\partial}{\partial \mathbf{p}} \cdot \frac{\partial H}{\partial \mathbf{r}} = 0 \tag{2.8}$$

because the Hamiltonian is s smooth function, the mixed partials commute. Thus, the continuity equation can be simplified to a statement about the convective derivative of

the distribution function being zero, i.e.:

$$0 = \nabla \cdot (f\mathbf{u}) + \frac{\partial f}{\partial t} = \mathbf{u} \cdot \nabla f + \frac{\partial f}{\partial t} = \frac{df}{dt} \tag{2.9}$$

which implies that the distribution function $f(\mathbf{r}, \mathbf{p}, t) = f(\mathbf{r}_0, \mathbf{p}_0, 0)$, is constant along single particle trajectories, indicating that the phase space density is transported along characteristic particle paths in the continuum, and is inherently preserved, this is the formal perspective of Liouvilles's Theorem.

When we consider stationary distributions, i.e., distributions having no explicit dependence on time, then the evolution of the distribution function is obtained once the single particle dynamical equations are known. Determining the time evolution of the distribution function becomes finding an invertible relationship between the initial and final conditions, i.e., $\mathbf{r}_0 = \mathbf{r}_0(\mathbf{r}(t), \mathbf{p}(t))$, and $\mathbf{p}_0 = \mathbf{p}_0(\mathbf{r}(t), \mathbf{p}(t))$. Inversion is simple when an invertible square matrix relates initial and final states, this is when the hamiltonian is a quadratic function of position and momentum coordinates. This is when the equations of motion are linear. The difficulty increases when the relationship between the initial and final states is more complicated because of non-linear forces. This solution method is called the method of characteristics [47]. Thus, having knowledge of the initial distribution function and the solution to Hamilton's equations of motion fully determines the evolution of the distribution function.

Many formulate the theorem from the perspective that the differential volume elements are invariant [48, 49, 50], that is, a parcel of particles will maintain its volume as it undergoes Hamiltonian flow. To illustrate this, let's consider a simplified scenario in a 2D phase space. If we allow our particles to evolve by an infinitesimal time step δt, we observe that each particle's phase space location changes according to:

$$\begin{cases} x = x_0 + \dot{x_0}\delta t, \\ p_x = p_{x0} + \dot{p_{x0}}\delta t, \end{cases} \tag{2.10}$$

The initial phase area is given by $dx_0 dp_{x0}$. Thus, by considering the edges of a square region in phase space, we find that its edges will change as follows:

$$\begin{cases} dx = dx_0 + \frac{\partial \dot{x_0}}{\partial x_0} dx_0 \delta t, \\ dp_x = dp_{x0} + \frac{\partial \dot{p_{x0}}}{\partial p_{x0}} dp_{x0} \delta t. \end{cases} \tag{2.11}$$

The product of these quantities yields the new differential area element. To first order in δt, we have:

$$\begin{aligned} dxdp_x &= dx_0 dp_{x0}\left(1+\left(\frac{\partial \dot{x_0}}{\partial x_0}+\frac{\partial \dot{p_{x0}}}{\partial p_{x0}}\right)\delta t\right) \\ &= dx_0 dp_{x0}\left(1+\left(\frac{\partial^2 H}{\partial x_0 \partial p_{x0}}-\frac{\partial^2 H}{\partial p_{x0}\partial x_0}\right)\delta t\right) \\ &= dx_0 dp_{x0}. \end{aligned} \tag{2.12}$$

This implies that the jacobian for the transformation is 1. Generally, this same argument applies to showing the 6D differential volume element dV_{6D} constant through a Hamiltonian flow.

2.3 Paraxial approximation

A pulsed beam is characterized by a distribution of charged particles with small variations about the centroid and mean momentum. The centroid and average momentum are given by the first moments:

$$\langle \mathbf{r} \rangle = \frac{\int \mathbf{r} f(\mathbf{r},\mathbf{p}) dV_{6D}}{\int f(\mathbf{r},\mathbf{p}) dV_{6D}} \tag{2.13}$$

$$\langle \mathbf{p} \rangle = \frac{\int \mathbf{p} f(\mathbf{r},\mathbf{p}) dV_{6D}}{\int f(\mathbf{r},\mathbf{p}) dV_{6D}} \tag{2.14}$$

These first moments, characterizing the beam centroid, are ideally the design trajectory. We want to also linearize the distribution with respect to deviations from the average position and momentum of the electron beam. The deviations from the averages are given by:

$$\delta \mathbf{r} = \mathbf{r} - \langle \mathbf{r} \rangle \tag{2.15}$$

$$\delta \mathbf{p} = \mathbf{p} - \langle \mathbf{p} \rangle \tag{2.16}$$

Single particle dynamics capture the motion of the centroids, and we can ignore collective effects. The distribution can be parameterized in terms of the deviations from the centroid, which we try to match with an on-axis single particle design trajectory, i.e., $\langle \mathbf{r} \rangle = \int_0^t c\beta_c(\tau) d\tau \hat{z} = s(t)\hat{z}$, and $\langle \mathbf{p} \rangle = mc\gamma_c(s)\beta_c(s)\hat{z}$. To parameterize the entire distribution regarding the centroid and its momentum as the design trajectories, the distribution satisfies $\langle x \rangle = \langle y \rangle = 0$ and $\langle \beta_x \rangle = \langle \beta_y \rangle = \int_0^t \langle (\beta_z - \beta_c \rangle d\tau = 0$.

For a beam, the transverse velocities are typically orders of magnitude smaller than the longitudinal velocity, and the spread in longitudinal velocity components of each particle in the bunch, about the design trajectories, is also ideally small. We aim to express an arbitrary particle's position and momentum in terms of the design trajectory. So we first consider the longitudinal velocity, and write:

$$\beta = \sqrt{\beta_x^2 + \beta_y^2 + \beta_z^2} \rightarrow \beta_z = \beta\sqrt{1 - \frac{\beta_x^2 + \beta_y^2}{\beta^2}} \tag{2.17}$$

To account for a total velocity deviation with respect to the design trajectory, we let $\beta = \beta_c + \Delta\beta$, where β_c is the velocity of the design trajectory, which is only along the z-direction. We define particle angles $x' = \beta_x/\beta_c$ and $y' = \beta_y/\beta_c$. To second order in deviations, the longitudinal velocity, and variation with respect to the design trajectory are obtained:

$$\beta_z = \beta_c(1 + \Delta\beta/\beta_c)\sqrt{1 - \frac{x'^2 + y'^2}{(1 + \Delta\beta/\beta_c)^2}} \tag{2.18}$$

$$\approx \beta_c\left(1 + \Delta\beta/\beta_c - \frac{1}{2}(x'^2 + y'^2)\right) \tag{2.19}$$

$$\Rightarrow \frac{\Delta\beta}{\beta_c} \approx \frac{\Delta\beta_z}{\beta_c} + \frac{1}{2}(x'^2 + y'^2) \tag{2.20}$$

$\frac{\Delta\beta_z}{\beta_c} = z'$ is referred to as the fractional velocity difference.

The momentum components can be expressed in terms of ray angles and velocity deviations:

$$p_z = \frac{mc\beta_z}{\sqrt{1 - \beta_c^2(1 + \frac{\Delta\beta}{\beta_c})^2}} \tag{2.21}$$

$$= p_c(1 + \gamma_c^2 z' + \frac{3}{2}\beta_c^2\gamma_c^4 z'^2 - \frac{1}{2}(x'^2 + y'^2) + ...) \tag{2.22}$$

$$\Rightarrow \Delta p_z = p_c\left(\gamma_c^2 z' + \frac{3}{2}\beta_c^2\gamma_c^4 z'^2 - \frac{1}{2}(x'^2 + y'^2) + ...\right) \tag{2.23}$$

Similarly, second-order expansions for the momentum in the x and y direction are given by:

$$p_x = mc\gamma\beta_x = p_c(x' + \gamma_c^2\beta_c^2 x'z' + ...) \tag{2.24}$$

$$p_y = mc\gamma\beta_y = p_c(y' + \gamma_c^2\beta_c^2 y'z' + ...) \tag{2.25}$$

If the deviations are small enough to require only first-order terms, then to first order $p_x = p_c x'$, $p_y = p_c y'$, and $p_z - p_c = p_c\gamma_c^2 z'$. This truncation is what is known as the

paraxial approximation. Geometrically tracking particles is best understood in terms of trace variables. However, it is important to note that when x', y', or z' become large enough, the relation between trace space to phase space, when parameterized in this way, develops non-linearly.

2.3.1 Trace space emittance

In the paraxial approximation, the projected emittance can be expressed in terms of a trace space emittance:

$$\epsilon_{nx} = p_c \int_{\Omega'_x} dx dx' = p_c \epsilon_x \quad \text{and} \quad \epsilon_{ny} = p_c \int_{\Omega'_y} dy dy' = p_c \epsilon_y \tag{2.26}$$

$$\epsilon_{nz} = p_c \gamma_c^2 \int_{\Omega'_z} dz dz' = p_c \gamma_c^2 \epsilon_z \tag{2.27}$$

It must be stressed that these measures of trace space area are only constant if there is no acceleration and equal to the true normalized emittance only when the dynamics are truly paraxial. In the case of acceleration, while remaining paraxial, the trace space emittance can always be easily related to the normalized emittance, which is conserved regardless of how non-linear the forces are.

So to recap, we solve for the design trajectory's velocity $\beta_c(s)$, which determines it's momentum $p_c(s)$. Then, once these are known, we can describe off-axis/off-momentum particles, by parameterizing all subsequent particles to the arc length of the design trajectory, which is the straight path through the accelerator that is the electron emitted normally from the center of the cathode. We can already see that if we consider rays that are on-momentum, i.e., having the same energy as the design trajectory but small mismatch in initial transverse conditions, then the corrections are second order compared to the longitudinal velocity variation.

2.4 Linear transport theory of charged particle optics

A major challenge in transporting the charged particles is designing the transport system so that the forces about the centroid, including space charge, are linear, then phase space filamenting will not occur. Analyzing the general case of linear dynamics is the first step in a perturbative approach to a full solution, including higher-order perturbing forces. If

we could have, for all time, the forces about the centroid be linear, then corresponding paraxial equations for each projected phase space are:

$$\dot{p_x} = -k_x x \implies x'' + \frac{p_c'}{p_c}x' + K_x x = 0 \tag{2.28}$$

$$\dot{p_y} = -k_y y \implies y'' + \frac{p_c'}{p_c}y' + K_y y = 0 \tag{2.29}$$

$$\dot{p_z} = -k_z z \implies z'' + \frac{(p_c\gamma_c^2)'}{(p_c\gamma_c^2)}z' + K_z z = 0 \tag{2.30}$$

Here, k_x, k_y, and k_z are functions dependent on s, typically associated with the gradient of quadrupole fields or the on-axis field of solenoids, and

$$K_x = \frac{k_x}{p_c c\beta_c}, \quad K_y = \frac{k_y}{p_c c\beta_c}, \quad K_z = \frac{k_z}{(p_c\gamma_c^2)c\beta_c} \tag{2.31}$$

The specific characteristics of these functions will be discussed in detail later in this chapter, as it is a case-by-case function, looking particularly in this chapter at linear focusing in cylindrical symmetric electrostatic optics, then later into magnetic solenoids and quadrupoles; however, the discussion is meant to remain general for now.

By making scaling transformations to the equations of motion, i.e., letting $x = X/\sqrt{p_c}$, $y = Y/\sqrt{p_c}$, and $z = Z/\sqrt{p_c\gamma_c^2}$, the equations can be converted to normal form as well, in summary, the equations of motion in the linearized theory are given by:

$$X'' + \kappa_x(s)X = 0, \qquad \kappa_x(s) = \left(K_x + \frac{1}{4}\frac{(p_c')^2}{p_c^2} - \frac{1}{2}\frac{p_c''}{p_0}\right) \tag{2.32}$$

$$Y'' + \kappa_y(s)Y = 0, \qquad \kappa_y(s) = \left(K_y + \frac{1}{4}\frac{(p_c')^2}{p_c^2} - \frac{1}{2}\frac{p_c''}{p_0}\right) \tag{2.33}$$

$$Z'' + \kappa_z(s)Z = 0, \qquad \kappa_z(s) = \left(K_z + \frac{1}{4}\frac{((p_c\gamma_c^2)')^2}{(p_c\gamma_c^2)^2} - \frac{1}{2}\frac{(p_c\gamma_c^2)''}{(p_c\gamma_c^2)}\right) \tag{2.34}$$

For now, let U be a place holder for X, Y, or Z, and let κ take the place of κ_x, κ_y, or κ_z, so we need to solve:

$$U'' + \kappa U = 0 \tag{2.35}$$

in general. This second-order linear differential equation is called Hill's equation, named after George William Hill who originally studied this equation in the context of orbits, where κ becomes a periodic function. To solve Hill's equation [51], it is typical to employ the variation of parameters technique [52], and assume a form reminiscent to the harmonic oscillator, i.e.:

$$U = w\cos(\psi) \tag{2.36}$$

where w is an amplitude function and ψ is a phase function both depending on s (which for each dimension each should have a subscript x, y, or z). Substituting into Hill's equation results in:

$$(w'' - w\psi'^2 + \kappa w)\cos(\psi) - (2w'\psi' + w\psi'')\sin(\psi) = 0 \tag{2.37}$$

$$\tag{2.38}$$

$\sin(\psi)$ and $\cos(\psi)$ are linearly independent, so to have a solution the coefficients need to be zero, which implies:

$$\psi = \psi_0 + \int_{s_0}^{s} \frac{ds}{w^2}, \qquad \text{and} \qquad w'' + \kappa w = \frac{1}{w^3} \tag{2.39}$$

where ψ_0 is an initial phase and $\Delta\psi(s) = \int_{s_0}^{s} \frac{ds}{w^2}$.

So far, this discussion of the solution to the Hill's equation has been general, and we have not stipulated the dimensionality of the equation. If we want to think of U and w to have units of meters, and derivatives are inverse meters, then there is an issue with how the phase angle is interpreted, namely, the phase angle has units of inverse meters. Furthermore, in this context, $\sin(\psi)$ has units of meters. If however, U and w had units of $m^{1/2}$, then the phase angle is in radians, and the units of the trig functions are sensible. To remedy this counter-intuitive issue, we normalize by ϵ, a constant referred to as the Courant Snyder invariant which is a phase space area that has units of meters times radians, this allows us to redefine $U = \tilde{U}/\sqrt{\epsilon}$, and $w = \sigma/\sqrt{\epsilon}$, where σ, the envelope function, has units of meters, and $\tilde{U}$ is a physical state measured in meters. This formulation ensures the phase advance ψ has units of radians and corrects the trig functions. Making this substitution in the differential equation governing w, we get

$$\Delta\psi = \epsilon \int_{s_0}^{s} \frac{ds}{\sigma^2}, \qquad \text{and} \qquad \sigma'' + \kappa\sigma = \frac{\epsilon^2}{\sigma^3} \tag{2.40}$$

an envelope equation and a phase depending on the envelope evolution. A striking feature of this approach is that the equation for σ is also second order, so it also has two initial conditions. There are 3 integration constants, meaning we have an additional degree of freedom; how can we make sense of this peculiar feature? To make some headway, at any given moment, the position and angle are given by:

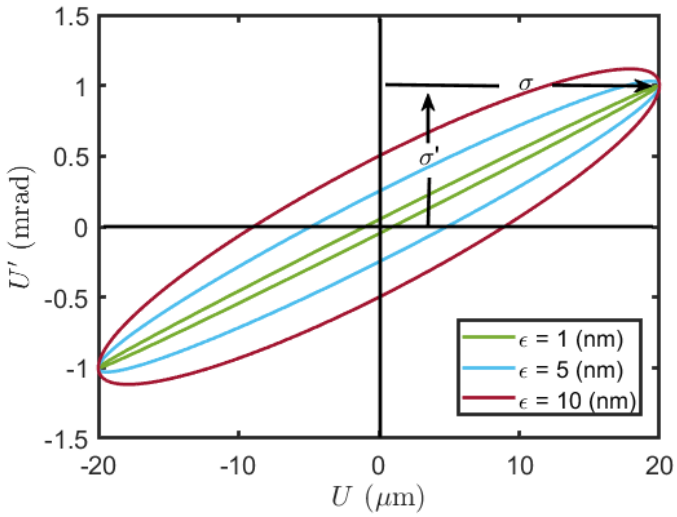


Figure 2.1: Example trace space distributions with different values of ϵ, with the same initial $\sigma = 20\mu m$ and $\sigma' = 1$ mrad.

$$\tilde{U} = \sigma \cos(\psi) \tag{2.41}$$

$$\tilde{U}' = \sigma' \cos(\psi) - \frac{\epsilon}{\sigma} \sin(\psi) \tag{2.42}$$

For a fixed set of values σ and σ', this system is a parametric system in ψ for an ellipse in the trace space, centered on the origin with an area of $\pi\epsilon$. As the initial phase is varied from 0 to 2π, a complete lap around the phase ellipse is completed. For example, Fig. 2.2 depicts different– diverging– phase ellipses. We can choose any value of ϵ, σ_0, and σ_0', if the initial conditions are on the boundary of that initial phase ellipse. As the particle undergoes its motion, its position in phase space is confined to remain on the ellipse, which maintains a fixed area while rotating and shearing in accordance with the evolution of σ. The equation for the trace ellipse is found, after eliminating the parameter ψ, to be:

$$\frac{U^2}{\sigma^2} + \frac{(\sigma' U - \sigma U')^2}{\epsilon^2} = \frac{U^2}{\epsilon w^2} + \frac{(w' U - w U')^2}{\epsilon} = 1, \tag{2.43}$$

Consequently, any particle within the region bounded by the ellipse will remain within it. This suggests we strive to match the phase space from the source to our desired phase ellipse for the application.

2.4.1 Transfer matrices

In terms of the initial conditions $\tilde{U}_0$ and $\tilde{U}_0'$, the general solution to Hill's equation is expressed as:

$$\tilde{U}(s) = \tilde{U}_0 w \left(\frac{1}{w_0} \cos\Delta\psi - w_0' \sin\Delta\psi \right) + \tilde{U}_0' w w_0 \sin\Delta\psi \tag{2.44}$$

the two functions scaling the initial conditions are linearly independent and are typically called cosine and sine-like trajectories [53, 48]. They are labeled as:

$$\begin{aligned} C(z) &= w \left(\frac{1}{w_0} \cos\Delta\psi - w_0' \sin\Delta\psi \right) \\ S(z) &= w w_0 \sin\Delta\psi \end{aligned} \tag{2.45}$$

We can construct a matrix transport from these two solutions so that we can push forward arbitrary initial conditions U_0 and U_0':

$$\begin{pmatrix} U \\ U' \end{pmatrix} = \begin{pmatrix} C(s) & S(s) \\ C'(s) & S'(s) \end{pmatrix} \begin{pmatrix} U_0 \\ U_0' \end{pmatrix} \tag{2.46}$$

The determinant of this matrix is 1, which makes sense considering if there is no acceleration, the trace emittance should not change. To convert the reduced variables back to trace space, for example, in the x-direction:

$$\begin{pmatrix} x \\ x' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{\gamma_c\beta_c(s)}} & 0 \\ -\frac{(\gamma_c\beta_c)'(s)}{\gamma_c\beta_c(s)\sqrt{\gamma_c\beta_c(s)}} & \frac{1}{\sqrt{\gamma_c\beta_c(s)}} \end{pmatrix} \begin{pmatrix} C_x(s) & S_x(s) \\ C_x'(s) & S_x'(s) \end{pmatrix} \begin{pmatrix} \sqrt{\gamma_c\beta_c(s_0)} & 0 \\ \frac{\sqrt{\gamma_c\beta_c(s_0)}(\gamma_c\beta_c)'(s_0)}{\gamma_c\beta_c(s_0)} & \sqrt{\gamma_c\beta_c(s_0)} \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix} \tag{2.47}$$

$$= R_x \begin{pmatrix} x_0 \\ x_0' \end{pmatrix} \tag{2.48}$$

Where the $\det(R_x) = p_c(s_0)/p_c(s)$. The transport, R_y, for the y-direction is the same form, but in the case of the z-direction, the momentum factors must switch from p_c to $p_c\gamma_c^2$.

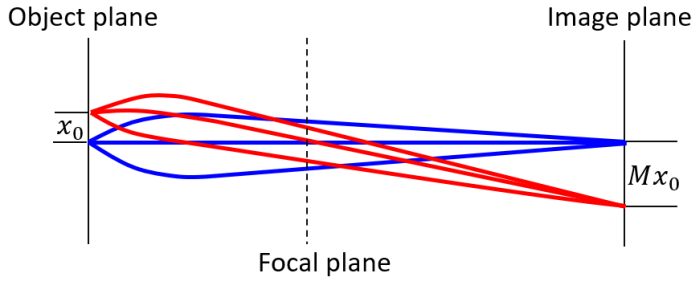


Figure 2.2: Image planes and focal planes (angles are not to scale). At image planes, bundles of rays from points in the object plane converge to points in the image plane, having a magnified offset. Rays having no angle in the object plane converge to a point in the focal plane.

For the full 6D case of uncoupled transport, we have:

$$\begin{pmatrix} x \\ x' \\ y \\ y' \\ z \\ z' \end{pmatrix} = \begin{pmatrix} R_x & 0_{2x2} & 0_{2x2} \\ 0_{2x2} & R_y & 0_{2x2} \\ 0_{2x2} & 0_{2x2} & R_z \end{pmatrix} \begin{pmatrix} x_0 \\ x'_0 \\ y_0 \\ y'_0 \\ z_0 \\ z'_0 \end{pmatrix} = R \begin{pmatrix} x_0 \\ x'_0 \\ y_0 \\ y'_0 \\ z_0 \\ z'_0 \end{pmatrix} \tag{2.49}$$

2.4.2 Imaging and momentum mapping conditions

Some key positions (or planes) occur during transport, respectively referred to as image planes and focal planes. An image plane occurs, for example in the horizontal plane, wherever the matrix element $R_{12} = 0$. When the transport meets this condition, it ensures that all final horizontal positions are correlated only with the initial horizontal positions and not on the initial angles. This condition is met whenever the phase advances by an amount $\Delta\psi_x = n\pi$, where n is a positive integer. In this case, the spatial distribution at the image plane is a magnified version of the original spatial distribution. Focal planes on the other hand, which, again considering motion in the horizontal plane, occur wherever $R_{11} = 0$, which is ensured for a transport having $w'_0 = 0$ and the phase advances by an amount $\Delta\psi_x = (2n + 1)\pi/2$. At these planes, the spatial distribution is an image of the

original momentum distribution.

2.4.3 Beam envelope equations

The second moments of the distribution concerning the deviations describe the spatial extent, correlations between phase space (or trace space) variables, and angular spread. For example in the x-direction:

$$\langle x^2 \rangle = \frac{\int x^2 f dV_6}{\int f dV_6} \tag{2.50}$$

$$\langle xx' \rangle = \frac{\int xx' f dV_6}{\int f dV_6} \tag{2.51}$$

$$\langle x'^2 \rangle = \frac{\int x'^2 f dV_6}{\int f dV_6} \tag{2.52}$$

The correlation $\langle xx' \rangle$ measures how much a beam is either converging or diverging. It can also be considered a measure of inward or outward flow in the transverse kinetic energy [46].

The root mean square (RMS) size of the beam in the x-direction is $\sigma_x = \sqrt{\langle x^2 \rangle}$, and is a commonly measured quantity. For distributions of paraxial particles, we can write an envelope equation for each respective subspace, e.g., in the x-direction, we compute the second derivative of the envelope:

$$\sigma_x'' = \frac{\langle xx'' \rangle}{\sigma_x} + \frac{\epsilon_{xr}^2}{\sigma_x^3} \tag{2.53}$$

where the quantity ϵ_{xr}^2, is the square of the RMS trace emittance is given, and is the determinant of the covariance matrix of the horizontal trace space moments:

$$\epsilon_{xr}^2 = \langle x^2 \rangle \langle x'^2 \rangle - \langle xx' \rangle^2 \tag{2.54}$$

By differentiating the RMS trace emittance with respect to s and again the equation of motion for x to simplify, we get the equation governing the evolution of the RMS trace emittance:

$$\frac{1}{\epsilon_{xr}^2} \frac{d\epsilon_{xr}^2}{ds} = -2\frac{p_c'}{p_c} \tag{2.55}$$

$$\Rightarrow \epsilon_{xr}(s) = \frac{p_c(s_0)}{p_c(s)} \epsilon_{xr}(s_0) = \frac{\epsilon_{nxr}}{p_c(s)} \tag{2.56}$$

where the normalized RMS emittance, $\epsilon_{nxr} = p_0(s_0)\epsilon_{xr}(s_0)$ is a constant. There is a critical difference between RMS emittance and actual emittance. Actual phase space

emittance is always conserved under Hamiltonian flows, but normalized RMS emittance measures the linearity of the phase space distribution, only remaining constant under linear forces. If an initially linear phase space undergoes non-linear forces, as in Fig. **??**, the shape of the distribution deforms, and subsequently, the normalized rms emittance changes, potentially increasing, so effectively, an increase in rms emittance measures the phase space filamentation.

With the RMS emittance evolution determined, we can focus on the equation of the beam's envelope. By using the equation of motion for x and the evolution of the trace emittance, we obtain the horizontal envelope equation:

$$\sigma_x'' + \frac{p_c'}{p_c}\sigma_x' + K_x\sigma_x = \frac{\epsilon_{nxr}^2}{p_c^2\sigma_x^3} \tag{2.57}$$

By rescaling $\sigma_x = \tilde{\sigma}_x/\sqrt{p_c}$, we reduce the envelope equation to normal form:

$$\tilde{\sigma}_x'' + \kappa_x\tilde{\sigma}_x = \frac{\epsilon_{nxr}^2}{\tilde{\sigma}_x^3}, \qquad \kappa_x(s) = \left(K_x + \frac{1}{4}\frac{(p_c')^2}{p_c^2} - \frac{1}{2}\frac{p_c''}{p_c}\right) \tag{2.58}$$

If p_c is constant, then the normal form is the same description as the original envelope equation, and we are free to swap to the trace RMS emittance. An identical analysis leads to the same equation for the y-direction, but in the z-direction, the trace emittance evolves instead as:

$$\epsilon_{zr} = \frac{\epsilon_{nzr}}{p_c(s)\gamma_c(s)^2} \tag{2.59}$$

which means the envelope equation in the z-direction is:

$$\sigma_z'' + \frac{(p_c\gamma_0^2)'}{(p_c\gamma_c^2)}\sigma_z' + K_z\sigma_z = \frac{\epsilon_{nzr}^2}{(p_c\gamma_c^2)^2\sigma_x^3} \tag{2.60}$$

Which has its own normal form, with $\sigma_z = \tilde{\sigma}_z/\sqrt{p_c\gamma_c^2}$:

$$\tilde{\sigma_z}'' + \kappa_z\tilde{\sigma}_z = \frac{\epsilon_{nzr}^2}{\tilde{\sigma}_z^3}, \qquad \kappa_z(s) = \left(K_z + \frac{1}{4}\frac{((p_c\gamma_c^2)')^2}{(p_c\gamma_c^2)^2} - \frac{1}{2}\frac{(p_c\gamma_c^2)''}{(p_c\gamma_c^2)}\right) \tag{2.61}$$

In this formalism, we can define a moment matrix as :

$$\Sigma = \frac{\int u^t u f dV_6}{\int f dV_6} \tag{2.62}$$

where $u = (x, x', y, y', z, z')$, u^t is the transpose of u (and is a column vector), and $u^t u$ is the dyadic product (or outer product). Given how the coordinates evolve, we can write how the beam matrix transforms:

$$\Sigma(s) = R^t\Sigma(s_0)R \tag{2.63}$$

Isolating the block corresponding to the x-direction, $\Sigma_x(s) = R_x^t \Sigma_x(s_0) R_x$, and evaluating the determinant reaffirms the notion of conserved normalized rms emittance when the forces are linear.

This analysis bridges single-particle dynamics with the envelope theory, illustrating that the same equation ultimately governs both perspectives; however, in a single-particle case, there is freedom in how the envelope is initialized. The envelope can be related to the evolution of any arbitrary elliptical region of phase space containing the initial conditions of interest on its boundary.

2.5 Perturbations of the transverse dynamics and RMS emittance growth

Up until now, we have only considered the case where the forces considered have been linear, and consequently, the rms normalized emittance has not changed. In this section, we discuss techniques for computing higher-order tensor transport to appropriately map particles through phase space when there are non-linear forces. For simplicity, we will show how the formalism works in the x-subspace, consider the beam to not be accelerating, and then discuss how to write general tensor transport.

When the force on the particle has a term that depends non-linearly on the particle position, e.g., (x^2, x^3, xy^2,...) then we often consider this to be a small force because the positions are considered close to the axis. These higher-order polynomial terms don't begin to matter until the ray is further off-axis, for instance in the case that we have quadratic dependence, we have:

$$x'' + \kappa_x(s)x = \Pi(s)x^3 \tag{2.64}$$

In this context, $\Pi(s)$ is a function of s scaled by physical constants, typically inversely with beam rigidity, which gives the strength of the perturbation. Generally, a perturbation will have a form that is a product $\Pi(s)P(x, y, z)$, where $P(x, y, z)$ is a polynomial in x, y, and potentially z having constant coefficients. Returning to the specific example where the perturbation is proportional to x^3, we apply perturbation theory to solve this problem. We assume $x = x_c + \delta x$, where x_c is the solution to the homogeneous case and

substitute into the dynamical equation collecting terms first order in δx. This process leads to:

$$\delta x'' + \kappa_x \delta x = \Pi(s) x_c^2 \left(1 + 3\frac{\delta x}{x_c}\right) \tag{2.65}$$

Often, it is acceptable to neglect the dependence of δx in the driving force, such as when $\delta x << x_c$. When this approximation is made, then we have successfully simplified the problem to a driven oscillator, which is handled with the Green's function approach, ubiquitous in solving perturbation problems in accelerator physics [48].

When the right hand side is considered just a function of s and the characteristic solution, then the solution can be written as a convolution integral:

$$\delta x(z) = \int^z \mathcal{G}_x(z,s) f(s) ds \tag{2.66}$$

where the integration interval is from the object to any position along the column up to the image plane and $G_x(z,s)$ is the Green's function of the problem which satisfies:

$$\frac{\partial^2 \mathcal{G}_x(z,s)}{\partial z^2} + \kappa_x(z)\mathcal{G}_x(z,s) = \delta(z-s) \tag{2.67}$$

Considering separately the cases when $z \neq s$ we can write:

$$\mathcal{G}_x(z,s) = \begin{cases} A_1(s)C_x(z) + A_2(s)S_x(z) & z < s \\ B_1(s)C_x(z) + B_2(s)S_x(z) & z > s \end{cases} \tag{2.68}$$

as a linear combination of the cosine-like and sine-like basis functions.

Applying the proper boundary conditions at $z = s$ (i.e. continuity of G and discontinuity for the derivative as required by integrating once around the δ-function in Eq. 2.67) we get:

$$\mathcal{G}_x(z,s) = \begin{cases} C_x(s)S_x(z) - S_x(s)C_x(z) & z > s \\ 0 & z < s \end{cases} \tag{2.69}$$

This allows us to solve for the excursion from the reference orbits

$$\delta x(z) = \int_0^z \left[C_x(s)S_x(z) - S_x(s)C_x(z)\right] f(s) ds \tag{2.70}$$

In this case, we have sub-scripted to x to indicate it corresponds to motion in the x-direction. Note that each dimension will have its own Green's function. We write

the first approximation to the solution for δx as $\delta x_{(1)}$ explicitly in terms of the Green's function as:

$$\delta x_{(1)} = \int_{s_0}^{s} \mathcal{G}_x(s,\tau)\Pi(\tau)x_c(\tau)^3 d\tau \tag{2.71}$$

Often, this is far enough, and this is what will typically be done to evaluate aberrations. However if the perturbations are large enough, then by the method of successive approximations, we can substitute $\delta x_{(1)}$ back into the equation into the RHS and iteratively re-solve the driven equation for a better result. For example, when P is only a function of x, we can write the solution as the limit of a sequence:

$$\delta x_{(n)} = \int_{s_0}^{s} \mathcal{G}_x(s,\tau)\Pi(s)P\left(x_c(\tau) + \delta x_{(n-1)}(\tau)\right) d\tau \tag{2.72}$$

Then $\delta x = \lim_{n\to\infty} \delta x_{(n)}$. Modern computing can handle this approach, which is analogous to a born-series approach to solving scattering problems in quantum mechanics.

The orbit corrections are ultimately higher-order polynomials of the initial conditions, and the collection of coefficients is often referred to as a Taylor map. To third order we may write:

$$u_i(s) = R_{ij}u_j(s_0) + T_{ijk}u_j(s_0)u_k(s_0) + U_{ijkl}u_j(s_0)u_k(s_0)u_l(s_0) \tag{2.73}$$

Where T_{ijk} and U_{ijkl} and rank 3 and 4 tensors respectively [54].

It should be noted that there are times when the transport is non-linear, but can be solved without relying on the green's function approach. For instance, there are scenarios where the function $\Pi(s)$, which serves as a placeholder in this context, can be replaced with a delta function. In that case, the higher-order effects only depend on the coordinates at the delta function's site. This is the case for longitudinal dynamics when the impulses are strong but short in time.

Once the Taylor map is obtained, it can be used to evaluate the change in RMS emittance for a given particle distribution.

2.5.1 Emittance growth in lenses

Now consider an example of an emittance growth calculation. As is often the case in TEM, a beam is focused onto a sample, so that at the sample plane, the distribution is characterized by having a small spot but large angular spread [55], this is illustrated in Fig

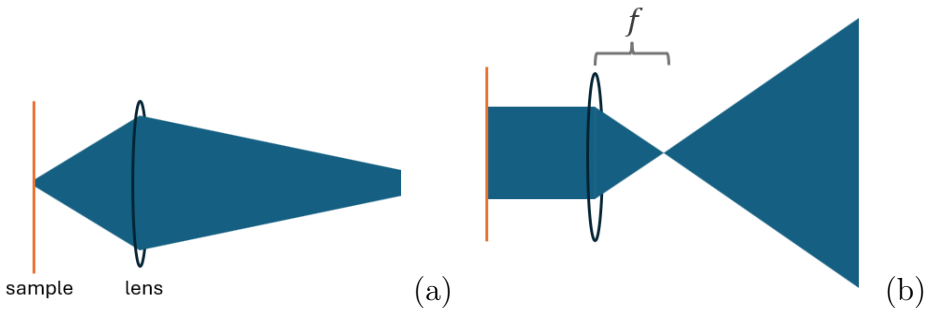


Figure 2.3: Diagram elucidating the typical illumination geometries and subsequent imaging transport. (a) From left to right, a tight focus on the sample is imaged. (b) A large collimated spot is imaged.

2.3 (a). The subsequent transport is set to image, so that, $R_{11} = -m$ the magnification, $R_{12} = 0$, $R_{21} = -1/f$ where f is the lenses focal length. The dominant aberration in this type of transport is the spherical aberration, for instance in the x-direction, the coefficient is given by U_{1222} and its derivative U_{2222} appears for the image mapping in the horizontal plane:

$$x = -mx_0 + U_{1222}x_0'^3$$
$$x' = -\frac{1}{f}x_0 - \frac{1}{m}x_0' + U_{2222}x_0'^3$$

Now we exemplify how to evaluate emittance growth for this map. If the initial distribution on the sample is an uncorrelated Gaussian with spatial and angular standard deviations σ_x and σ_θ , then the emittance changes by:

$$\Delta\epsilon_{xr}^2 = \frac{6\sigma_\theta^8 U_{1222}^2}{m^2} + \frac{3\sigma_x^2\sigma_\theta^4(U_{1222} - fmU_{2222})\left(5\sigma_\theta^2(U_{1222} - fmU_{2222}) + 2f\right)}{f^2}$$

The emittance growth from the first term can become problematic depending on how large the initial angular spread is, and typically the worse the initial emittance is, the larger the growth will be from the aberrations. U_{1222}/m is often quoted as the spherical aberration C_s. The second term can be dropped if the initial spot size is small. In that case, we have $\sqrt{\Delta\epsilon_{xr}^2}/\sigma_\theta \sim \sqrt{6}C_s\sigma_\theta^3 \sim d$, which is the resolution. Thus, when imaing from the smallest spots, the resolution of the imaging transport is directly tied to growth in emittance.

Alternatively, if a large collimated beam hits the sample, having low angular divergence, as illustrated in Fig 2.3 (b), then the dominant aberrations are actually; U_{1111} and

U_{2111}. Repeating the calculation, the strongest term is $\Delta\epsilon_{xr}^2 = 6\sigma_x^8 U_{1111}^2/f^2$. Identifying $\theta_l = \sigma_x/f$ as the convergence angle after the lens, the resolution is again related to the emittance growth $d \sim \sqrt{\Delta\epsilon_{xr}^2}/\theta_l$. Thus in these two cases, the resolution is directly related to the growth in emittance by the divergence or convergence angle depending upon the illumination geometry.

2.6 Space charge fields

We model the beam as a system of particles, treating their positions and momentum as a continuous 6D distribution. Charged particle beams are nonneutral plasmas, exhibiting collective behaviors, such as instabilities and electromagnetic wave propagation. Particularly dense beams also experience intrabeam scattering. Any local disturbance in the beam's equilibrium charge distribution is attenuated over a distance equivalent to the Debye length, which screens the electrostatic potential of individual test charges [46].

This shielding distance, or Debye length, is crucial for assessing collisional effects, especially under conditions of sharp focus caused by strong magnetic lenses. The Debye length, λ_d, is defined as the ratio of the RMS transverse velocity to the plasma frequency of the beam. If the Debye length being significantly smaller than the beam's dimensions and much larger than the interparticle spacing, i.e., $\rho^{-1/3} \ll \lambda_d \ll A$, where ρ is the beam density, and A represents a length measure of the beams. Under such conditions, we can use the mean field approximation, where a particle's interaction with its nearest neighbors can be effectively disregarded in favor of the collective field generated by the entire beam. The Debye length satisfies the necessary limits, provided the beam is not focused to exceptionally tight spots smaller than a Debye length.

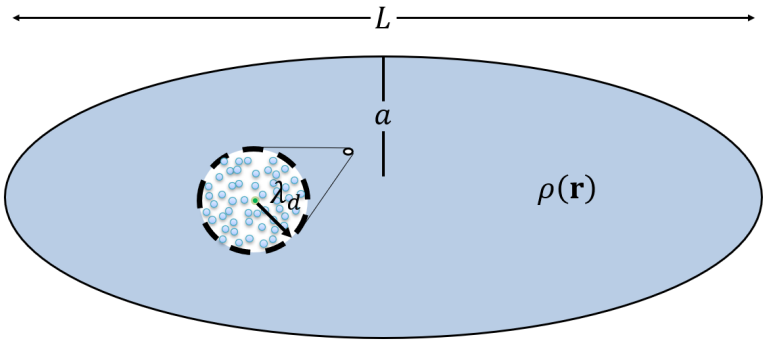
The electric and magnetic fields are determined self-consistently by Maxwell's equations:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.74}$$

$$\nabla \cdot \mathbf{B} = 0. \tag{2.75}$$

$$\nabla \times \mathbf{B} = \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} + \mu_0 \int \dot{\mathbf{r}} f d^3p \tag{2.76}$$

$$\nabla \cdot \mathbf{E} = \epsilon_0 \int f d^3p \tag{2.77}$$



(a)

Figure 2.4: A cartoon of debye shielding (not to scale), where a test charge within the beam effectively experiences a smooth field from a smooth charge distribution, and the effects of local density fluctuations are negligible.

Where the electric and magnetic fields can be separated into an applied and self-consistent field, i.e., $\mathbf{E} = \mathbf{E}_s + \mathbf{E}_{app}$, and $\mathbf{B} = \mathbf{B}_s + \mathbf{B}_{app}$.

In our previous discussion on beam evolution, we focused solely on linear forces, making a significant assumption by considering the forces to be linear. However, the only charge distribution that produces linear forces in all three dimensions is a uniformly filled ellipsoid. Practically, creating and maintaining such a distribution is impossible due to the non-zero initial temperature of the beam. Although mathematical solutions exist for this ideal case, they are not realistic because they typically assume zero longitudinal emittance, as discussed in [56].

Here, we are primarily concerned with how space charge fields depend on the beam's aspect ratio, how the beam distribution evolves according to linear transport, and when the beam can be approximated as two-dimensional. The method of characteristics is used to evaluate the charge density, and working primarily within cylindrical symmetry, an-off axis expansion for the space charge forces is presented. Then, a review of the Kapchinskij-Vladimirskij (KV) distribution is provided, highlighting its characteristics and the KV phase space as a mathematical idealization of the ideal 2D distribution. We then examine the evolution of two more practical distributions, namely, uniform ellipsoid and gaussian ellipsoid charge distributions, depicted in Fig. 2.5, showing how they naturally evolve to exhibit non-linearity and determining when they can still be approximated well enough with two-dimensional fields.

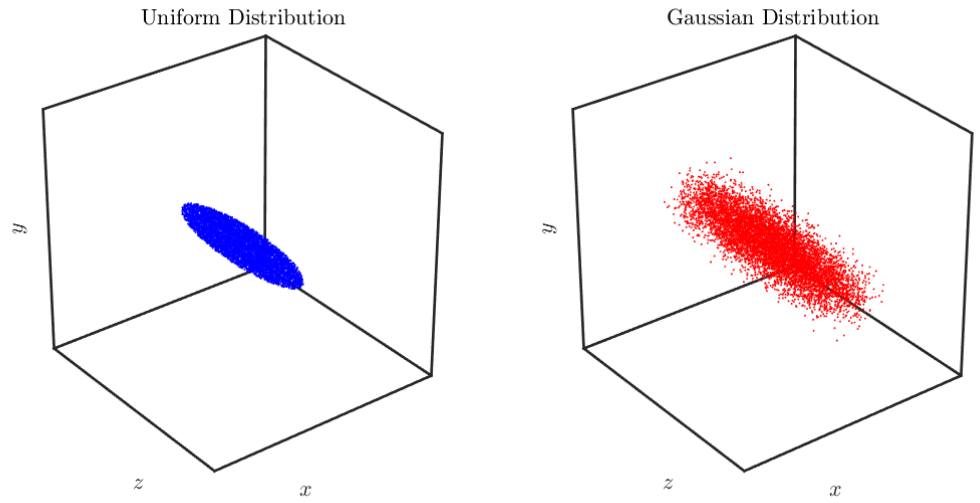


Figure 2.5: Visual comparison of 3D ellipsoidal particle distributions: a uniform distribution (left) and a Gaussian distribution (right).

2.6.1 Linear evolution of space charge fields

Assuming linear uncoupled dynamics, after making the appropriate change of variables, the charge density at any given position is given by:

$$\rho(x,y,z) = \int f(x,p_x,y,p_y,z,p_z)dp_x dp_y dp_z \tag{2.78}$$

$$= p_c(s)^3\gamma_c(s)^2 \int f(x,x',y,y',z,z')dx'dy'dz' \tag{2.79}$$

$$= p_c(s_i)^3\gamma_c(s_i)^2 \int \frac{f\left(\frac{x-R_{12}u}{R_{11}},u,\frac{y-R_{34}v}{R_{33}},v,\frac{z-R_{56}w}{R_{55}},w\right)}{R_{11}R_{33}R_{55}}dudvdw \tag{2.80}$$

[1] For a general cylindrically symmetric distribution, the charge density can be expanded in an even-powered Maclauren series:

$$\rho(r;s) = \sum_{n=0}^{\infty} \rho^{(2n)}(0;s)\frac{r^{2n}}{(2n)!} \tag{2.81}$$

where the sum only runs through even indices as required by the symmetry and $\rho^{(2n)}$ is the $2n$ radial derivative of the space-charge evaluated on the optical axis. The explicitly indicated dependence on s is due to the evolution along the beamline. In that case,

[1]The final integral is obtained by taking advantage of the distribution's stationary nature, changing the integration variables to the initial momenta, and factoring out the transport's determinant. R_{ij} are the transport matrix elements, and u,v, and w are the initial momenta.

gauss's law can be applied to obtain the radial field:

$$E_r(r;s) = \frac{1}{r}\int_0^r \frac{\rho(\xi;s)}{\epsilon_0}\xi d\xi \tag{2.82}$$

Substituting the charge density and performing the integral yields:

$$E_r(r;s) = \rho^{(0)}(s)\frac{r}{2\epsilon_0} + \rho^{(2)}(s)\frac{r^3}{8\epsilon_0} + \mathcal{O}(r^5) \tag{2.83}$$

So, we must strive to preserve the beam's uniformity to mitigate the non-linear evolution of the charge density or keep the density sufficiently low wherever the beam is not uniform so that the space charge field can be neglected. The space charge field will be strongest at beam waists, characterized by positions where $\sigma_x' = 0$, near the focus of a lens, so we may generally want to preshape a distribution so that in transport, the beam becomes uniform at waists.

2.6.2 Space charge field example distributions

Now, we analyze the well-known self-consistent Kapchinskij-Vladimirskij (KV) solution to the Vlasov equation, which describes the evolution of a continuous charged particle beam. This ideal case serves as a model that we aim to closely mimic with 3D beams to benefit from the simplicity of a linear field. However, real beams prepared in the laboratory often exhibit a range of aspect ratios, from unity to very large, due to the laser's finite duration and the use of RF cavities to bunch the beam.

2.6.2.1 KV distribution

The Kapchinskij-Vladimirskij (K-V) distribution represents a self-consistent function for an axially continuous beam with zero longitudinal emittance, $\epsilon_z = 0$ [57]. Here we review its properties. It uniquely exhibits a linear self-consistent space charge force.

Explicitly, the KV distribution function is written as:

$$f = \frac{\lambda_0}{\pi^2\epsilon_{xr}\epsilon_{yr}}\delta\left(1 - \frac{A_x^2}{\epsilon_{xr}} - \frac{A_y^2}{\epsilon_{yr}}\right)\delta\left(z'\right) \tag{2.84}$$

Where the first delta function's argument depends only on the invariants of Hill's equation

that were previously discussed:

$$\frac{A_x^2}{\epsilon_{xr}} = \frac{1}{\epsilon_{xr}}\left[\frac{x^2}{w_x^2} + \left(w_x' x - w_x x'\right)^2\right] \tag{2.85}$$

$$\frac{A_y^2}{\epsilon_{yr}} = \frac{1}{\epsilon_{yr}}\left[\frac{y^2}{w_y^2} + \left(w_y' y - w_y y'\right)^2\right] \tag{2.86}$$

The second delta function ensures all particles have the same longitudinal momentum. The distribution function is not dependent on z, so the distribution projection into the (z, z') subspace is an infinite line with no area. The projection into the 4-dimensional transverse phase space is a 4-dimensional ellipsoidal shell.

The projection into a transverse trace space is a uniform ellipsoidal region, which is shown by making the change of variables, $y = \sqrt{\epsilon} w_y cos(\psi_y)$, and $w_y' y - w_y y' = \sqrt{\epsilon} sin(\psi_y)$, the projection into the (x, x') trace space is:

$$f_x(x, x') = \int f dz' dy dy' \tag{2.87}$$

$$\propto \int \delta\left(1 - \frac{A_x^2}{\epsilon_{xr}} - \frac{\epsilon}{\epsilon_{yr}}\right)\frac{d\epsilon d\psi_y}{2} \tag{2.88}$$

$$\tag{2.89}$$

which evaluates to a constant as long as $1 > \frac{A_x^2}{\epsilon_{xr}}$ and defines the uniform elliptical region in the horizontal trace space that the particles occupy.

To show the density is constant throughout transport, we perform the integral over momentum by making a change of variables $w_x' x - w_x x' = \sqrt{\epsilon_{xr}}\alpha\cos(\zeta)$ and $w_y' y - w_y y' = \sqrt{\epsilon_{yr}}\alpha\sin(\zeta)$. With this substitution, the integration element transforms according to the jacobian by:

$$dx'dy' = \frac{\sqrt{\epsilon_{xr}\epsilon_{yr}}}{w_x w_y}\alpha d\alpha d\zeta, \tag{2.90}$$

then the argument of the first delta function can be simplified to $1 - \frac{A_x^2}{\epsilon_{xr}} - \frac{A_y^2}{\epsilon_{yr}} = \alpha_0^2 - \alpha^2$, where $\alpha_0^2 = 1 - x^2/w_x^2\epsilon_{xr} - y^2/w_y^2\epsilon_{yr}$, so the transverse charge density can be integrated with the following expression:

$$\rho(x, y) = \frac{\lambda_0}{\pi^2\sqrt{\epsilon_{xr}}w_x\sqrt{\epsilon_{yr}}w_y}\int_0^{2\pi}\int_0^{\infty}\delta(\alpha^2 - \alpha_0^2)\alpha d\alpha d\zeta \tag{2.91}$$

$$= \frac{\lambda_0}{\pi\sqrt{\epsilon_{xr}}w_x\sqrt{\epsilon_{yr}}w_y}\int_0^{\infty}\frac{\delta(\alpha - \alpha_0)}{\alpha_0}\alpha d\alpha \tag{2.92}$$

$$= \frac{\lambda_0}{\pi\sqrt{\epsilon_{xr}}w_x\sqrt{\epsilon_{yr}}w_y}. \tag{2.93}$$

This finally shows that the density remains constant throughout transport within the elliptical region defined by the inequality $1 > x^2/w_x^2\epsilon_{xr} + y^2/w_y^2\epsilon_{yr}$.

By identifying $A = \sqrt{\epsilon_{xr}}w_x$ and $B = \sqrt{\epsilon_{yr}}w_y$ as the edges of the uniformly charged ellipse, which vary along the optical axis according to the enevelope functions w_x, and w_y, we can immediately write the potential, in the small aspect ratio limit, i.e., that $C \to \infty$ with Q/C is kept constant, the potential of the uniformly charged ellipse is:

$$\phi(x, y; s) = -\frac{3Q/C}{16\pi\epsilon_0}\int_0^\infty \frac{\frac{x^2}{A^2+t} + \frac{y^2}{B^2+t}}{\sqrt{(A^2+t)(B^2+t)}}dt \tag{2.94}$$

$$= -\frac{\lambda_0}{2\pi\epsilon_0}\left(\frac{x^2}{A(A+B)} + \frac{y^2}{B(A+B)}\right) \tag{2.95}$$

where again $\lambda_0 = 3Q/4C$. Being a quadratic potential that yields the desired linear force, we have closed the loop and have an ideal, self-consistent distribution. Therefore, overall, our goal is to shape a beam's phase space distribution to mimic a KV distribution as closely as possible to maintain the optical system's linearity.

2.6.2.2 Uniform ellipsoid beam

Ideally, we would generate a KV distribution, but in practice real beams fall short due to non-zero cathode emittance and no 3D analog to the KV distribution exists [58], having finite emittance in all dimensions.

So we consider a distribution having an initially uncorrelated Gaussian momentum space, that in real space is a uniformly filled ellipsoid spatial profile at an initial plane s_i:

$$f_e = \frac{3Q}{4\pi ABCp_c(s_i)^3\gamma_c(s_i)^2}H\left(1 - \frac{x^2}{A^2} - \frac{y^2}{B^2} - \frac{z^2}{C^2}\right)\frac{\exp\left(-\frac{x'^2}{2\sigma_{x'_0}^2} - \frac{y_0'^2}{2\sigma_{y'_0}^2} - \frac{z'^2}{2\sigma_{z'_0}^2}\right)}{(2\pi)^{3/2}\sigma_{x'_0}\sigma_{y'_0}\sigma_{z'_0}} \tag{2.96}$$

In the initial plane, the charge density in the elliptical region set by the step function is $\rho = \frac{3Q}{4\pi ABC}$. Uniform ellipsoidal beams are highly desireable, their generation, transport properties, and phsycal realization is active field of research [59, 60, 61]. State-of-the-art injectors come close to delivering beams of sufficient quality to almost be called a uniform ellipsoid [62], employing the blowout technique described in [41] that is essentially the inverse of gravitational collapse of a spheroid [63], but as we will see, they do not retain perfect uniform ellipsoidal density in transport.

To get an idea of how the charge density evolves through a linear transport, let's assume the beam and transverse transport are axially symmetric, and that the longitudinal velocity spread is sufficiently small so that the spread can be considered approaching zero, then $A = B$, $\exp\left(-\frac{z'^2}{2\sigma_{z'}^2}\right)/\sqrt{2\pi}\sigma_{z'} \to \delta(z')$, in this case, after changing the transverse variables of integration to polar coordinates, the charge density is given by:

$$\rho(r, z; s) = \int f_e dp_x dp_y dp_z \tag{2.97}$$

$$= \frac{3Q}{2\pi(R_{11}A)^2 C} \exp\left(-\frac{r^2}{2R_{12}^2\sigma_\theta^2}\right) \int_0^{\lambda_m} \lambda \exp(-\lambda^2) I_0\left(\frac{\sqrt{2}r}{R_{12}\sigma_\theta}\lambda\right) d\lambda \tag{2.98}$$

where I_0 is a modified Bessel function of the first kind, and $\lambda_m = \frac{R_{11}a}{\sqrt{2}R_{12}\sigma_\theta}\sqrt{1-\frac{z^2}{C^2}}$. At an image plane, where $R_{12} \to 0$, the charge density $\rho \to \frac{3Q}{4\pi R_{11}^2 a^2 c}$ inside the ellipse, which, as expected, is still a uniform ellipsoid but with a magnified transverse radius $R_{11}a$. In the center of the bunch, we can neglect the longitudinal dependence, and the beam is essentially a uniform cylinder of charge having line charge density$\lambda_0 = 3Q/4C$ and the field is $E_r = \lambda_0 r/2\pi\epsilon_0 A^2$. Meanwhile, at focal planes, whenever $R_{11} \to 0$, the transverse distribution transitions to a Gaussian:

$$\rho(r, z; s_f) = \frac{3Q}{4C}\left(1 - \frac{z^2}{C^2}\right) \frac{\exp\left(-\frac{r^2}{2R_{12}^2\sigma_\theta^2}\right)}{2\pi R_{12}^2\sigma_\theta^2} \tag{2.99}$$

Lastly, throughout transport, the line density remains the same, i.e., $\lambda(z) = \int_0^{2\pi}\int_0^{\infty} \rho(r, z) r dr d\theta = \lambda_0(1 - \frac{z^2}{C^2})$ because the longitudinal motion relative to the centroid is frozen[2]. So, as the beam is transported, the transverse charge density transitions between a Gaussian and an ellipsoid as it undergoes transverse focusing. However, if we did not assume the longitudinal velocities were frozen, then eventually after drifting far enough, the line density would become Gaussian as well. In the core of the beam, where $z \ll C$, the field can be obtained with Gauss Law, $E_r = \frac{\lambda_0}{2\pi\epsilon_0 r}\left(1 - e^{-\frac{r^2}{2R_{12}^2\sigma_\theta^2}}\right)$, which is a gives a non-linear space charge force.

By exploiting elliptical coordinates, we can evaluate the potential of the initial ellipsoidal distribution and reveal how the field components explicitly depend on the beam's aspect ratio, which translates well to other cylindrically symmetric distributions. The

[2]To prove all of these limiting cases the asymptotic expansions of the modified Bessel function are useful, i.e., for a large argument $I_0(z) = \frac{e^x}{\sqrt{2\pi x}}$ and $I_0(0) = 2\pi$

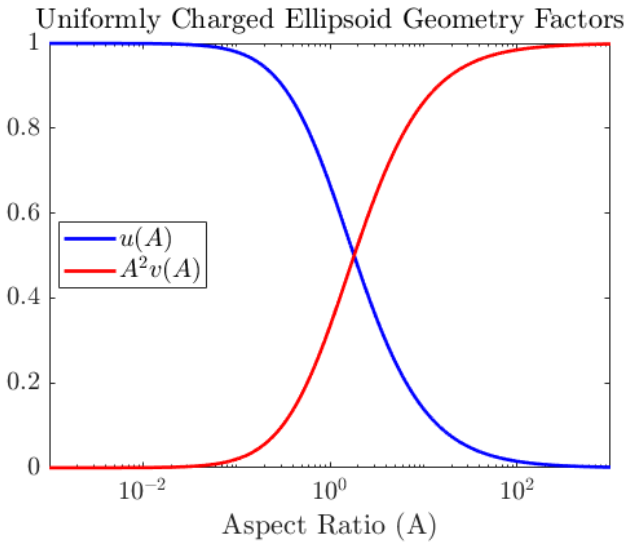


Figure 2.6: The transverse geometry factor is shown in blue and the longitudinal form factor is shown in red (scaled by $\mathcal{A}^2$). While the transverse size of the beam is finite, we consider the limiting cases of a pancake and cigar beam. As the beam becomes a cigar only the transverse field remains, and if the beam becomes a pancake, then the longitudinal field dominates.

potential of the ellipsoid can be expressed as a single integral, the mathematical technique is described in [64], we merely quote the result here:

$$\phi(x,y,z) = -\frac{3Q}{16\pi\epsilon_0}\int_0^\infty \frac{\frac{x^2}{A^2+t}+\frac{y^2}{B^2+t}+\frac{z^2}{C^2+t}}{\sqrt{(A^2+t)(B^2+t)(C^2+t)}}dt \tag{2.100}$$

Let's consider the case where initially, the beam has cylindrical symmetry (i.e., $A = B$). The electric field from this potential has a "nice" closed solution that is linear in the linear and transverse coordinates, for instance in cylindrical coordinates, r and z:

$$\mathbf{E}_u = \frac{3Qu(\mathcal{A})}{8\pi\epsilon_0 A^2 C} r\hat{\mathbf{r}} + \frac{3Qv(\mathcal{A})}{4\pi\epsilon_0 C^3} z\hat{\mathbf{z}}, \tag{2.101}$$

where $\mathcal{A} = A/C$ is the beam aspect ratio and the geometry factors are given by:

$$u(\mathcal{A}) = \frac{\xi(\mathcal{A}) - (1-\xi(\mathcal{A})^2)\coth^{-1}\left(\frac{1}{\xi(\mathcal{A})}\right)}{\xi(\mathcal{A})^3}, \tag{2.102}$$

$$v(\mathcal{A}) = \frac{\coth^{-1}\left(\frac{1}{\xi(\mathcal{A})}\right) - \xi(\mathcal{A})}{\xi(\mathcal{A})^3}, \tag{2.103}$$

and $\xi(\mathcal{A}) = \sqrt{1-\mathcal{A}^2}$.

If we always think of Q/C being finite, we can see how field components depend on the aspect ratio of the beam. In the first case, by letting $A \to 0$, for instance, by letting $C \to \infty$, the beam is in the cigar regime; in that scenario, the longitudinal field component tends to zero. Meanwhile, the transverse force remains finite as it's geometry factor approaches 1. In that case, the electric field is $\mathbf{E} \to 3Qr\hat{\mathbf{r}}/8\pi\epsilon_0 A^2 C$. Letting, $\lambda_0 = 3Q/4C$, the radial field converges to $E_r = \lambda_0 r/2\pi\epsilon_0 A^2$, again the uniformly charged cylinder.

On the other hand, if $A \to \infty$, for instance, if the beam is compressed by letting $c \to 0$, then the transverse field vanishes, and resulting longitudinal field is given by $E_z = \frac{\lambda_0}{\epsilon_0 \pi A^2} z$.

To summarize these features of the charge distribution, a plot of the relevant geometry factors are shown in Fig. 2.6. Thus, in the cigar regime, the self-fields of beams are completely independent of longitudinal coordinates, implying that any rms emittance growth is exclusively correlated with transverse coordinates.

2.6.2.3 Gaussian ellipsoid beam

Beams observed in laboratory settings consistently deviate from uniformity, displaying profiles that often look Gaussian. We extend the effective 2D model to 3D, incorporating the longitudinal field to address this. We start our analysis with the 6D uncorrelated Gaussian phase space distribution:

$$f_g = \frac{Q \exp\left(-\frac{x^2}{2\sigma_{x_0}^2} - \frac{y^2}{2\sigma_{y_0}^2} - \frac{z^2}{2\sigma_{z_0}^2} - \frac{x'^2}{2\sigma_{x_0'}^2} - \frac{y'^2}{2\sigma_{y_0'}^2} - \frac{z'^2}{2\sigma_{z_0'}^2}\right)}{(2\pi)^3 p_c(s_i)^3 \gamma_c(s_i)^2 \sigma_{x_0}\sigma_{y_0}\sigma_{z_0}\sigma_{x_0'}\sigma_{y_0'}\sigma_{z_0'}} \tag{2.104}$$

The Gaussian phase space distribution maintains a Gaussian charge density in linear transport:

$$\rho(x,y,z) = \int f_g \, dp_x \, dp_y \, dp_z = \frac{Q \exp\left(-\frac{x^2}{2\sigma_x(s)^2} - \frac{y^2}{2\sigma_y(s)^2} - \frac{z^2}{2\sigma_z(s)^2}\right)}{(2\pi)^{3/2}\sigma_x(s)\sigma_y(s)\sigma_z(s)} \tag{2.105}$$

where,

$$\sigma_x(s) = \sqrt{R_{11}^2\sigma_x^2 + R_{12}^2\sigma_{x'}^2}, \tag{2.106}$$

$$\sigma_y(s) = \sqrt{R_{33}^2\sigma_y^2 + R_{34}^2\sigma_{y'}^2}, \tag{2.107}$$

$$\sigma_z(s) = \sqrt{R_{55}^2\sigma_z^2 + R_{56}^2\sigma_{z'}^2}. \tag{2.108}$$

Although the fields of the Gaussian are non-linear, if the beam's peak current is low enough, the trace emittance should remain mostly constant. The linear components of the field becomes crucial for understanding beam evolution. Particles within one sigma of the core primarily experience this linear component. For an axially symmetric case, the linear field component can be calculated directly by Fourier transforming the charge density:

$$\phi(x,y,z) = \frac{1}{(2\pi)^3\epsilon_0}\iiint \frac{\tilde{\rho}(\mathbf{k})}{\mathbf{k}^2}\exp(i\mathbf{k}\cdot\mathbf{r})\,d^3k, \tag{2.109}$$

where $\tilde{\rho}$ is the Fourier transform of the charge density, which also has a Gaussian shape. Expanding the complex exponential in a Taylor series in $(\mathbf{k}\cdot\mathbf{r})$ and omitting the odd terms due to the symmetry of $\tilde{\rho}$, the second-order term $(\mathbf{k}\cdot\mathbf{r})^2$ yields the uncorrelated linear electric field components:

$$\mathbf{E}_g^{(1)} = \frac{Q/2\sigma_z}{(2\pi)^{3/2}\epsilon_0\sigma_x^2}u(\mathcal{A})r\hat{\boldsymbol{r}} + \frac{Q}{(2\pi)^{3/2}\epsilon_0\sigma_z^3}v(\mathcal{A})z\hat{\boldsymbol{z}}, \tag{2.110}$$

reflecting the same geometry factors as a uniformly filled ellipsoid, now with $\mathcal{A} = \sigma_x/\sigma_z$ representing the beam aspect ratio.

2.6.3 Lorentz boosting the space charge potential to the lab frame

The potential in the beam's rest frame, $\tilde{\phi}$, is a function of $\tilde{x}$, $\tilde{y}$, and $\tilde{z}$ in the beam's rest frame. The transverse coordinates are identical in each reference frame, but the temporal and longitudinal coordinates are related by a Lorentz transformation:

$$\begin{pmatrix} c\tilde{t} \\ \tilde{z} \end{pmatrix} = \begin{pmatrix} \gamma_c & -\gamma_c\beta_c \\ -\gamma_c\beta_c & \gamma_c \end{pmatrix}\begin{pmatrix} ct \\ z \end{pmatrix} \tag{2.111}$$

To obtain the potential in the lab frame, we apply the inverse Lorentz boost in the $-z$ direction:

$$\begin{pmatrix} \phi/c \\ A \end{pmatrix} = \begin{pmatrix} \gamma_c & \gamma_c\beta_c \\ \gamma_c\beta_c & \gamma_c \end{pmatrix}\begin{pmatrix} \tilde{\phi}/c \\ 0 \end{pmatrix} \tag{2.112}$$

This shows that the scalar and vector potentials in the lab frame are $\phi = \gamma_c\tilde{\phi}$ and $\mathbf{A} = A\hat{\mathbf{z}} = \gamma_c\beta_c\tilde{\phi}\hat{\mathbf{z}}/c$. The electric and magnetic fields are obtained using the relations

$\mathbf{E}_{sc} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$ and $\mathbf{B}_{sc} = \nabla \times \mathbf{A}$. The transverse components of the Lorentz force are given by:

$$\mathbf{F}_\perp = -e\nabla_\perp \phi + ec\beta_c \hat{\mathbf{z}} \times (\nabla \times A\hat{\mathbf{z}}) = -e\frac{\nabla_\perp \tilde{\phi}}{\gamma_c} \tag{2.113}$$

and the longitudinal component is given by:

$$F_z = -e\left(\frac{\partial \phi}{\partial z} + \frac{\partial A}{\partial t}\right) = -e\left(\frac{\partial \tilde{z}}{\partial z}\frac{\partial \phi}{\partial \tilde{z}} + \frac{\partial \tilde{z}}{\partial t}\frac{\partial A}{\partial \tilde{z}}\right) = -e\frac{\partial \tilde{\phi}}{\partial \tilde{z}} \tag{2.114}$$

By identifying $s = c\beta_c t$ as the position of the beam rest frame relative to the lab frame, we can re-reference $z \to z - s$, then we have $\tilde{z} = \gamma_c z$, where z is the particle position relative to the beam centroid, measured from the lab system, effectively length-contracted distances. Likewise, the longitudinal extent of the beam is length-contracted; for instance, the length of the uniform ellipsoid in the rest frame is $\tilde{C} = \gamma_c C$, where C is the contracted bunch length measured in the lab frame. The aspect ratio of the beam, appearing in the geometry factors of the potentials of uniform ellipsoids and gaussians, should be in reference to the beam's rest frame, i.e., $\mathcal{A} = A/\gamma_c C$. Note, in the transverse force, there is apparently an inverse proportionality with γ_c when we expect proportionality with γ_c^2. However, once length contraction of the charge density is accounted for, the correct proportionality is recovered.

2.6.4 Beam perveance and coupled envelope equations

We derived the space charge potential for a uniform ellipsoid. Although this distribution does not fully capture the beam dynamics due to the non-zero temperature of the beam from the cathode, it can still be used to assess the limit of bunching the beam. To use those results, we must first Lorentz transform the potential. Defining $k_x = k_{x,\text{ext}} + k_{x,sc}$ and $k_z = k_{z,\text{ext}} + k_{x,sc}$ in 2.30 to be sums of space charge and external forces, After applying the Lorentz transformation to the potential of the uniform ellipsoid, we can use the fields 2.101 to incorporate the linear term coefficients given in 2.31. For the ellipsoid, we can express the coefficients for the linear repulsive space charge force as:

$$K_{x,sc} = \frac{3Nr_c}{2\beta_c^2\gamma_c^3 A^2 C} u(\mathcal{A}) \tag{2.115}$$

$$K_{z,sc} = \frac{3Nr_c}{\beta_c^2\gamma_c^5 C^3} v(\mathcal{A}) \tag{2.116}$$

where N is the number of electrons in the beam, and $r_c = e^2/4\pi\epsilon_0 mc^2$ is the classical electron radius.

For the envelope description, we refer to the rms values. For an ellipsoidal distribution, $\sigma_z = C/\sqrt{5}$ and $\sigma_x = A/\sqrt{5}$. By switching from edges to rms values in the space charge terms, we can develop a coupled rms envelope system. We define the longitudinal perveance $K_L = 3Nr_c/5\sqrt{5}\beta_c^2\gamma_c^5$. The envelope equations then become:

$$\sigma_x'' + \frac{p_c'}{p_c}\sigma_x' + K_{x,\text{ext}}\sigma_x = \frac{\gamma_c^2 K_L}{2\sigma_x\sigma_z} u(\mathcal{A}) + \frac{\epsilon_{nxr}^2}{p_c^2\sigma_x^3} \tag{2.117}$$

$$\sigma_z'' + \frac{(p_c\gamma_0^2)'}{(p_c\gamma_c^2)}\sigma_z' + K_{z,\text{ext}}\sigma_z = \frac{K_L}{\sigma_z^2} v(\mathcal{A}) + \frac{\epsilon_{nzr}^2}{p_c^2\gamma_c^4\sigma_z^3} \tag{2.118}$$

This effectively treats the beam as a uniform ellipsoid that remains self-similar throughout the transport. Note that this is not a completely accurate kinetic model, but when assuming dominantly linear dynamics, the equations reasonably model beam envelope behavior.

When the aspect ratio is small, i.e., in the cigar regime, the longitudinal space charge force vanishes, and the transverse force no longer depends on the aspect ratio. If there is no correlation $\langle zz' \rangle$ and the energy spread is small, then the line density is essentially constant, and there is only transverse space charge expansion that is determined by the strength of the transverse beam perveance defined by:

$$K_T = \frac{\gamma_c^2 K_L}{2\sigma_z} \propto \frac{\lambda_0}{\gamma_c^3} \tag{2.119}$$

In this scenario, the transverse envelope equation is given by:

$$\sigma_x'' + \frac{p_c'}{p_c}\sigma_x' + K_{x,\text{ext}}\sigma_x = \frac{K_T}{\sigma_x} + \frac{\epsilon_{nxr}^2}{p_c^2\sigma_x^3} \tag{2.120}$$

2.7 200 keV electrostatic DTEM source

As an exemplifying application of the general linear transport theory previously described, we explore the optimization of an electrostatic source for high peak current electron bunch

dynamics in time-resolved electron microscopy (DTEM). It should be noted that "high-peak current" in the context of DTEM is the milliampere range compared to nanoamps typical of traditional TEM. We apply general paraxial theory to optimize a commercially available 200 keV JEM-2100Plus source, modified for time-resolved TEM [65]. This side-project, supported by and in collaboration with IDES.INC, a company working with JEOL, aims to evaluate performance improvements using a modified field emitter gun (FEG) geometry with a flattened cathode instead of a Mueller tip.

Although tip based field emission guns offer excellent beam quality, a modified FEG-based DTEM with a flat cathode is expected to surpass thermionic DTEMs in peak beam current, transverse emittance, and energy spread due to improved control and mitigation of space charge effects. While finer tip emitters offer excellent beam quality in time-resolved settings, they are limited to stroboscopic modes and are unsuitable for high currents required for DTEM, especially in single-shot mode. These emitters cannot withstand the high current densities and would be destroyed [66]. In collaboration with IDES, the purpose of this work was to provide a means to quantify the anticipated performance improvements through computational modeling. This model was applied to a "generic gun" (GG) model to provide preliminary performance estimates. Subsequently, IDES used these simulations to enhance a commercial FEG-based DTEM design, focusing on optimizing components such as the flat cathode, extractor, and gun lens.

DTEM requires each pulse to contain enough electrons to produce a complete image, typically at least 10^7 electrons, with improved performance up to 10^9 electrons. At 200 keV, these charges must be delivered in nanosecond pulses, resulting in peak currents on the order of milliamperes. The space charge effects, in this case, are still relatively weak. This enables the use of analytical space charge models for optimization of beam brightness specifically for DTEM applications, with GPT serving as the final test of the model. Mitigating space charge effects involves rapidly accelerating electrons to minimize their time at low speeds and high densities. This requires a large electric field across a flat cathode surface used for pulsed-laser photoemission. A FEG-type gun geometry with dedicated electrodes optimizes each function, such as extraction and beam focus, enabling higher current and spatial coherence compared to thermionic guns.

We first describe the matrix formalism for electrostatic transport systems, illustrating

how to include space charge effects; followed by a description of the GG column and simulations. The numerical model of the beamline is validated by comparing Superfish and GPT to the matrix propagator. Aberration estimates using Green's function method will be presented, followed by a discussion on space charge-induced aberrations, ultimately shown to be driving more geometrical aberrations from the gun lens.

2.7.1 Electrostatic optics formalism

The formalism used to evaluate the linear focusing properties of the GG column are described in this section. The third-order ray equation is developed, from which the linear ray equation is extracted. Reduced equations are given, and within MATLAB, the cosine and sine-like rays are obtained and subsequently used to develop the transport map as a function of the desired voltage setpoints of the column. The matrix model is based on an axially symmetric off-axis expansion.

Assuming axial symmetry greatly simplifies the problem of charged particle dynamics in electrostatic accelerators. The formalism developed in Chapter 2 proves its utility in the case where the externally applied potential can be off-axis expanded to arbitrary order in axial offset:

$$\phi_{ext}(r,z) = \sum_{n=0}^{\infty}(-1)^n \frac{V^{(2n)}(z)}{(n!)^2}\left(\frac{r}{2}\right)^{2n} \tag{2.121}$$

where $V^{(2n)}(z) = \phi^{(2n)}(0,z)$ represents even powered axial derivatives of the potential at $r = 0$. To obtain up to third order from the potential, we need only include up to 4th order in the potential:

$$\phi_{ext}(r,z) = V - V''\frac{r^2}{4} + V''''\frac{r^4}{64} + \mathcal{O}(r^6) \tag{2.122}$$

The explicit equations of motion are obtained from the relation:

$$\frac{d\mathbf{p}}{dt} = -q\nabla\phi \tag{2.123}$$

In this context, q can represent either an ion charge or an electron charge, and the same goes for m. For instance, the formalism is applied to optimize transport for an ion electrostatic column used for imaging of impact ionization of a low density gas jet presented in the appendix. Defining $x' = \beta_x/\beta_z$ and transforming the equation of motion

to paraxial coordinates, the equations of motion can be expressed in terms of trace space. The longitudinal equation of motion is:

$$(\gamma\beta_z)' = -\frac{q}{mc^2\beta_z}\frac{\partial\phi}{\partial z} \tag{2.124}$$

with the prime indicating axial derivatives. The transverse equation of motion can be parameterized in terms of axial distance, making use of the longitudinal equation, leading to:

$$x'' = \frac{q}{mc^2\gamma\beta_z^2}\left(\frac{\partial\phi}{\partial z}x' - \frac{\partial\phi}{\partial x}\right) \tag{2.125}$$

The equation can be expressed in terms of the total velocity and Lorentz factor using the fact $\beta^2 = \beta_z^2(1 + x'^2 + y'^2)$, giving us:

$$x'' = \frac{q}{mc^2\gamma\beta^2}(1 + x'^2 + y'^2)\left(\frac{\partial\phi}{\partial z}x' - \frac{\partial\phi}{\partial x}\right) \tag{2.126}$$

There is conservation of energy, so at any given position, the energy is related to the change in potential:

$$\gamma(r, z) = \gamma_0 - \frac{q\Delta\phi(r, z)}{mc^2} \tag{2.127}$$

where the contribution from collective fields has been neglected so far. However, the Lorentz-boosted potential can be included to account for space charge effects.

We can parameterize the equations of motion relative to the axial coordinate of a particle that starts on axis with no transverse momentum—the design trajectory. In that case, the change in the relativistic gamma factor is independent of r, and the beam velocity is fully determined:

$$\gamma_c(z) = \gamma_0 - \frac{q\Delta V(z)}{mc^2} \tag{2.128}$$

Then, by performing a perturbative expansion, writing a general rays factor of $1/\gamma\beta^2$ to second order in the potential, we get:

$$\frac{1}{\gamma\beta^2} = \frac{1}{\gamma_c\beta_c^2}\left(1 - \frac{q}{4mc^2}\frac{\gamma_c^2 + 1}{\gamma_c(\gamma_c^2 - 1)}V''(z)r^2 + \mathcal{O}(r^4)\right) \tag{2.129}$$

Substituting this into the transverse equation of motion, we obtain the third-order ray equation of electrostatic optics:

$$x'' = \frac{q}{mc^2\gamma_c\beta_c^2}\left(1 - \frac{q}{4mc^2}\frac{\gamma_c^2 + 1}{\gamma_c(\gamma_c^2 - 1)}(x^2 + y^2)V''\right)\left(1 + x'^2 + y'^2\right)\left(\frac{\partial\phi}{\partial z}x' - \frac{\partial\phi}{\partial x}\right) \tag{2.130}$$

In this factorized form, any terms higher than third order can be ignored after multiplying out. Higher orders enter into the paraxial equation through the terms involving partial derivatives of the potential and through the expansion of $1/\gamma\beta^2$.

Including space charge forces calculated in the lab frame up to third order, the scaling of the linear driving terms by the second-order correction to the factor $1/\gamma\beta^2$ leads to the ray equation:

$$x'' + \frac{(\gamma_c\beta_c)'}{\gamma_c\beta_c}x' = \left(\frac{qV''}{2mc^2\gamma_c\beta_c^2} + \frac{q\rho^{(0)}}{2\epsilon_0 mc^2\gamma_c^3\beta_c^2}\right)x \\ - \frac{q\rho^{(2)}}{8\epsilon_0 mc^2\gamma_c^3\beta_c^2}(x^2+y^2)^{3/2} - \left(\frac{q}{mc}\right)^2\left(\frac{\gamma_c^2+1}{8(\gamma_c^2-1)^2}\right)(x^2+y^2)\left(2V'V''x' + (V'')^2x + \frac{V''\rho^{(0)}}{\epsilon_0\gamma_c^2}x\right) \tag{2.131}$$

where the higher-order terms are collected on the second line; $\rho^{(0)}$ and $\rho^{(2)}$ are the on axis charge density and it's second derivative respectively. If we keep only terms linear in x' and x, we obtain the linear equation of motion including space charge:

$$x'' + \frac{(\gamma_c\beta_c)'}{\gamma_c\beta_c}x' = \left(\frac{qV''}{2mc^2\gamma_c\beta_c^2} + \frac{q\rho^{(0)}}{2\epsilon_0 mc^2\gamma_c^3\beta_c^2}\right)x \tag{2.132}$$

which is a damped harmonic oscillator equation, also provided in [67, 68, 69] for classical takes, albeit with the exclusion of space charge repulsion and less detail given to higher-order. By axial symmetry, the same form applies to the y-direction. The space charge terms depend on the charge density, and it's axial derivatives, however if we assume a Gaussian distribution, then the charge density evaluated on the axis is given by $\rho^{(0)} = \rho(0;s) = \frac{\lambda_0}{2\pi\sigma_x(s)^2}$, and can be substituted to obtain:

$$x'' + \frac{(\gamma_c\beta_c)'}{\gamma_c\beta_c}x' = \frac{qV''}{2mc^2\gamma_c\beta_c^2}x + K_T(s)\frac{x}{\sigma_x^2} \tag{2.133}$$

where $K_T(s) = \frac{q\lambda_0}{4\pi\epsilon_0 mc^2\gamma_c^3\beta_c^2}$ is the transverse beam perveance, which depends on the s because of the explicit dependence on the relativistic Lorentz factors varying through the electrostatic column.

After applying the transformation to normal coordinates, $x = X/\sqrt{\gamma\beta}$, and expressing all derivatives in terms of the potential, obtaining $\kappa_x = \frac{\gamma_c^2+2}{(\gamma_c\beta_c)^4}\left(\frac{qV'}{2mc^2}\right)^2 + \frac{K(s)}{\sigma_x^2}$, the reduced transverse equation of motion is given by:

$$X'' + \left[\frac{\gamma_c^2+2}{(\gamma_c\beta_c)^4}\left(\frac{qV'}{2mc^2}\right)^2 + \frac{K(s)}{\sigma_x(s)^2}\right]X = 0. \tag{2.134}$$

In this case, σ_x must be obtained from the corresponding envelope equation, self consistently. Hence, provided their axial voltage distribution, transport through electrostatic accelerators is simplified to obtaining the beam velocity from energy conservation and then solving Hill's equation for the sine-like and cosine-like trajectories S_x and C_x, to obtain the transverse evolution of particle coordinates. Once obtained, the full solution in matrix format, is obtained after applying the transformation given in 2.48.

The method to obtain the Green's function is identical to what was laid out in the previous section, just needing to use the normal form transformation. This leads to a modified Green's function which takes into account damping:

$$\mathcal{G}_x(z,s) = \sqrt{\frac{\gamma_c\beta_c(z)}{\gamma_c\beta_c(s)}}\left[C_x(s)S_x(z) - S_x(s)C_x(z)\right] \tag{2.135}$$

Thus, the general formalism for matrix transport and perturbations in electrostatic columns is established. This formalism is useful for evaluating the performance of any generic gun and can be used to optimize for desired imaging properties. Having direct access semi-analytically to the higher order terms and subsequent aberrations, and understanding how they depend on beam density and the axial potential, is very useful in understanding how to suppress them.

2.7.2 Generic gun

The GG model was designed to be representative of TEM electron guns and accelerators regardless of manufacturer. We calculate the performance using an acceleration voltage of 200 kV, as would be used in the modified JEM-2100. While TEMs can operate at reduced voltage, performance in a space-charge-limited regime notorious in single shot modes will generally perform better at higher accelerating voltages.

The model consists of: A photoemitting cathode region, initially specified as a sharp tip (as appropriate for the field emission gun design that inspired the model), but flattened for purposes of calculating the performance of a DTEM. This cathode is set at a voltage of -200 kV so that negatively charged electrons launched at the cathode will obtain a kinetic energy of 200 keV when they reach the main part of the TEM column, which is at the electrostatic ground.

An extractor electrode is also called the first anode. This is a thin electrode with a

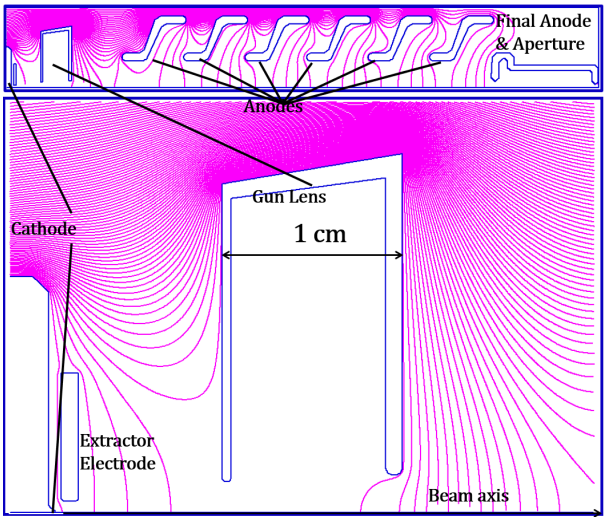


Figure 2.7: The GG column, shown as a cross-section of a cylindrically symmetrical electrostatic system. Electrons are rapidly accelerated from the cathode surface by the large electric field imparted by the extractor electrode, focused and collimated with the independent voltage control on the gun lens electrode, and accelerated up to 200 keV over the next 20 cm.

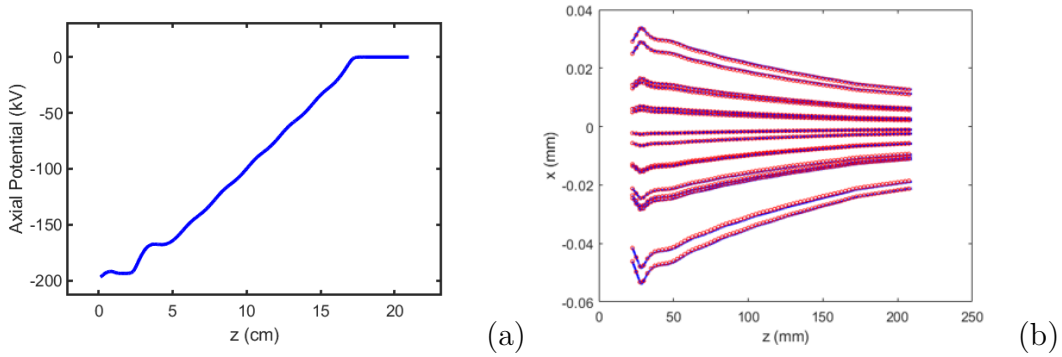


Figure 2.8: (a) The potential distribution on the axis of the GG model. (b) Particle tracking comparison with matrix solution. Blue shows the matrix solution, and the red dots show the outputs from GPT for the corresponding initial ray.

small aperture placed directly below the cathode. A variable electric potential of roughly 2-5 kV relative to the cathode is applied to A1 in order to produce a large extraction and initial acceleration field across the cathode surface.

A gun lens electrode is also called the second anode. This is a shaped piece of metal with two apertures and a variable electric potential of roughly 2-10 kV relative to the cathode. The purpose of this electrode is to focus and collimate the beam so that a very large fraction of the current ultimately can pass through the final aperture at the bottom of the accelerator.

A stack of accelerator anodes, equally spaced in voltage between the ground and the initial voltage, provides a relatively uniform accelerating field between the gun lens and the final anode.

The final anode is at zero volts and has a fixed aperture at the end. This "final aperture" is where the throughput and beam properties are supposed to be quantified. In a real microscope, this would be shortly above the first magnetostatic condenser lens. However, there was interest in utilizing the gun lens for this purpose.

Before proceeding to model space charge in the GG, electrostatic transport was validated. An axially symmetric 3-dimensional field map of the GG was generated using POISSON, with the cathode radius set to 150 μm. The geometry and corresponding equipotentials are shown in Figure 2.7.

First, the full cylindrically symmetric map was used and compared with the off-axis expanded map element using the on-axis field data. The potential on the axis is shown in Figure 2.8(a), from which the lorentz factors and all relevant derivatives are computed. The axial data was also implemented into GPT. Preliminary particle tracking studies were conducted to assess the field map's convergence by comparing the off-axis expansion map with a finely meshed full POISSON solution. Once an agreement was reached on that front, the off-axis expansion particle tracking was then to serve as a benchmark for the matrix model, and agreement was obtained. In Figure 2.8(b), individual particle trajectories obtained from the electrostatic matrix propagator (blue) are overlaid with rays (red) obtained from the GPT simulation. In this case, the point of propagation is at the midpoint of the gun lens. Subsequently, the initialization point for the matrix

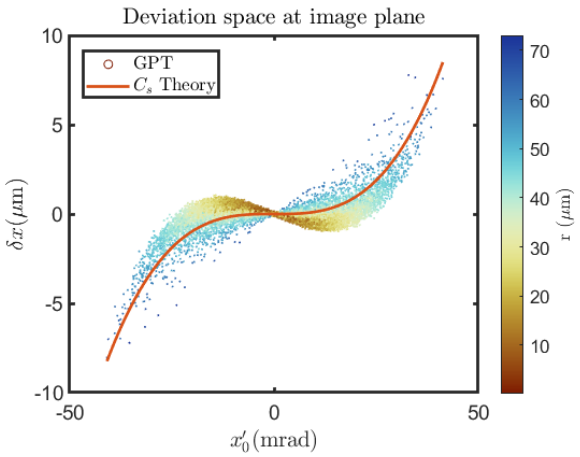


Figure 2.9: Spherical aberration from the gun lens when the image plane was set to be at the final aperture.

propagator was moved to 1.3 mm from the cathode.

Without proper consideration, particles emitted from the cathode will clip on gun lens apertures. Scanning the laser spot size revealed the transverse extent needed to fully accept the beam. The apertures and the distance from the cathode to the gun lens were adjusted to accept more charge.

GPT was used to assess particle losses on apertures for a beam having an initial MTE of 0.1eV. A smaller laser spot is favorable. Considering just MTE, for the emission from a sub 100um laser spots to be fully accepted, the gun lens aperture needed to be opened to 1.5mm and brought to within 1cm of the cathode. For a 150 μm cathode radius, it was found that the threshold laser spot needed to be about half the cathode radius to avoid geometric emittance growth from the cathode emission process.

2.7.3 Geometrical/Spherical aberration

In our study of the GG column, we employed the perturbation technique to compute aberration terms in electrostatics. The most significant perturbations were observed from the second-order correction to the factor $1/\gamma\beta^2$. Using the Green's function method and neglecting space charge, we obtain:

$$\delta x(z) = \int_0^z \frac{\gamma_c^2 + 1}{8(\gamma_c^2 - 1)^2} \left(\frac{qV''}{mc^2}\right)^2 x^3 \mathcal{G}_x \, ds \tag{2.136}$$

By applying the perturbation formalism and substituting $x = R_{11}x_0 + R_{12}x'_0$, we can construct the individual components of higher-order transport.

The large angles off the cathode cause the beam to expand to transverse sizes where chromatic effects become significant, leading to cubic aberration at the image plane (Figure 2.9), as well as emittance growth. Since the chromatic term is also scaled by the linear space charge force, increased beam current drives emittance growth from this term. Both space charge and a large mean transverse energy (MTE) have a major impact. Unlike an RF source, the acceleration off the cathode in our model is not as strong, resulting in substantial intrinsic angles until the axial velocity increases enough to suppress the beam blow-up.

By making the laser spot size smaller than the cathode's flat surface, we can mitigate far off-axis effects. However, the MTE still causes rapid beam blow-up, from a 10 μm radius to 200 μm within 2 cm. The initial angles within the beam, resulting from the MTE, approach 50 mrad, and increasing the current further exacerbates this effect. These large angles contribute significantly to higher-order distortions in the third-order paraxial equation.

2.7.4 Space charge driven aberrations

With the inclusion of space charge effects, we find that geometric aberrations are increased and largely dominate the emittance growth off the cathode as the beam passes through the gun lens. As the beam current increases to the desired levels for single-shot imaging, reaching the mA scale, the smallest achievable spot size at the aperture increases, making emittance growth a concern. The voltage of the gun lens can typically be reoptimized to ensure the waist occurs at the aperture, see Fig. 2.10(a). When space charge effects are sufficiently negligible, the best phase space at the aperture maintains the cathode emittance and achieves an RMS spot size of nearly 1 μm. However, increasing the charge to the mA level causes the spot size at the aperture to increase by a factor of 10. This increase is due to two competing effects: the sharp emittance growth caused by geometric aberrations and some space charge emittance growth. This is illustrated in Fig. 2.10(b).

Maintaining a smaller beam size throughout the electrostatic column is challenging

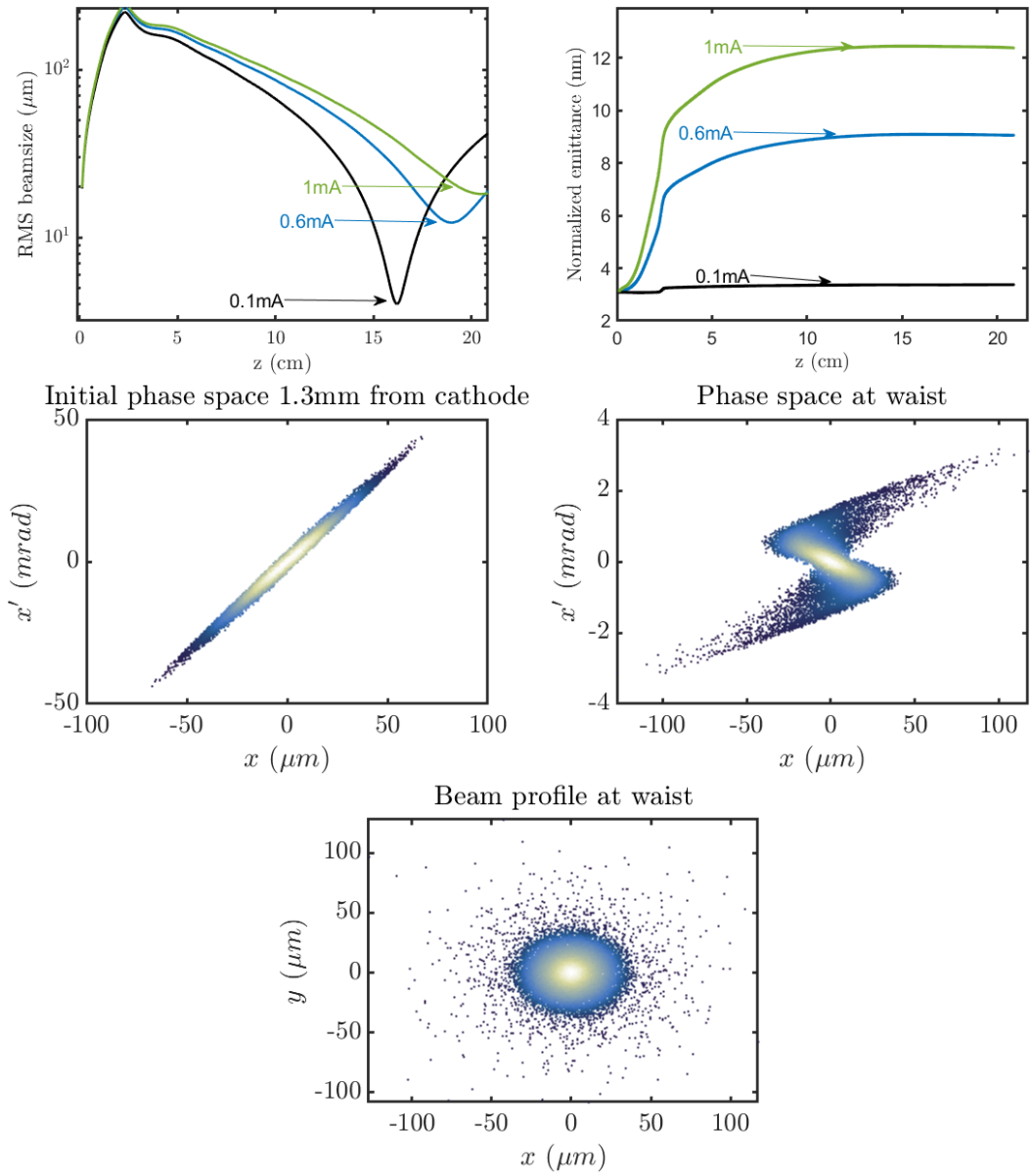


Figure 2.10: (a) Evolution of the beam envelope and (b) normalized emittance through the gun lens, demonstrating the effects of geometric and space charge aberrations. (c) Initial phase space of the beam for a current of 1 mA, roughly 1.3 mm from the cathode, showing linearity and dominance by thermal emittance. (d) Final phase space exhibiting cubic correlations due to spherical aberration from both space charge and geometric aberrations. (e) Beam profile indicating the potential for emittance improvement by clipping particles with an aperture of 2x the RMS beam size.

but necessary to ensure the emittance growth from geometric aberrations is mitigated enough so that at the exit of the source so the beam quality is suitable for imaging in subsequent magnetic objective and projection optics. However, if the beam density is maintained in a tight coasting configuration, then space charge aberrations will become more prominent. In Fig. 2.10(c), the initial phase space of the beam is shown for a current of 1 mA, roughly 1.3mm from the cathode, and is linear and dominated by the thermal emittance, with space charge only increasing the angles without adding any non-linear correlations. In Fig. 2.10(d), the final phase space exhibits cubic correlations attributed to the calculated spherical aberration of the system, which in this case is attributed to both space charge and the geometric aberrations, mainly dominated by the geometric aberrations being driven by large initial angles inherited by the MTE and further driven by space charge.

An aperture can help clean the phase space. As shown in Fig. 2.10(e), the spot size indicates that most particles contributing to the larger emittance can be clipped away with an aperture, with a radius of 2x the RMS size of the beam. This clipping improves the emittance by more than a factor of 2. This highlights how the model can aid in designing a better final aperture for future imaging applications and improved performance.

2.7.5 Future Direction

Future work can focus on further optimizing brightness by modifying the geometry of the electron gun. While the GG column represents one possible geometry, various other geometries can also be individually optimized. This includes adjusting the apertures of the gun lens, as well as the position and diameter of the extractor electrode. Additionally, modifications to the Pierce electrode geometry [70], such as varying its angles, can be explored to suppress space charge expansion off the cathode and mitigate emittance growth in the gun lens. These different geometries can be cataloged and utilized in robust multi-objective optimization schemes, as demonstrated in [71]. However, there is still significant progress to be made using semi-analytical optimizations.

At a designed beam energy, the laser can be shaped and its intensity adjusted to balance charge and temporal resolution while maintaining spatial resolution in single-shot

mode. Furthermore, implementing more complex electrode geometries may be beneficial in compensating for aberrations caused by space charge, particularly after addressing the initial emittance growth mechanism from rapid expansion into the gun lens.

CHAPTER 3

Longitudinal dynamics

3.1 Introduction

This chapter, which is divided into two parts, is a discussion of the applications of RF cavities in high-performing electron-based imaging systems.

We start with an introduction to longitudinal phase space variables and present different conventions in the definition of longitudinal emittance. Then, velocity bunching dynamics are analyzed starting with single particle dynamics. Space charge effects in the velocity bunching scheme are theoretically investigated to determine the optimal initial conditions for the process. Techniques of longitudinal emittance growth compensation in velocity bunching are elucidated, and scaling laws for optimal compensation and dependence on charge, beam energy, and RF parameters are obtained. Then, energy spread minimization techniques based on RF cavities in relativistic and sub-relativistic imaging instruments are explored. In relativistic beamlines, the source of energy spread comes from the RF fields in the gun. It is identified that with an additional short 3rd harmonic cavity the energy spread can be reduced by two orders of magnitude. The final part of the first section of the chapter is about the application of RF cavities in sub-relativistic time-resolved EELS for the purpose monochromatization. Here, we present a concept where RF cavities temporally stretch an initially short (100fs) electron pulse to ps scales, which minimizes the beam energy spread after the removal of an energy chirp [72], which is shown to be capable of improving energy resolution in EELS by an order of magnitude.

In the second part of the chapter, the Pegasus beamline is introduced, along with experimental work regarding the commissioning of the X-band linearizer and the demonstration of X-band harmonic linearization. We measured energy spreads in tens of parts per million for 10 ps long beams. Results are found to agree with start-to-end simulations

of the beamline. Although left to future experimental work, start-to-end velocity bunching simulations demonstrate the feasibility of generating bunches having 0.25pC charge approaching a fs.

3.2 Theory

3.2.1 Common alternatives in the definition of longitudinal emittance

To facilitate a description in terms of an envelope equation, longitudinal trace variables (z, z') are used, where z is the longitudinal position relative to the centroid and $z' = \Delta\beta/\beta_c$ represents the fractional velocity difference. We now review the relationship between longitudinal trace space (z, z') with the phase space (z, p_z).

Typically, the pulse duration is what is measured. The pulse duration is directly related to the variance in position, $\sigma_t = \sigma_z/c\beta_c$. Assuming the initial relative velocity distribution is uncorrelated with particle positions, we can write for the trace space emittance:

$$\epsilon_{zr} = \sigma_z \sigma_{z'} \tag{3.1}$$

To first order, a fractional velocity difference corresponds to an energy deviation:

$$z' = \frac{1}{\gamma_c^3 \beta_c^2} \Delta\gamma + \mathcal{O}(\Delta\gamma^2) \tag{3.2}$$

Likewise, a momentum deviation is related to a fractional velocity difference:

$$\frac{\Delta p_z}{mc} = \gamma_c^3 \beta_c z' + \mathcal{O}(z'^2) \tag{3.3}$$

The higher-order terms are due to the relativistic relationship between velocity and energy touched on in Chapter 2. If the differences are sufficiently large, then more terms from the Taylor expansions are needed. Therefore, for small enough deviations, there is a direct proportionality relating the spread in fractional velocity differences to energy spread or momentum spread:

$$\sigma_{z'} = \frac{1}{\gamma_c^3 \beta_c^2} \sigma_\gamma = \frac{1}{\gamma_c^3 \beta_c} \frac{\sigma_{p_z}}{mc} \tag{3.4}$$

Where σ_{p_z} and σ_γ are variances in momentum and energy respectively. Often quoted in accelerator physics is the relative energy spread denoted by $\sigma_\delta = \sigma_\gamma/\gamma$. Therefore,

for an initially uncorrelated longitudinal phase space, the normalized longitudinal RMS emittance can be expressed in terms of either momentum spread, velocity spread, or energy spread. In this case, the invariant longitudinal normalized RMS emittance can be expressed as:

$$\epsilon_{nzr} = \sigma_z \sigma_{p_z} = \gamma_c^3 \beta_c \sigma_z \sigma_{z'} = \frac{\sigma_z \sigma_\gamma}{\beta_c} \tag{3.5}$$

Better longitudinal emittance enables shorter bunch lengths or lower energy spreads after proper longitudinal phase space manipulation. If the bunch duration remains sufficiently small so that sampled RF phase intervals are essentially linear, then the entire transport theory is linear and can be analyzed using transport matrices. However, long beams sample larger phase intervals and can attain large correlated energy spread, which causes non-linear correlations.

To investigate RF bunching, we use a non-linear impulse approximation to represent the effect of RF cavities. In this model, the main effect of the RF buncher is that it changes centroid energy depending on the phase and applies a negatively sloped energy chirp to allow the tail of the beam to catch the head. RF nonlinearities and relativistic dynamics in drifts result in longitudinal trace emittance growth, which must be evaluated to effectively use an envelope description.

3.2.2 Velocity bunching

Ultrafast electron scattering requires the generation of very short electron bunches to capture the fastest physical processes [8, 73]. Due to the repulsive effect of space-charge forces, one critical challenge in this field is related to packing as many electrons as possible in a short bunch [74]. In UED, pushing the electron energy to relativistic levels has helped in minimizing the space-charge effects, concurrently bringing other advantages such as longer penetration depths, reduced group velocity mismatch, and suppressed inelastic scattering background [75, 76, 77, 78]. Over recent years, UED beamlines have seen continuous improvement in the achievable temporal resolution thanks to the introduction of techniques borrowed from accelerator physics based on the use of time-dependent radiofrequency (RF) electric field to compress the electron bunch during its propagation in the beamline [79]. RF compression using 3 GHz resonant cavities has been applied

to both non-relativistic and relativistic electron beamlines for UED [80, 81, 82], yielding bunch lengths down to the single-digit fs in the latter case.

While the discussion in this section focuses on the electron bunch length, it is important to recognize that there are many additional factors, other than the temporal duration of the probe pulse, that contribute to the actual temporal resolution limit in a specific UED setup such as temporal jitter, group velocity mismatch, laser pulse length. For example, to counteract the additional temporal jitter introduced by RF-based compression, naturally synchronized laser-generated higher frequency waves have been used to impart an energy chirp on the beam in more complex coupling structures and drive the compression dynamics [83, 84].

In any case, though, to push the boundary of the UED technique, it is critical to understand the limits in beam compression and how the various beamline parameters such as charge, energy, cavity voltage, and frequency affect the shortest bunch duration achievable. The minimum bunch length at the sample results from a complex interplay between the details of the bunching dynamics and the longitudinal space-charge forces in the beam so that typically UED practitioners have resolved to particle tracking simulation codes to design the beamline and predict the beam dynamics. The agreement with experimental results has been excellent [85]. Still, particle simulations only deal with specific beamline setups, typically lack generality, and might not offer an immediate answer to how to improve the compression in a given configuration.

It would be beneficial to have a unified formalism describing beam dynamics in RF-compression UED beamlines, covering both relativistic and non-relativistic cases while including the space charge effects. To this end, we employ the longitudinal envelope equation formalism to highlight the interplay between longitudinal emittance and space-charge forces on the pulse evolution. The single-particle dynamics presentation builds on previous works of Floettman and Zeitler [86, 87] that pointed out the role of the non-linearities in the beam compression process. The collective effects are then considered in the approximation that the beam aspect ratio remains constant along the beamline, thus decoupling the longitudinal dynamics from the transverse beam size evolution. While this is a somewhat restrictive assumption, it is experimentally relevant (the beams in UED are usually focused transversely and longitudinally at the sample) because this approach

yields an upper bound estimate for the minimum bunch length. In this case, space-charge forces are over-estimated for the situation in which the transverse spot size is kept large during the compression. Using the constant aspect ratio approximation, we can extend on the previous work and obtain analytical formulas for the minimum bunch length at the longitudinal waist that are valid in the presence of space charge. The expressions presented can then be used to guide the system optimization, compare parameter choices at different facilities, and evaluate mechanisms for further improving the bunch length.

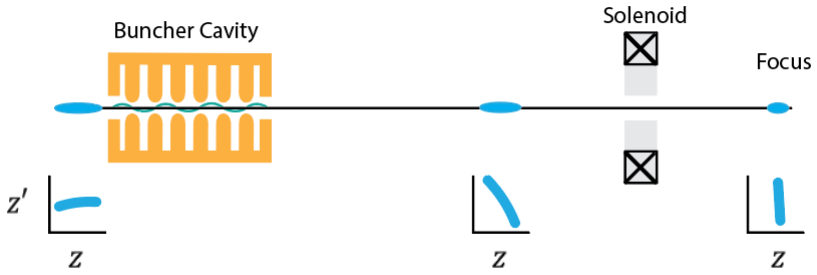


Figure 3.1: Illustration of RF ballistic bunching scheme. A velocity chirp is imparted on an electron using an RF cavity so that the tail of the beam has higher energy than the head. During the following drift the particles in the tail catch up with the particles in the head resulting in strong longitudinal compression

The simple cartoon in Fig. 3.11 illustrates the dynamics under study. Essentially, a finite electron beam propagates through an RF buncher cavity where electromagnetic fields oscillate with angular frequency $\omega = kc$ at the zero-crossing phase. Ideally, the input bunch length satisfies $k\sigma_z << 1$, so only a small phase window of the wave is sampled by the beam and the chirp imparted on the beam is predominantly linear. However, in our discussion, we will keep the higher-order terms in the energy modulation expansion to elucidate their role in the final bunch length. In the propagation region after the buncher, due to the strong energy chirp, the tail of the beam begins to catch up, while the head of the beam slows down. Finally, at some location downstream of the buncher, ideally arranged to be the sample plane or the interaction point of the UED experiment, the minimum bunch length occurs when the phase space distribution is vertically aligned.

We will strive to keep all the formulas in this section as general as possible (for example, not assuming $\beta = 1$) so that they could be applied to different RF compression setups (non-relativistic, MeV UED beamlines, as well as higher frequency compression

schemes) once the parameters are scaled accordingly. For this reason, we will use two different example cases, loosely based on the UED beamlines at the UCLA Pegasus laboratory [81], to benchmark the agreement between the analytical framework and particle tracking simulations. The reference parameters used for this study are reported in Table 4.1.

The section is structured as follows. We will first review gentle acceleration to obtain the harmonic impulse transport maps representing the cavities. Then we solve the single particle dynamics and introduce the envelope equation formalism to describe the evolution of the bunch duration through the system [88]. We approximate the buncher as a thin lens, compute the main contribution to emittance growth and analyze the ballistic dynamics in the drift while neglecting space-charge. Thus, essentially re-obtaining the results first presented in [86], then in [87]. We will then include the space-charge repulsion term in the envelope equation for an ellipsoidal beam, solving in a space-charge-dominated regime, and arrive at an expression for the minimum achievable bunch length in the beam aspect ratio. We will build on the formalism, using it to describe different current profiles. Finally, in light of the findings presented in this section, we review the use of an additional higher frequency RF cavity to compensate non-linearities to reach sub-fs bunch lengths as originally proposed by [86].

Table 3.1: Simulation beam parameters

Parameter	**High Energy**	**Low Energy**
Focal length	1.88 m	1 m
Beam kinetic energy	4.6 MeV	150 keV
Norm. transverse emittance	100 nm	8.3 nm
RMS transverse beam size	100 um	100um
Cavity Frequency	2.856 GHz	2.856 GHz
Relative energy spread	10^{-5}	10^{-5}

3.2.2.1 Gentle acceleration

Now, we consider the dynamics of charged particles in linear accelerators, where the incident particles are already moving relativistically. We derive the longitudinal impulse imparted on the beam by the cavity and examine how it induces a correlation in the longitudinal phase space at the cavity's exit.

Close to the axis, the longitudinal field within the RF cavity can generally be expressed as:

$$E_z = E(z)\cos(\omega t + \phi_0)$$

The energy gain through the entire structure, denoted as $\Delta\gamma$, where the full structure has length L, is given by:

$$\Delta\gamma = \frac{e}{mc^2}\int_{-L/2}^{L/2} E(z)\cos[\omega t(z) + \phi_0]\,dz$$

$$= \frac{e}{mc^2}\int_{-L/2}^{L/2} E(z)[\cos(\omega t)\cos\phi_0 - \sin(\omega t)\sin\phi_0]\,dz$$

By introducing the axial RF voltage, defined by $V_0 = \int_{-L/2}^{L/2} E(z)\,dz$, the energy gain can be expressed compactly as:

$$\Delta\gamma = \frac{qV_0\tau}{mc^2}\cos\phi_0$$

where the transit-time factor τ has been introduced, and explicitly is given by:

$$\tau = \frac{\int_{-L/2}^{L/2} E(z)\cos(\omega t(z))\,dz}{\int_{-L/2}^{L/2} E(z)\,dz} - \tan\phi_0\frac{\int_{-L/2}^{L/2} E(z)\sin(\omega t(z))\,dz}{\int_{-L/2}^{L/2} E(z)\,dz} \tag{3.6}$$

RF structures are ideally designed to minimize the second term, eliminating the transit time's dependence on the initial phase. Note this is only possible when $\int_{-L/2}^{L/2} E(z)\,dz \neq 0$ [89].

It is often permissible to assume the near speed of light approximation; that the velocity is effectively constant through the cavity, then $\omega t(z) = \omega z/c\beta_c$, greatly simplifying the transit time factor[1]. We define $k = \omega/c\beta_c$ to be the RF wave number divided by

[1]If we had not integrated over the full cell for the impulse, we could write the energy gain at an arbitrary axial point within the RF field.

the normalized longitudinal velocity, so we can conveniently analyze how the impulse acts on a longitudinal distribution. We make a change of $\phi_0 \rightarrow \phi_0 + \Delta\phi$, where a phase offset $\Delta\phi = kz_0$ is relative to the beam centroid arriving at ϕ_0. This enables us to write longitudinal phase space correlations at the exits of RF cavities as:

$$\Delta\gamma = \alpha \cos(kz_0 + \phi_0) \tag{3.7}$$

$\alpha = \frac{qV_0\tau}{mc^2}$ is the amplitude of the imparted energy modulation.

3.2.2.2 Single particle dynamics and non-linear phase-space correlations in the RF buncher

When the beam arrives at the zero-crossing point, i.e., when $\phi_0 = \pi/2$, the energy modulation is given by:

$$\Delta\gamma = -\alpha \sin(kz_0) \tag{3.8}$$

where $\alpha = eV_0/mc^2$ and eV_0 is the cavity voltage or maximum energy gain seen by an ideally phased particle (transit time factored in). The phase of the cavity is tuned so that the center of the bunch experiences no net energy gain and particles at the tail gain energy, while particles at the head of the bunch lose energy.

There are two distinct sources of non linearities in the trace space dynamics resulting from the applied energy change to the particles. Firstly, for finite duration input bunches, the curvature of the RF wave will cause significant non linear effects in the trace space. In addition, the relativistic relation between normalized velocity and beam energy $\beta = \sqrt{1 - 1/\gamma^2}$ adds an important degree of non linearity to the transport as pointed out in zeitler. Following the discussion therein, Taylor-expanding the relative velocity deviation $\frac{\Delta\beta}{\beta}$ in terms of the energy deviation we can write

$$\frac{\Delta\beta}{\beta} = \sum_m \eta_m \Delta\gamma^m \tag{3.9}$$

where η_m are proportional to the mth-derivatives $\frac{d^m}{d\gamma^m}\beta$ and in particular

$$\eta_1 = \frac{1}{\beta^2\gamma^3} \tag{3.10a}$$

$$\eta_2 = \frac{2 - 3\gamma^2}{2\gamma^6\beta^4} \tag{3.10b}$$

$$\eta_3 = \frac{2 - 5\gamma^2 + 4\gamma^4}{2\gamma^9\beta^6} \tag{3.10c}$$

where γ and β are the mean values of the normalized energy and velocity distributions respectively. The coefficients η_m scale as $\gamma^{-(m+2)}$ so that at high relativistic energies the higher order non linear terms in the transport can be neglected. Considering the lowest order dynamics, we can simply replace $\sin(kz_0)$ with $kz_0 - \frac{(kz_0)^3}{6}$, and truncate the series to obtain:

$$\frac{\Delta\beta}{\beta} \approx -\eta_1\alpha(kz_0) + \eta_2\alpha^2(kz_0)^2 + \left(\frac{\eta_1\alpha}{6} - \eta_3\alpha^3\right)(kz_0)^3 \tag{3.11}$$

We verify this expression at high energy (4.6 MeV) and low energy (150 keV) by considering a particle tracking simulation of the buncher configuration listed in Table 4.1 with an initial bunch length of 195 μm and 1.87 mm respectively. The buncher was modeled by a 2.856 GHz standing wave cylindrically symmetric TM010 cavity with an amplitude adjusted to reach a longitudinal focus 1.88 m and 1 m downstream respectively for the high and low energy cases. Since the RF cavity length is 0.05 m, it is reasonable to approximate it as a thin lens. The longitudinal phase spaces from GPT at the exit of the buncher are shown in Fig. 3.2(a) and 3.2(c) for the high and low energy case respectively with subtracted linear correlations. The quality of the agreement between GPT and our analytical framework can be assessed by comparing the distributions with the lines corresponding to Eq. 3.11 which are also shown. The parameters chosen for these examples highlight the different possibilities for the dominant non-linearity in the system. In the high-energy case, the relativistic effects are responsible for the parabolic shape seen in the simulation. While in the low energy case, the injected bunch length is longer and the third-order non-linearity associated with the sinusoidal RF fields is the main effect in the beam distribution shape.

The convenience of working in the trace space is the linearity of the dynamics in the drift which fully preserves the trace space area. Explicitly, in the drift after the buncher, the longitudinal particle position can be written as

$$z = z_0 + s\frac{\Delta\beta}{\beta} = z_0 + s\sum_{n=1}^{\infty}\eta_n\Delta\gamma^n \tag{3.12}$$

The initial coordinate z_0 is expressed in terms of the induced energy modulation $\Delta\gamma$ by inverting Eq. 8. Then $\Delta\gamma$ is Taylor expanded in terms of $\Delta\beta/\beta$. Substituting into Eq. 3.12, keeping only terms up to third order, we can write:

$$z \approx \left(s - \frac{1}{\eta_1\alpha k}\right)\frac{\Delta\beta}{\beta} + + \frac{\eta_2}{\eta_1^3\alpha k}\frac{\Delta\beta^2}{\beta^2} - \frac{\left(\frac{\eta_1\alpha}{6} - (\eta_3 - 2\eta_2^2/\eta_1)\alpha^3\right)}{\eta_1^4\alpha^4 k}\frac{\Delta\beta^3}{\beta^3} \tag{3.13}$$

The longitudinal waist occurs where the linear chirp is cancelled at distance $s = \frac{1}{\eta_1 \alpha k}$ along the beamline allowing us to define the buncher longitudinal focal length:

$$f = \frac{1}{\eta_1 \alpha k} = \frac{m_0 c^2 \gamma^3 \beta^2}{e V_0 k} \tag{3.14}$$

which indicates that very high voltage cavities are needed to obtain short focal lengths for relativistic electrons. It is also useful to note the k-dependence of this expression which favors the use of very high frequencies for this application. At the focal plane,

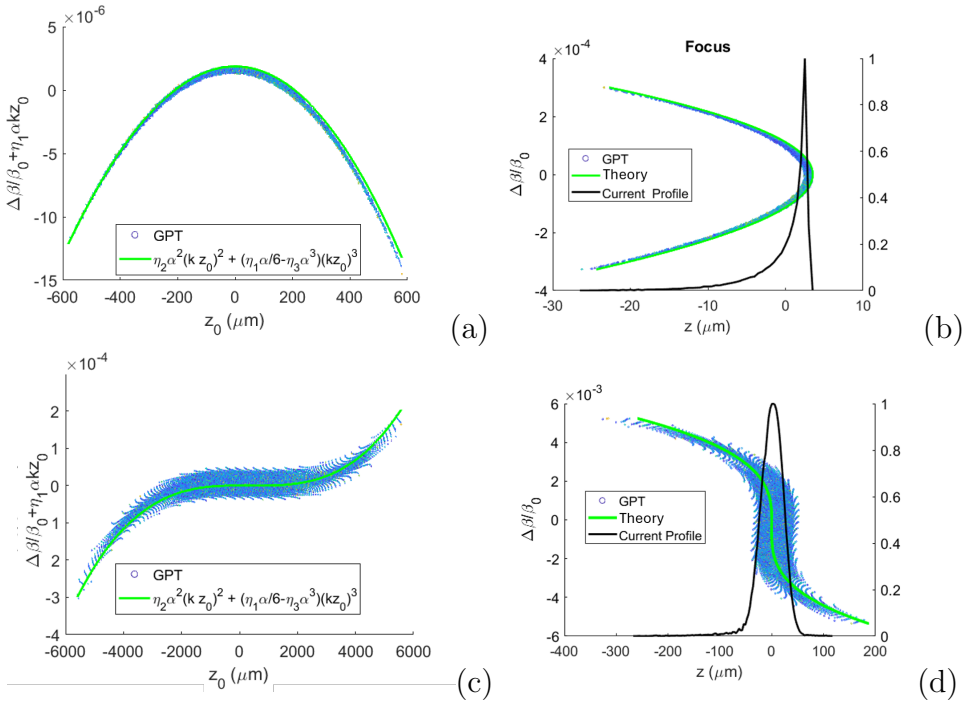


Figure 3.2: Left) Trace spaces of the beam at the exit of the prebuncher after the linear chirp has been subtracted from the distribution for high energy (top) and low energy (bottom) cases, compared with the analytical predictions from Eq. 3.11 (green curves). Right) Longitudinal trace spaces at the temporal waist for the high energy (top) and low energy (bottom) cases compared with the predictions from Eq. 3.33 (green curves). The current profiles at the focus are also shown in black.

the residual correlation is quadratic or cubic in $\Delta\beta$ depending on the relative importance of the non-linearity in the drift propagation concerning the RF curvature. As discussed above, lower beam energies and longer input bunches tend to show higher third-order non-linearities, while relativistic energies typically have dominant second-order contributions. The predictions from Eq. 3.33 can be again verified by comparing to the phase spaces

at the temporal waist plane from the same GPT simulation, as shown in Fig. 3.2(b) and (d).

3.2.2.3 Emittance growth mechanisms and the relationship between different longitudinal phase space definitions.

3.2.2.3.1 (z, z') trace space emittance

Since the drift dynamics in the trace space are entirely linear, and the emittance growth is all accrued in the buncher, the envelope equation formalism is a convenient choice to follow the RMS bunch length evolution. To evaluate the RMS emittance growth induced by the RF compressor, we start from an initial longitudinal phase space with RMS emittance $\epsilon_{z_0,z_0'}$. After the energy chirp is applied, the single-particle velocity variation maps to $z_0' \to z_0' + \Delta\beta/\beta$, where $\Delta\beta/\beta$ represents the velocity variation imparted by the buncher, which is correlated with particle position. In the thin lens approximation, the particles do not change position as the beam goes through the cavity.

The moments of the new distribution can be calculated and the relation between initial and final emittance after the buncher written as

$$\begin{aligned}\epsilon_{zz'}^2 &= \epsilon_{z_0 z_0'}^2 + \epsilon_{RF}^2 \\ &= \epsilon_{z_0 z_0'}^2 + \langle z_0^2 \rangle \langle \left(\frac{\Delta\beta}{\beta}\right)^2 \rangle - \langle z_0 \left(\frac{\Delta\beta}{\beta}\right) \rangle^2\end{aligned}$$

Note that each respective expectation value needed to calculate the emittance, with the exception of $\langle z_0^2 \rangle = \sigma_{z0}^2$, requires integrating $\sin^m(kz)$ or $z\sin^m(kz)$ over the beam distribution. Assuming an initial Gaussian current profile, these integrals have closed-form expressions up to arbitrary order of m, but keeping the leading contributions to the emittance growth, we obtain

$$\Delta\epsilon_{RF}^2 \approx \sigma_{z0}^2 \left[2\eta_2^2 \alpha_1^4 k_1^4 \sigma_{z0}^4 + \frac{1}{6}\left(\eta_1\alpha_1 - 6\eta_3\alpha_1^3\right)^2 k_1^6 \sigma_{z0}^6 \right], \tag{3.15}$$

Alternatively, assuming an initial quadratic current profile, the emittance growth is:

$$\Delta\epsilon_{RF}^2 \approx \sigma_{z0}^2 \left[\frac{8}{7}\eta_2^2 \alpha^4 k^4 \sigma_{z0}^4 + \frac{50}{1323}\left(\alpha\eta_1 - 6\alpha^3\eta_3\right)^2 k^6 \sigma_{z0}^6 \right]. \tag{3.16}$$

In most cases, these expressions are much larger than the initial longitudinal emittance because the buncher induces a large velocity spread (with correspondingly significant non-linear contributions) to achieve strong compression. As long as the space-charge forces

are negligible in the drift, ϵ_{RF} is also equal to the final emittance and can be used to calculate the shortest bunch length achievable at the waist using Eq. 3.22. The emittance growth in the buncher and its preservation in the drift is shown in Fig. 3.9, where the trace space RMS emittance evolution calculated from GPT is plotted along the beamline. The black dotted line shows Eq. 3.15 which provides a good approximation for the final emittance after a thin lens buncher cavity located at the origin.

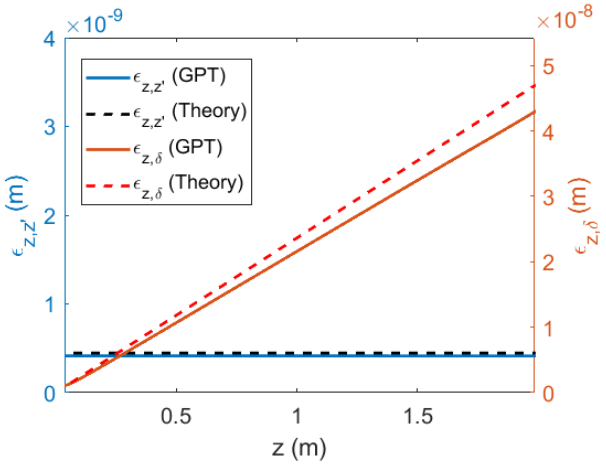


Figure 3.3: Comparison between analytical estimates (dashed lines) and GPT simulations (solid lines) for an initial bunch length of 0.65 ps. The trace space emittance after the buncher is shown in blue, and emittance growth in (z, δ) phase space in orange.

3.2.2.3.2 (z, δ) phase space It is important to note at this point that if instead, we had utilized more common choices of defining the longitudinal trace space in terms of the relative energy spread $\delta = \frac{\Delta\gamma}{\gamma}$ or momentum spread $\Delta p_z/p_z$, the drift dynamics would become highly non-linear especially for mildly relativistic particles.

Initially, the trace space emittance $\epsilon_{z,z'}$ is related to the (z, δ) emittance by the following relationship:

$$\epsilon_{z_0,\delta} = \gamma^2\beta^2\epsilon_{z_0,z_0'} \tag{3.17}$$

This relationship holds because $z' \approx \eta_1\Delta\gamma$ and explains the order of magnitude difference in absolute emittance values in Fig. 3.9. It is then fairly common to see in the literature the envelope equation written in terms of the (z, δ) emittance:

$$\sigma_z'' = \frac{\langle zz'' \rangle}{\sigma_z} + \frac{\epsilon_{z,\delta}^2}{\beta^2\gamma^2\sigma_z^3} \tag{3.18}$$

Nevertheless, in a drift, the particle positions evolve according to Eq. 3.12 where the transport is inherently non-linear, especially for mildly relativistic particles, causing emittance growth and limiting the usefulness of the envelope equation approach. Ultimately, for large enough initial energy spreads, the higher-order terms proportional to η_m lead to significant (z, δ) emittance growth, which can be estimated using the same techniques as in the previous subsection

$$\epsilon_{z\delta}^2 = \epsilon_{z_0,\delta_0}^2 + 2s^2\eta_2^2\alpha^6 k^6 \sigma_z^6 \tag{3.19}$$

where ϵ_{z_0,δ_0} is the emittance at the beginning of the drift. This expression predicts a nearly linear growth with propagation distance for a small initial emittance. This is shown in Fig. 3.9 where the (z, δ) phase space evolution from GPT is compared with Eq. 3.19 with the inclusion of the initial emittance as well.

The seemingly counterintuitive behavior of the longitudinal emittance (growing linearly in the drift) is the main reason we adopt the (z, z') trace space emittance in calculating the final bunch length when using the envelope equation formalism. Finally, for completeness, we observe that if the un-normalized momentum $(z, \Delta p_z/p_z)$ was used as a trace space variable, all expressions could be simply modified substituting $z' = \frac{1}{\gamma_0^2}\Delta p_z/p_z$. Nevertheless, due to the relativistic non-linear relation between momentum and velocity, even in this case, one would have significant emittance growth in the drift propagation.

3.2.2.4 Envelope analysis

3.2.2.4.1 Bunch limits without space charge

The simplest scenario is where propagation occurs in a drift and space-charge forces are negligible, meaning the beam remains dominated by emittance up to the waist. This situation applies to beams with very low charge.

To determine the shortest bunch length in the absence of space charge, we solve the longitudinal envelope equation independently of the transverse equation. The envelope equation can be expressed as:

$$\frac{1}{2}\frac{d}{ds}\left(\sigma_z'^2\right) = \frac{\epsilon_{z,z'}^2}{\sigma_z^3}\sigma_z', \tag{3.20}$$

which can be exactly integrated when the trace emittance is constant:

$$\sigma_{zf}'^2 - \sigma_{z0}'^2 = \epsilon_{z,z'}^2 \left(\frac{1}{\sigma_{z0}^2} - \frac{1}{\sigma_{zf}^2} \right). \tag{3.21}$$

At the waist, σ_z reaches a local minimum, allowing us to set $\sigma'_{zf} = 0$. Immediately following the buncher, we can write $\sigma'_{z0} = \frac{\langle z_0 z_0' \rangle}{\sigma_{z0}} - \frac{\sigma_{z0}}{f}$, where the first term represents incoming phase space correlations and the second term represents the linear chirp imparted by the buncher cavity. The waist size can be expressed as:

$$\sigma_{zf} = \frac{1}{\sqrt{\frac{1}{\epsilon_{z,z'}^2}\left(\frac{\langle z_0 z_0' \rangle}{\sigma_{z0}} - \frac{\sigma_{z0}}{f}\right)^2 + \frac{1}{\sigma_{z0}^2}}} \approx \frac{f\epsilon_{z,z'}}{\sigma_{z0}} \left| 1 - \frac{f\langle z_0 z_0' \rangle}{\sigma_{z0}^2} \right|^{-1}, \tag{3.22}$$

where we neglect the term $\frac{1}{\sigma_{z0}^2}$, which is small compared to $\frac{1}{\sigma_z^2}$ for large compression factors.

Examining the factor that accounts for initial correlations, we find that an initial negative chirp (i.e., $\langle z_0 z_0' \rangle < 0$) effectively reduces the focal length and final bunch length. Conversely, in most scenarios, an RF buncher is used to achieve strong compression, making $\frac{f\langle z_0 z_0' \rangle}{\sigma_{z0}^2} << 1$, meaning incoming correlations are minor compared to the linear correlation introduced by the RF fields. In this case, the final bunch length at the waist is $\frac{f\epsilon_{z,z'}}{\sigma_{z0}}$, directly proportional to the focal length times the energy spread, assuming thermal contributions dominate the longitudinal emittance. However, non-linear correlations induced by the buncher significantly distort the trace space and dominate the emittance in the final drift. Therefore, beams with smaller longitudinal emittance allow for shorter bunch lengths. Additionally, shorter final bunch durations can be achieved by reducing the focal length f of the RF buncher.

To estimate the minimum achievable bunch duration based on the main beamline parameters (when space-charge effects are negligible), we insert the trace space longitudinal emittance estimate incorporating Equation (3.16) or 3.15 depending if a quadratic or gaussian profile is initialized, into the envelope-equation solution (Equation (3.22)), yielding an expression as the quadrature sum of the contributions to the final emittance from (i) the initial uncorrelated relative energy spread σ_δ, (ii) non-linearities from relativistic corrections to the transport, or (iii) RF-induced emittance.

For easier comparison between beamlines of different energies, we rewrite the terms

as a function of the buncher's focal length, a useful parameter related to standard requirements on the sample chamber dimensions, pumping geometry, and transverse optics. Assuming one term dominates, we can summarize this result as:

$$\sigma_{zf} \sim \begin{cases} f\sigma_{z0'} = \frac{f}{\beta^2\gamma^2}\sigma_\delta \\ \sqrt{2}\frac{|\eta_2|}{\eta_1}\alpha k\sigma_{z0}^2 \approx \frac{3\sqrt{2}\gamma^2}{2}\frac{\sigma_{z0}^2}{f} \\ \frac{1}{\sqrt{6}}k^2\sigma_{z0}^3. \end{cases} \tag{3.23}$$

In this particular example, a gaussian distribution was assumed. For any reasonable distribution, the first equation is the same. The second and third equations will scale the same with rms length measures, but with different numeric prefactors. The total quadrature sum of these three expressions can generally be used when two or more terms are comparable in magnitude.

In all three cases, to be expected, removing non-linearities, and starting with a shorter bunch having less energy spread is the direction to go to shrink the beam even more. Eventually hitting an optimum where the charge density in transit to the focal point is just right so that space charge repulsion disrupts away from the optimum.

This is validated by observing the scalings in the formulas. In the first case, non-linearities in the transport can be ignored, and the final bunch length is proportional to the initial relative velocity spread σ'_{z0}. For the same uncorrelated relative energy spread σ_δ, a relativistic energy system has an advantage in achieving ultrashort bunch lengths due to the inverse square γ dependence in the formula. However, as the energy increases, non-linearities due to relativistic dynamics in the drift likely become the dominant contribution to the final emittance and bunch length. In this regime, after approximating $\eta_2/\eta_1^2 \approx 3/2\gamma^2$ and using the definition of the focal length (Equation (3.14)), the final bunch duration scales with the square of the beam energy. The RF curvature, the third equation, dominates the final phase space shape for very long initial bunch lengths, setting the limit for the shortest achievable bunch duration. This contribution is crucial for both relativistic and non-relativistic energies and strongly favors using lower RF frequencies. Hence, in all THz compression experiments conducted so far, this term has been a significant limit to the final bunch length [83, 90].

The analytical formulas summarized in Eq. 3.23 were found in excellent agreement with GPT in both the relativistic and non-relativistic regimes. The results are shown in Fig. 3.4 top and bottom for the 4.6 MeV and 150 keV cases, respectively. In the low-energy case, the cubic non-linearity from the RF curvature dominates due to the longer initial bunch length.

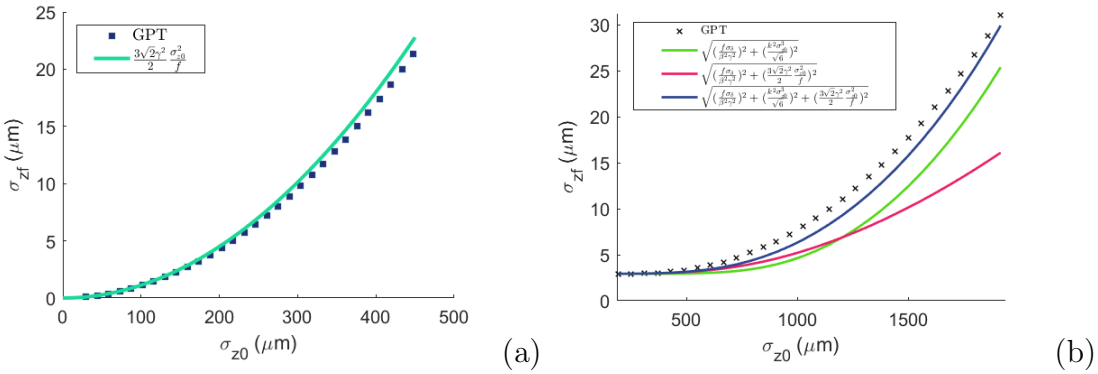


Figure 3.4: Final bunch length as a function of the initial bunch length for the (a) high energy, and (b) low energy cases. The analytical curves are also shown and are found to be in very good agreement with the simulations.

In the absence of space-charge, the shortest bunch durations are obtained by minimizing the emittance growth in the temporal lens, which can be done by using shorter input beams. At the same time, as we will see, decreasing the initial bunch length also increases the initial peak current so that the space-charge effects in the longitudinal envelope equation can no longer be neglected. As a result, space-charge effects begin to take over at a certain point and prevent further bunch compression. Thus, there must be an optimum initial bunch length to inject, which exactly balances space charge and RF emittance growth.

3.2.2.4.2 Bunch limits with space charge

To evaluate the limits imposed by space charge, we solve the envelope equations assuming the beam is dominated by space charge, thereby neglecting the emittance term. In this space-charge-dominated regime, the envelope equations for each dimension can be written as:

$$\frac{d}{ds}\left(\frac{\gamma^2\sigma_z'^2}{2}\right) = \gamma^2 K_L \frac{v(\mathcal{A})}{\sigma_z^2}\sigma_z' \tag{3.24}$$

$$\frac{d}{ds}\left(\frac{\sigma_x'^2}{2}\right) = \gamma^2 K_L \frac{u(\mathcal{A})}{2\sigma_x\sigma_z}\sigma_x' \tag{3.25}$$

Assuming axial symmetry, the total kinetic energy in the beam's rest frame (which is non-relativistic) is the sum $\gamma^2\sigma_z'^2/2 + \sigma_x'^2$. Thus, we have:

$$\frac{d}{ds}\left(\frac{\gamma^2\sigma_z'^2}{2} + \sigma_x'^2\right) = \gamma^2 K_L \left(\frac{v(\mathcal{A})}{\sigma_z^2}\sigma_z' + \frac{u(\mathcal{A})}{\sigma_x\sigma_z}\sigma_x'\right) \tag{3.26}$$

The right-hand side is a perfect differential related to the electrostatic self-energy of the beam in the rest frame, i.e., $\frac{v(\mathcal{A})}{\sigma_z^2}\sigma_z' + \frac{u(\mathcal{A})}{\sigma_x\sigma_z}\sigma_x' = -\frac{d}{ds}\left(\frac{\coth^{-1}(\frac{1}{\xi})}{\sigma_z\xi}\right)$. Therefore, when the beam is dominated by space charge, the dynamics of the envelope edges resemble Hamiltonian mechanics. After some algebraic simplifications, we obtain a conservation equation:

$$\frac{d}{ds}\left(\frac{\sigma_z'^2}{2} + \frac{\sigma_x'^2}{\gamma^2} + K_L\frac{\coth^{-1}(1/\xi)}{\sigma_z\xi}\right) = 0 \tag{3.27}$$

Here, the first two terms represent kinetic expansion, while the last term represents the electrostatic self-energy of the beam in the rest frame.

Experimentally, to shrink the beam in all dimensions, one would use a bunching cavity followed by a solenoid or quadrupole multiplet. The focal length of a buncher ($f_{||}$) is typically much longer than that of a solenoid or quadrupole lenses ($f_\perp$). To utilize the conservation equation, we assume the beam starts emittance-dominated and remains so up to the lens, with σ_z' staying constant leading up to the lens. We also assume the beam is collimated transversely, with low divergence. Thus, just after the lens, the slopes are $\sigma_z' = -\sigma_{z0}/f_{||}$ and $\sigma_x' = -\sigma_{x0}/f_\perp$. For a simultaneous waist, the envelope edge derivatives are zero, leading to the final bunch length:

$$\sigma_z = \frac{2f_{||}^2 K_L}{\sigma_{z0}^2}\left(\frac{\coth^{-1}(1/\xi_f)/\xi_f}{1 + 2(\mathcal{A}_0 f_{||}/f_\perp)^2}\right) \tag{3.28}$$

where $\xi_f = \sqrt{1-\mathcal{A}_f^2}$, and $\mathcal{A}_f$ and $\mathcal{A}_0$ are the final and initial aspect ratios, respectively. Typically, the initial aspect ratio is small enough to neglect the term $\mathcal{A}_0 f_{||}/f_\perp$ in the denominator.

Two scenarios are noteworthy: First, when aiming to focus the beam in all dimensions to achieve a unit aspect ratio (a spherical shape in the beam's rest frame), the final aspect ratio approaches unity. The factor $\coth^{-1}(1/\xi_f)/\xi_f \to 1$, and the final bunch length is:

$$\sigma_{zf} \sim \frac{2f_{||}^2 K_L}{\sigma_{z0}^2} \tag{3.29}$$

For instance, if the RF emittance growth is not compensated and dominated by quadratic non-linearity, setting Equation (3.29) equal to the second expression in Equation (3.23), the optimum initial bunch size is:

$$\sigma_{z0}^2 \sim \sqrt{\frac{f_{||}^3 K_L}{\gamma^2}} \tag{3.30}$$

Substituting this into Equation (3.29), in the high energy limit ($\gamma_0 \gg 1$), the minimum bunch duration at the temporal waist is:

$$\sigma_{zf} = C\sqrt{\frac{f_{||} N r_c}{\gamma^3}} \tag{3.31}$$

where C is a dimensionless factor of order unity that depends on the longitudinal distribution, particularly in the computation of RF emittance growth and linear components of the space charge field. In the case of a uniform distribution, $C_u = 1.02$ and for a Gaussian, $C_g = \sqrt{2}C_u$.

The nonlinear forces in the Gaussian beam's space charge field self-consistently cause some emittance evolution, which is not accounted for in our model. The emittance is not exactly constant as the beam density becomes larger near optimal conditions. It is not straightforward to include this evolution; yet it has only a small error on the solution. Effectively, because the linear driving force plays the dominant roll in the rms envelope evolution and the emittance doesnt change signifantly near optimum. In Fig. 3.5(a), for a Gaussian beam, we compare the exact field experienced by the particles of the simulation

within the Gaussian beam before the focus with both the linear component and beam averaged force that appears in the envelope equation.

In Figure 3.5(b), a simulation of the optimum compensation for 10^5 electrons is shown. For an initial Gaussian beam, The optimum initial pulse duration is 72μm and is compressed to 1μm, or 3 fs. Notably, the final phase space shows space charge correlations, with the longitudinal phase space partially reflecting the Gaussian beam's field distribution and partially exhibiting quadratic distortion. Nonetheless, the final bunch length is in good agreement with the analytical approximation, differing by less than 15%.

In Figure 3.5(c), the comparison of Equation (3.31) with GPT simulations, while setting the initial bunch length according to Equation (3.30) as the charge varies from $10^5 e$ to $10^6 e$, shows good agreement for both ellipsoidal and Gaussian distributions. As expected, the Gaussian distribution yields a longer final bunch length by a factor of $\sqrt{2}$. In Figure 3.5(d), we use Equation (3.31) to visualize the dependence on energy and charge for a constant focal length of the bunching system set to 1.88 m. Due to the energy dependence intrinsic to the focal length, increasing beam energy while keeping the focal length constant poses technological challenges involving RF cavity breakdown limits and available power sources at higher frequencies.

The second scenario, if there is no lens ($f_\perp \to \infty$), while the beam is collimated, then the aspect ratio of the beam becomes large, as it does the transverse force becomes constant suggesting we assume the transverse dimension of the beam is mostly constant, then the final bunch length satisfies the transcendental equation:

$$\sigma_{zf} = C\sqrt{\frac{f_{||} N r_c}{\gamma^3} \frac{\coth^{-1}(1/\xi_f)}{\xi_f}} \tag{3.32}$$

Note that the factor $\coth^{-1}(1/\xi_f)/\xi_f$ is less than 1 in this case, indicating a reduction from the optimum achievable with a unit aspect ratio. In fact, plotting $\sqrt{\frac{\coth^{-1}(1/\xi_f)}{\xi_f}}$ between aspect ratios of 1 and 10 shows a decrease in the space charge contribution by more than a factor of 2, as shown in Fig. 3.6.

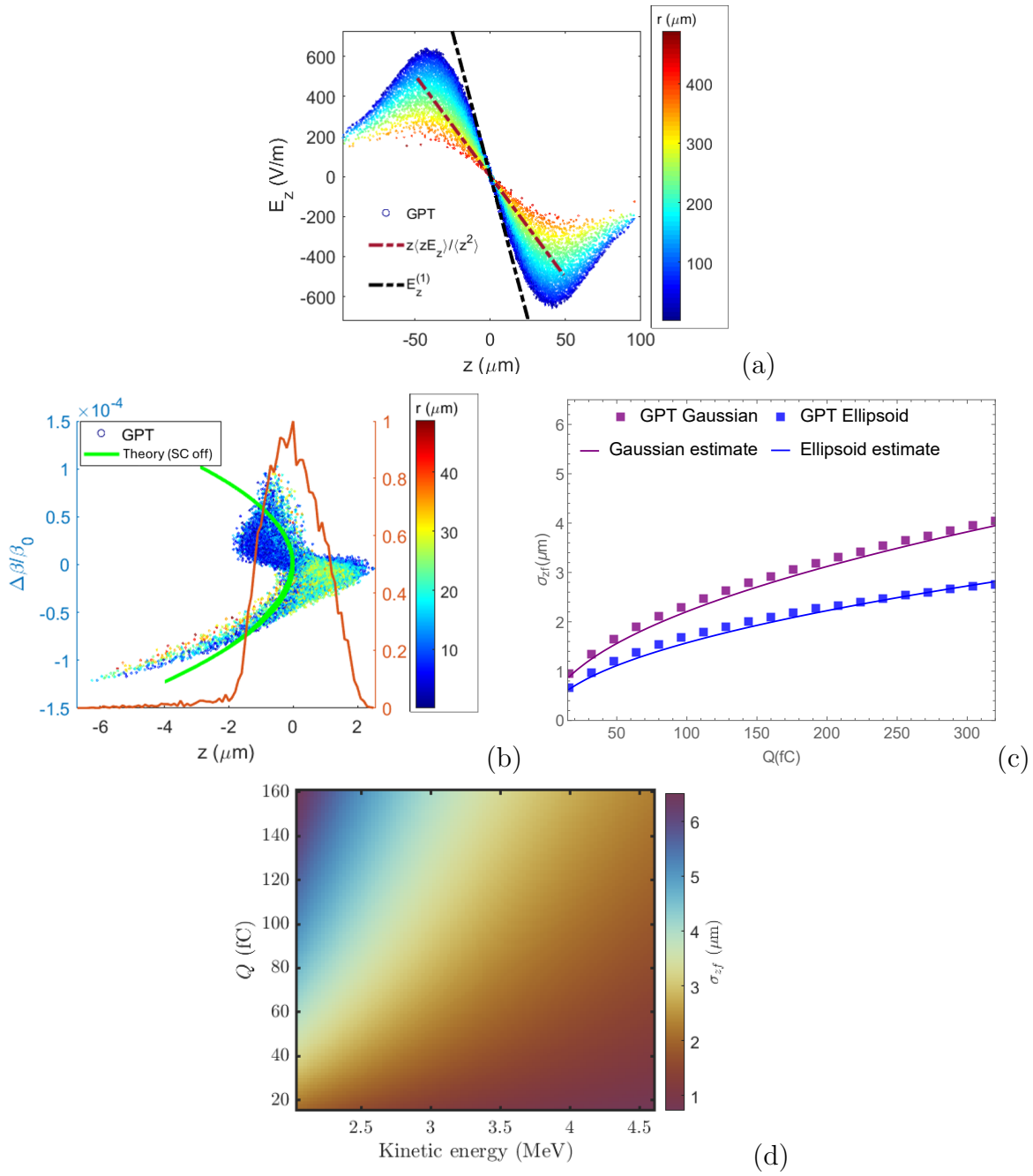
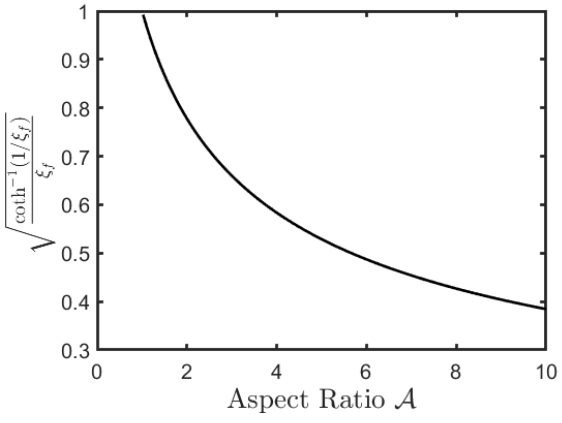


Figure 3.5: (a) Longitudinal field of the bunch compared with the linear field component of the 3D Gaussian field and with the beam-averaged generalized force term that appears in the envelope equation. The total charge in the bunch in this simulation is $Q = 10^5 e$.(b) The final phase space for an optimized Gaussian bunch. (c) Minimum bunch length analytical estimates compared with GPT simulation as the charge in the 4.6 MeV energy beam is varied for Gaussian (purple) and uniformly filled ellipsoidal (blue) beam distributions. (d) Minimum bunch length versus γ and N for the Gaussian distribution assuming a constant focal length $f = 1.88$ m.



(a)

Figure 3.6: Plot of $\sqrt{\frac{\coth^{-1}(1/\xi_f)}{\xi_f}}$ as a function of the aspect ratio $\mathcal{A}$. The graph illustrates how the space charge contribution decreases significantly, by more than a factor of 2, as the aspect ratio increases from 1 to 10.

3.2.2.5 Longitudinal emittance compensation using a higher harmonic cavity

Assuming the input beam has an uncorrelated LPS and traverses the higher harmonic cavity first, then traverses a bunching cavity second while operated at zero crossing. Then, the energy correlation can be written as:

$$\begin{aligned}\gamma = \gamma_0 - \alpha_1 \sin(k_1 z) - \alpha_2 \cos(k_2 z + \phi_2) \\ \approx (\gamma_0 - \alpha_2 \cos(\phi_2)) - z(k_1\alpha_1 - k_2 \sin(\phi_2)) \\ + \frac{1}{2} z^2 k_2^2 \alpha_2 \cos(\phi_2) + \frac{1}{6} z^3 (k_1^3 \alpha_1 - k_2^3 \alpha_2 \sin(\phi_2)) \quad (3.33)\end{aligned}$$

where γ_0 is the initial energy, α_1 is the amplitude of the bunching cavity, k_1 the wave number, α_2 is the amplitude of the x-band, k_2 is the x-band wave number, and ϕ_2 is the phase of the x-band. The constant term in (3.33) defines the resonant energy in the drift after the buncher leading up to the focus, i.e., $\gamma_c = \gamma_0 - \alpha_2 \cos(\phi_2)$. After subtracting the resonant energy, the correlation between energy deviation and position is denoted as $\Delta\gamma(z)$

After substitution of $\Delta\gamma(z)$ into (3.9), and collecting terms, we obtain $z' = z'_0 + c_1 z + c_2 z^2 + c_3 z^3 + O(z^4)$, where z'_0 is added to account for an uncorrelated spread in velocity,

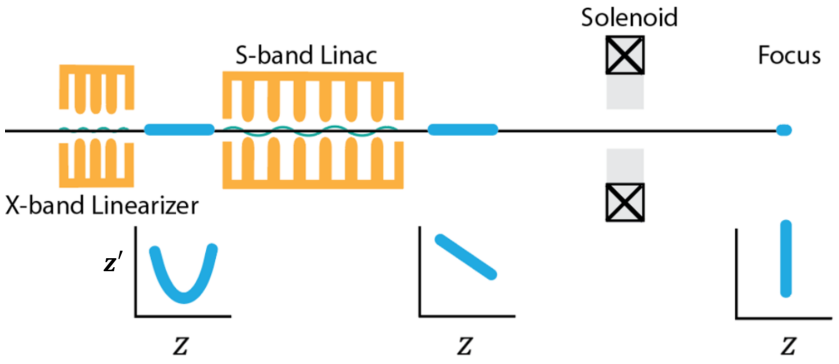


Figure 3.7: The addition of a short x-band cavity compensates for the curvature in trace space imparted by the S-band linac and gun and vertically aligning the output trace space.

and the coefficients are given by:

$$c_1 = -\eta_1(k_1\alpha_1 - k_2\alpha_2\sin(\phi_2)) \tag{3.34a}$$

$$c_2 = \frac{1}{2}\eta_1 k_2^2\alpha_2\cos(\phi_2) + \eta_2(k_1\alpha_1 - k_2\alpha_2\sin(\phi_2))^2 \tag{3.34b}$$

$$\begin{aligned} c_3 = \frac{1}{6}\eta_1(k_1^3\alpha_1 - k_2^3\alpha_2\sin(\phi_2)) \\ - \eta_3(k_1\alpha_1 - k_2\alpha_2\sin(\phi_2))^3 \\ - \eta_2\alpha_2 k_2^2\cos(\phi_2)(k_1\alpha_1 - k_2\alpha_2\sin(\phi_2)) \end{aligned} \tag{3.34c}$$

The longitudinal positions change in the drift after the buncher according to $z_f = z + sz'$. The linear correlation of the longitudinal phase space is canceled after the beam drifts a distance $f = 1/|c_1|$, which in the absence of higher harmonic k_2, gives the general expression for the focal length of the buncher 3.14.

The higher order coefficients contribute to emittance growth, and manifest distortions of the longitudinal phase space at the focus, which ultimately limit the electron beam from achieving the shortest bunch length.

By inspection, second-order correlations can be canceled by setting $\phi_2 = 0$ and $\alpha_2 = 2\frac{|\eta_2|}{\eta_1}\left(\frac{k_1}{k_2}\right)^2\alpha_1^2$. Although this solution still retains third-order curvature on the final phase space, it reveals a compensation amplitude with quadratic dependence on relative frequencies, the amplitude of the buncher, as well as the resonant energy (which in turn scales linearly with the initial accelerating voltage). Meanwhile, the amplitude needed

to compensate for any quadratic curvature imbued upon the longitudinal phase space during the initial injection scales only linearly with gun voltage [91], thus our assumption that the initial correlations are negligible is mildly justified.

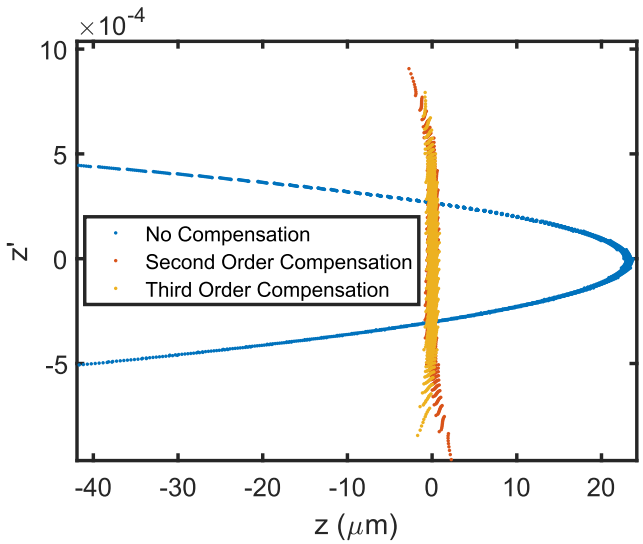


Figure 3.8: Longitudinal phase spaces for three cases, no compensation, only second order compensation, and third order compensation.

Although there is an analytic solution where $c_2 = 0$ and $c_3 = 0$ simultaneously, the expression is omitted. The solution is typically close in phase-amplitude parameter space to the second-order compensation. Therefore, to find the cavity phase and amplitude which eliminates both second and third-order coefficients, a numerical approach is employed whereby analyzing the neighborhood of phases around $\phi_2 = 0$ where $\alpha_2(\phi_2)$ is such that $c_2 = 0$ reveals a simultaneous compensation of c_2 and c_3. In the case that all higher order coefficients are negligible, the final bunch length is $\sigma_{zf} = \sigma_\gamma/(k_1\alpha_1 - k_2\alpha_2\sin(\phi_2)) \approx f\eta_1\sigma_\gamma$, where σ_γ is the uncorrelated energy spread. So in the ideal case, the bunch length is determined by the focal length, beam energy, and uncorrelated energy spread. The analytical description is illustrated in Fig. 3.8.

3.2.2.6 Managing non-linear space charge forces

Various factors can prevent reaching the limit set by the initial uncorrelated relative energy spread. If space charge effects are negligible and ballistic non-linearities are addressed, then a wide range of initial bunch lengths can be compressed nearly to the

thermal limit, assuming the bunch lengths are short enough that fourth-order distortions are insignificant. When space charge is significant, compression is constrained by radial and longitudinal slice-dependent energy spread.

Uniform picosecond-long cigar-shaped beams are desirable due to their linear fields, which help manage the growth of slice-dependent energy spread and fully optimize compression. Other charge distributions can cause unfavorable distortions at the final focus. In practice, we plan to generate electron beams of this quality by stretching the temporal profile of the drive laser into a quadratic shape with an RMS duration on the order of picoseconds while focusing the initial cathode spot size to approximately 10 μm. Imperfections in the temporal and transverse profiles can significantly hinder this approach, but diagnostics may allow for corrections during operation.

With compensation, the need for a precisely selected initial bunch length to achieve optimum compression is relaxed. In the uncompensated case, the optimum initial bunch length is sharply defined; however, with compensation, a broad range of bunch lengths achieve nearly the same performance, reaching the thermal limit. This is illustrated in Fig. 3.9(a), showing results of a GPT simulation scan of the bunching process with a charge of 250 fC for a range of initial bunch lengths. The final bunch length is plotted against the initial bunch length for both compensated and uncompensated cases. In the uncompensated case, the best compression occurs just above 1 μm, requiring the initial bunch length to be nearly 300 fs, with a quadratic dependence of the final bunch length on the initial bunch length. This optimum working point was established at Pegasus [92]. However, when non-linearities are compensated, any initial bunch length greater than 1.5 ps nearly achieves the thermal limit, with an optimum at 2.5 ps, as long as fourth-order curvature remains negligible. The necessity for optimal shaping of the electron distribution is evident in Fig. 3.9(b). In this scan, the initial bunch length is set to the optimum, and the charge is varied from 0 up to 2 pC for two distributions: Gaussian and uniform ellipsoid. The compressed Gaussian bunch length experiences substantial non-linearities at relatively low charges compared to the uniform ellipsoid, which can pack up to 1 pC of charge to sub-micron lengths while the transverse dimensions are focused to a few microns by a 20 cm focal length solenoid. Fig. 3.9(c) shows the space charge fields of the two distributions at 250 fC as each beam approaches focus. The linearization of

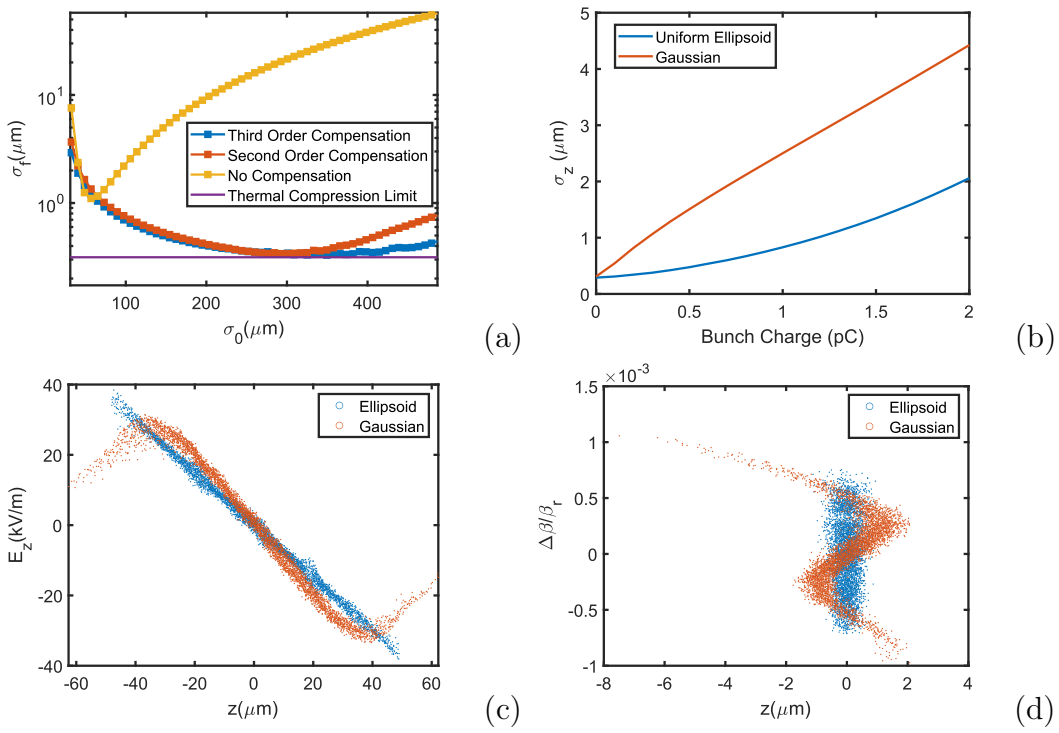


Figure 3.9: (a) Compressed bunch length as a function of the input bunch length for the uncompensated and compensated cases of a uniform ellipsoid distribution. (b) Compressed bunch length as a function of bunch charge for both the uniform ellipsoid and Gaussian spatial distributions. (c) Longitudinal component of the electric field experienced by particles at respective z-coordinate in the bunch. (d) Longitudinal phase space at the focus for the two distribution cases.

the ellipsoid and the curvature of the Gaussian are apparent. Fig. 3.9(d) shows the field curvature imprint on the longitudinal phase space of the Gaussian beam as a distortion, while the ellipsoid maintains linearity. In practice, adjusting the phase and amplitude of the x-band cavity can compensate to some extent for these additional distortions caused by space charge fields, but no exact analytical prescription currently exists.

Integrating these findings with the previous analytical estimates and simulations, it is clear that managing non-linear space charge effects heuristically, through compensation and optimal shaping of the electron distribution, is crucial for achieving the desired compression and minimizing emittance growth.

We have developed an analytical framework that yields a simple estimate for the minimum bunch length achievable in an RF compression beamline. Besides UED, the

formulas in this section might be helpful in the optimization of RF compression for other ultrashort electron beam applications, including radiation generation and injection into very high-frequency THz-driven [93] and laser-driven accelerators [94].

The envelope formalism allowed us to estimate the final bunch length accurately; our analysis showed that one can integrate the competing non-linear RF effects into the envelope evolution with space charge accurately with relative ease. The results indicate that an optimum initial bunch length condition (which can be satisfied by adjusting the laser pulse length on the cathode, for example) compromises the RF-induced emittance growth and the effects of the longitudinal self-fields. The simplicity of the reported expressions mainly stems from the fact that the effects from the longitudinal perveance are integrable. We have also limited the expansion yielding the non-linear terms in the emittance growth to second and third-order. Still, in principle, we could have evaluated all the higher-order terms (and compensated if enough independent knobs are added/available on the beamline).

Although realistic beams produced by photoinjectors typically present more complex phase-space distributions, the initial conditions assumed in the derivations are an initially unchirped longitudinal phase space. Nonetheless, the results obtained here still provide valuable estimates of the compression limits in a given configuration, which prove helpful as a starting point for numerical optimizations. In addition, the scaling laws have proven capable of guiding parameter choices in the design of new setups. Most importantly, these results highlight the main contributions to the final bunch length and suggest possible paths to improve the compression further and achieve sub-fs bunch lengths.

3.2.3 Energy spread minimization techniques

3.2.3.1 Relativistic RF harmonic energy spread linearization

A functional time-resolved electron microscope relies on beam sources with sufficiently low energy spread to mitigate chromatic aberration [11]. Adding a linearizing cavity of a different frequency can enhance the use of RF photoinjectors for this application [43]. The temporal nature of these instruments necessitates a detailed understanding of harmonically linearizing a longitudinal phase space to improve beam compression or

reduce energy spread. Simplified scaling laws and longitudinal envelope equations provide insights into the interplay between longitudinal emittance and space-charge forces on pulse evolution.

In an RF photoinjector, particles are violently accelerated off the cathode to relativistic energies in a few cm by a strong RF field; this process can also lead to high levels of energy spread for longer pulse durations – ones spanning multiple degrees of phase. Inside the cavity, The on-axis electric field component inside the cavity is given by:

$$E_z = E_0 \cos(kz) \sin(\omega t + \phi_0) \tag{3.35}$$

where E_0 is the peak electric field, k is the RF wave number. The cavity nominally spans a distance equal to $3\lambda/4$, ω is the RF angular frequency, and ϕ_0 is the initial phase.

Assuming a 1D scenario, the energy gain can be expressed as:

$$mc^2 \frac{d\gamma}{dt} = ec\beta_z E_z \Rightarrow \frac{d\gamma}{dz} = \frac{eE_0}{mc^2} \cos(kz) \sin(\omega t + \phi_0) \tag{3.36}$$

The harmonic factors can be rewritten as a sum of forward and backward traveling waves:

$$\cos(kz) \sin(\omega t + \phi_0) = \frac{1}{2} \left(\sin(\omega t + \phi_0 - kz) + \sin(\omega t + \phi_0 + kz) \right) \tag{3.37}$$

The first term represents the forward traveling wave, which dictates the electron's energy gain. It is what we strive to keep particles in phase with (typically by pursuing larger accelerating gradients). The phase the electron experiences is generally given by:

$$\phi = \phi_0 + k \int_0^z \left(\frac{\gamma}{\sqrt{\gamma^2 - 1}} - 1 \right) dz \tag{3.38}$$

Expressing the energy gain in terms of this phase, we get:

$$\frac{d\gamma}{dz} = \frac{eE_0}{2mc^2} \left(\sin(\phi) + \sin(\phi + 2kz) \right) \tag{3.39}$$

The first term here represents the forward wave, which we want the particles to remain in phase with, and the other term is the backward propagating wave. Approximating the dynamics to be close to a resonant phase condition, we consider $\phi = \phi_0$ to essentially remain constant during the gain process; we find $\frac{d\gamma}{dz} \approx \frac{eE_0}{mc^2} \sin(\phi_0)$, indicating rapid energy change near the cathode when considering a 70 MV/m gradient at an S-band wavenumber $k = 59m^{-1}$; the particle would gain approximately 1 MeV in the first few

cm [2]. The phase slippage complicates the problem, and requires the equations to be numerically solved. Approximate solutions reveal the scalings for the ideal cavity case. Notably, the longitudinal phase space correlations are harmonic, and to leading order, the energy spread scales linearly with pulse duration $\sigma_\gamma = \alpha k \sigma_z$ [95], which is inherently much larger for longer beams. For picosecond beams, this spread can exceed 1000 ppm, which is unusable for high-performing imaging applications.

Generally, the fields and particle dynamics in realistic gun geometries must be numerically simulated, but the overall shape of the longitudinal phase space still exhibits a harmonic correlation. For narrow phase intervals, at high energy, it is often permissible to assume the energy correlation is of the form $\gamma = \gamma_0 \cos(kz + \phi_0)$. Here, γ_0 represents the maximal energy gain from the gun, k is the spatial wave number of the gun, and ϕ_0 is it's operational phase. Still, it should be noted that the longitudinal phase space at the exit of a realistic cavity will have a polynomial correlation of the form:

$$\gamma = a_0 + a_1(kz) + a_2(kz)^2 + a_3(kz)^3 + \cdots \tag{3.40}$$

where the expansion coefficients are analogous to the coefficients in a Taylor series.

By adding a booster linac and a higher harmonic x-band linearizer, we can ensure that the desired beam energy is obtained for the application while compensating to remove the undesirable correlations, maintaining tolerable levels of energy spread.

For the analysis of minimizing the energy spread, we assume at the exit of the gun, the longitudinal phase space (LPS) exhibits a dominant correlation between energy and position relative to the beam centroid, expressed as $\gamma(z) \sim \gamma_0 \cos(kz + \phi_0)$. We also assume that the longitudinal positions of the particles remain frozen during propagation between cavities. In practice, particles may move toward the tail or head of the beam depending on their energy. This assumption holds if $\sin(\phi_0) kL/\gamma^2 \ll 1$ for a drift distance characterized by L.

Under these approximations, the LPS RF correlations after the sequence of the gun, X-band cavity, and linac can be expressed as:

$$\gamma(z) = \gamma_0 \cos(kz + \phi_0) + \alpha_x \cos(k_x z + \phi_x) + \alpha_L \cos(kz + \phi_L) \tag{3.41}$$

[2]This assumes the particles remain in phase with the forward-propagating wave; however, as the initial accelerating gradient increases, this condition becomes more readily satisfied.

where α_x and α_L are the normalized voltages of the X-band cavity and linac, respectively. This correlation can be Taylor-expanded to third order as:

$$\gamma(z) = c_0 + c_1 z + c_2 z^2 + c_3 z^3 + \mathcal{O}(z^4) \tag{3.42}$$

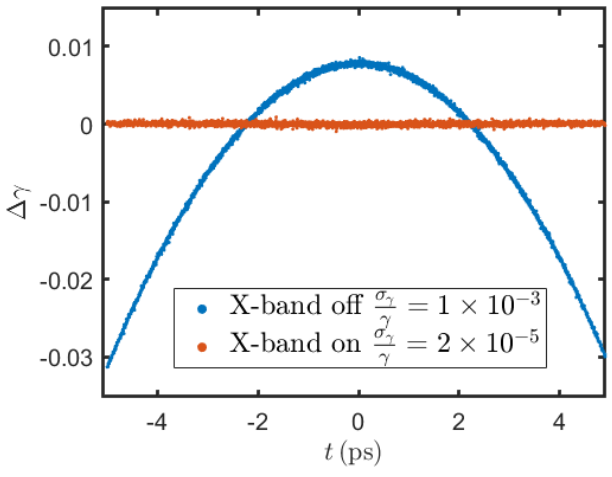
The expansion coefficients c_1, c_2, and c_3 are the primary contributors to the energy spread. One solution where these coefficients vanish is if the following conditions are met: the X-band phase is set to π, the linac phase satisfies $\sin(\phi_L) = -\gamma_0 \sin(\phi_0)/\alpha_L$, and the X-band voltage is:

$$\alpha_x = \left(\gamma_0 \cos(\phi_0) + \sqrt{\alpha_L^2 - \gamma_0^2 \sin^2(\phi_0)}\right) \frac{k^2}{k_x^2} \tag{3.43}$$

Under these conditions, c_0 is given by:

$$c_0 = \left(\gamma_0 \cos(\phi_0) + \sqrt{\alpha_L^2 - \gamma_0^2 \sin^2(\phi_0)}\right) \left(1 - \frac{k^2}{k_x^2}\right) \tag{3.44}$$

Since $c_0 \to 0$ when $k_x = k$, the need for higher harmonics is apparent to ensure the beam is not overly decelerated by the compensation cavity. Typically, the optimal gun phase and voltage are determined by factors such as transverse emittance and charge. However, including the linac allows for tunable beam energy, with compensation across the entire range of accessible energies.



(a)

Figure 3.10: LPS of the beam before (blue) and after (red) the use of an x-band cavity. The correlated energy spread has been removed.

In Fig. 3.10, an illustration of harmonic linearization of the initial energy correlation from the gun is shown. A beam accelerated to 4.6 MeV of kinetic energy having a 10

ps duration and 10^{-5} uncorrelated energy spread spans nearly 10° of phase at S-band, which causes RF significant curvature, culminating in 10^{-3} rms relative energy spread. Introducing a cavity at the third harmonic, i.e., $k_x = 3k$, it is shown to completely remove the RF curvature without a significant loss of beam energy, thus reducing the relative spread to the 10^{-5} level or 10 ppm.

After all correlations are removed, the thermal spread in energy is the final limit. The thermal spread can grow by intrabeam scattering. The change in energy spread from intrabeam scattering can be estimated through the use of the following expression:

$$\frac{d\sigma_\gamma^2}{dz} = \frac{2r_e^2 N_b}{\sigma_x \sigma_z \epsilon_{nxr}} \tag{3.45}$$

where $r_c = \frac{e^2}{4\pi\epsilon_0 mc^2}$, is the classical electron radius and N_b is the number of particles in the bunch [96]. Assuming a coasting, 10ps long beam, with an average rms spotsize $\sigma_x = 100\mu$m, having 100nm normalized emittance, after a 1 meter of drift distance, we get $\sqrt{\Delta\sigma_\gamma^2} \approx 2 \cdot 10^{-5}$, which is negligible. However, beams having better emittance are more susceptible to growth, and considering the beam evolution and tight focusing in these scenarios may increase the spread by an order of magnitude from what was just calculated as well, but even at this level, it is still manageable for high-performance imaging.

3.2.3.2 Sub-relativistic time-resolved EELS RF cavity based monochromatization

Now we briefly explore a concept pulsed transmission EELS scheme for non-relativistic 40-100 keV table-top setups, enabling access to meV-scale energy resolution and significantly higher beam brightness compared to current EELS instruments. The scheme uses a very short (40 fs FWHM) laser pulse to emit a burst of electrons from a flat photocathode in a high-tension DC gun. Two properly phased radiofrequency (RF) cavities act as a temporally magnifying telescope, stretching and collimating the electron bunch by nearly two orders of magnitude to 5-10 ps while proportionally reducing its energy spread. Unlike the original approach by [72], which involved strong coupling between transverse and longitudinal dynamics due to the nanoscale dimensions of the electron emitter, our use of a flat photocathode decouples beam size evolution, allowing for larger electron charges per

bunch. This simplification enables an analytical description of the longitudinal dynamics, revealing the scaling laws and limits in energy spread reduction.

Using a 2D energy analyzer, the energy spectra can also be momentum-resolved, revealing the dispersion characteristics of excitations. With the electron beam intrinsically synchronized with an ultrafast laser, our scheme is ideal for pump-probe studies, where energy resolution can be traded off for temporal resolution. An EELS instrument like this will represent a significant advancement in electron-beam technologies, enabling quantification of the energy loss function in quantum materials and the first momentum-resolved collective-mode studies of far-from-equilibrium electronic systems. Here, we review an analytical model describing the two RF cavity monochromatization scheme.

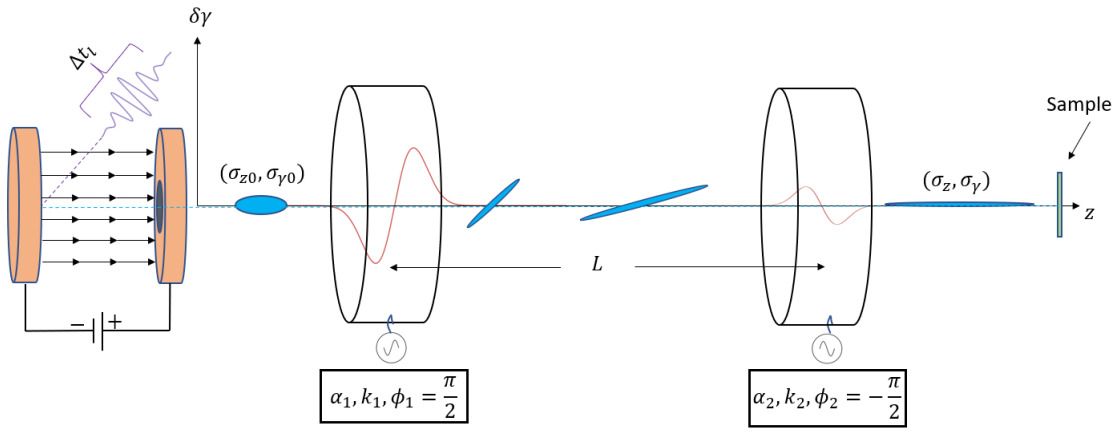


Figure 3.11: Cartoon depiction of compensation scheme. All elements are axially symmetric. Two RF cavities working in the TM010 mode are used to stretch the initial electron bunch length while simultaneously reducing its energy spread. Bunch arrives at the first cavity at the zero crossing of the field, receiving a positive chirp (energy-time correlation) which causes it to elongate in the following drift. The second cavity is tuned in amplitude and phase in order to remove the energy chirp and minimize the final energy spread.

A cartoon of the proposed beamline is depicted in Fig. 3.11. Therein, it is shown a DC photo-injector with a photocathode which is used to emit a short (40 fs) pulse electron beam containing a small amount of charge ($<$ 5000e potentially down to a single electron per pulse). The beam is then accelerated to 40keV and through two TM010-like cavities. The first cavity is set to act as a temporal defocusing lens which

stretches the beam. Due to Liouville's theorem, a temporally stretching beam must simultaneously have its intrinsic energy spread compressed to conserve phase space area. This principle is captured in the cartoon where the thickness of the longitudinal phase space ellipse is shown to decrease after the first cavity. As the beam exits the first cavity, the longitudinal phase space exhibits a positive correlation between position and energy (positive chirp). As the beam drifts to the next cavity, the pulse stretches in time, meanwhile the uncorrelated energy spread (the thickness of the longitudinal phase space ellipse) decreases. When the beam finally enters the second cavity, the positive chirp is cancelled. The end result is a significantly reduced energy spread.

Neglecting space charge (which will always be true in the limit of single-electron pulses) and transverse effects to focus only on longitudinal single particle dynamics. For an on-axis particle the effect of a cavity can be simply described by a sinusoidal change in its kinetic energy depending on the phase experienced by the particle. The energy gain imparted by the first cavity as a function of phase offset within the bunch can then be written as:

$$\Delta\gamma_1 = \Delta\gamma_0 + \alpha_1 \sin(k_1 z_0/\beta) \tag{3.46}$$

$\Delta\gamma_1$ is the energy deviation from the reference particle at the exit of the first cavity, $\Delta\gamma_0$ is the energy deviation prior entering the RF fields, typically associated with the emission process, $\alpha_1 = e\Delta V_1/mc^2$, ΔV is the peak accelerating voltage with transit time factored into the voltage, and with m and e the electron charge and mass, c is the speed of light, and z_0 is the particles longitudinal coordinate relative to the reference particle (which is assumed to be at the center of the bunch).

Adopting the impulse approximation, which amounts to the longitudinal coordinates not changing while passing through the RF cavity. This is reasonable when considering RF cavities occupying a very short distance in the beamline; then their only effect is to change the particle energy (and not their position within the bunch).

To combine the effects of two RF cavities and the drift in between them, a non-linear dynamical map determines the evolution of the position and energy coordinates $(z, \Delta\gamma)$ along the beamline; the map also determines the beam's final energy spread.

By design, the initial bunch length entering the first cavity is sufficiently small to

consider only the first order in the Taylor series expansion i.e. $\sin(k_1 z_0/\beta) \approx k_1 z_0/\beta$. Each electron then ballistically propagates for a distance L to the second cavity. Its relative longitudinal position in the bunch after the drift can be written as:

$$z_2 = z_0 + L\frac{\Delta\beta}{\beta} \tag{3.47}$$

$$= z_0 + L\left(\eta_1\Delta\gamma_1 + \eta_2\Delta\gamma_1^2 + \mathcal{O}(\Delta\gamma_1^3)\right) \tag{3.48}$$

using the expansion $\frac{\Delta\beta}{\beta} = \sum_m \eta_m \Delta\gamma^m$ to second order.

The second cavity is tuned so that the reference particle arrives at a zero crossing where the fields impart a negative chirp on the beam so that the resulting final energy difference is:

$$\Delta\gamma_2 = \Delta\gamma_1 - \alpha_2 \sin(k_2 z_2/\beta) \tag{3.49}$$

The amplitude of the second cavity is set to exactly cancel the positive linear chirp imparted by the first cavity, but at this point it is important to note that the beam is much larger at the second cavity, so we can not simply ignore the non-linear terms in the Taylor series expansion. Keeping the third order contribution, we get:

$$\Delta\gamma_2 = \Delta\gamma_1 - \alpha_2\left(k_2 z_2/\beta - (k_2 z_2/\beta)^3/6\right) \tag{3.50}$$

By composing mappings, we obtain a final symplectic map in terms of the initial coordinates which enables a straightforward evaluation of the second moments of the longitudinal phase space at the exit of the second cavity once we know the input beam distribution. We can then use it as a guide to evaluate optimal settings for the energy spread reduction EELS scheme. For example, only retaining the linear terms, we can determine the relation between cavity amplitudes and frequency to cancel the linear correlation between final energy and initial position. The resulting condition is:

$$\alpha_2 = \frac{\frac{k_1}{k_2}\alpha_1}{1 + \eta_1\alpha_1 L k_1/\beta} \tag{3.51}$$

then, to linear order, the residual energy spread is given by:

$$\Delta\gamma_2 = \frac{\Delta\gamma_0}{1 + \eta_1\alpha_1 L k_1/\beta} \tag{3.52}$$

These expressions point out how the compensation of energy spread is enhanced by increasing the first cavity frequency, increasing the drift length, or increasing the amplitude

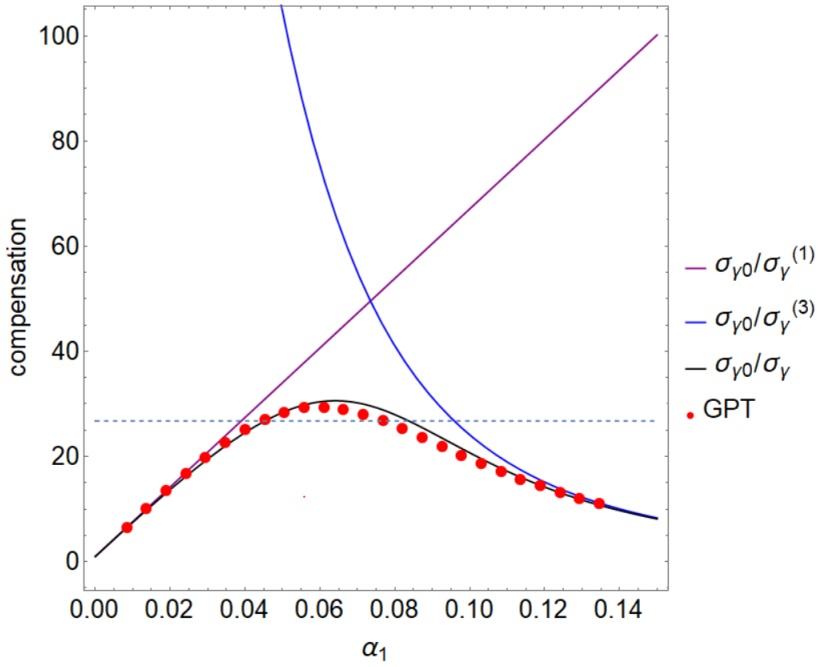


Figure 3.12: Final RMS energy spread (normalized to initial energy spread) near optimum as a function of the first cavities amplitude plotted using Eq. 3.54

of the first cavity. Although both scenarios increase sampled RF phase in either cavity, making the problem less linear as a whole.

Looking more in detail, while Eq. 3.52 suggests that the compensation is independent of the second cavity, provided its amplitude is set accordingly to Eq. 3.51. This issue can be addressed if we consider the higher-order terms in the Taylor expansion; it is easy to see a limit in compensation associated with the longitudinal emittance growth. In practice, the main issue that must be dealt with when optimizing this scheme is that the beam cannot sample too much RF phase at the second cavity, to avoid the non-linearities induced emittance growth which ruins the compensation. Optimal tuning of the first cavity amplitude is essential to minimize this effect.

To this end, we take advantage of the third order map to directly evaluate the final energy spread of the beam at the exit of the second cavity. In general, the second moment is given by:

$$\sigma_\gamma = \sqrt{\int \Delta\gamma_2(z_0, \Delta\gamma_0)^2 f(z_0, \Delta\gamma_0) dz_0 d(\Delta\gamma_0)} \tag{3.53}$$

If we make the assumption of an incoming gaussian distribution in both initial energy

and position, we can perform the integrals which result in the following quadrature sum:

$$\begin{aligned}
\sigma_\gamma &\approx \sqrt{\left(\sigma_\gamma^{(1)}\right)^2 + \left(\sigma_\gamma^{(3)}\right)^2} \\
\sigma_\gamma^{(1)} &= \frac{\sigma_{\gamma 0}}{1+\eta_1\alpha_1 L k_1/\beta} \\
\sigma_\gamma^{(3)} &= \sqrt{\frac{5}{12}}\alpha_1 k_1 k_2^2 \left[\left(1+\eta_1\alpha_1 L\frac{k_1}{\beta}\right)^2 \sigma_{z0}^3/\beta^3 + \frac{(\sigma_{\gamma_0}\eta_1 L)^3/\beta^3}{1+\eta_1\alpha_1 L k_1/\beta}\right]
\end{aligned} \tag{3.54}$$

Note that the linear order term, $\sigma_\gamma^{(1)}$, asymptotically approaches zero as the first cavity amplitude increases. Meanwhile, the third order term, $\sigma_\gamma^{(3)}$ begins to dramatically increase when the final rms bunch length at the second cavity entrance samples an appreciable amount of RF phase. Furthermore, the third order term has a cubic dependence on σ_{z_0} which clearly imposes a limit on the possible compensation by requiring a smaller initial bunch length to improve compensation. Given that the scheme relies on the elongation of the bunch to compensate for energy spread, beginning with a smaller bunch provides more room for the bunch to grow before it reaches a point of sampling too much RF phase at the second cavity. The initial bunch length is dependent on the laser pulse length which will thus determine the maximum compensation possible. For a initial bunch length of 14μm, the optimum is evident in Fig. 3.12, with the optimum compensation being 27, corresponding to a final energy spread of $\sigma_{\gamma_0} = 3.6 \times 10^{-6}$, and a final bunch length of 1.2 mm.

Note that the first cavity amplitude that sets the minimum is very close to the predicted crossing of linear contributions to energy spread with third-order contributions. The optimum first cavity setting taking into account the nonlinearities can be estimated by equating $\sigma_\gamma^{(1)} = \sigma_\gamma^{(3)}$, then solving for α_1. In the assumption of a strong initial chirp i.e. $\alpha_1\eta_1 k_1 L >> 1$, the optimum amplitude of the first cavity is well approximated by:

$$\alpha_{1,opt} \approx \frac{\sqrt[8]{\frac{3}{5}}\sqrt[4]{2}\sqrt[4]{\sigma_{\gamma 0}}\beta^{3/2}}{\eta_1^{3/4} k_1\sqrt{k_2} L^{3/4}\sigma_z^{3/4}}, \tag{3.55}$$

then substituting the acquired amplitude into Eq.3.52 and scaling by a quadrature factor leads to:

$$\sigma_{\gamma,opt} \approx \frac{\sqrt[8]{\frac{5}{3}} 2^{3/4}\sigma_{\gamma 0}^{3/4}\sigma_z^{3/4}\sqrt{\frac{k_2}{\beta}}}{\sqrt[4]{\eta_1 L}} \tag{3.56}$$

Eq. 3.55 together with Eq. 3.56 serve as excellent starting points for more complete simulations of the scheme, which take into account transverse focusing and require ro-

bust particle tracking optimization to realize compensation. The optimum of Eq. 3.54 represents the best energy spread achievable for a particular setup.

To further justify the applicability of the aforementioned formulas, we study a simplified GPT simulation model which utilizes ideal cavity field maps built into GPT to benchmark the analytic formulas. The longitudinal field component of the TM010 cavities used in the simulation have Gaussian spatial profile expressed as $E_z(s,t) = A\exp(-s^2/2\sigma^2)\sin(\omega t+\phi)$ where A is the cavity amplitude, σ is the effective length of the cavity, ω is the cavity frequency, and ϕ is the phase. The cavity phase is synchronized such that the beam centroid gains no energy. The resulting longitudinal energy modulation is given by:

$$\begin{aligned}\Delta\gamma &= \Delta\gamma_0 + \int_{-\infty}^{\infty} E_z(s, s/c\beta)ds/mc^2 \\ &= \Delta\gamma_0 + \frac{\sqrt{2\pi}\sigma A\exp(-k^2\sigma^2/2\beta^2)}{mc^2}\sin(kz_0/\beta)\end{aligned} \tag{3.57}$$

which can used to derive α. To mimic the impulse approximation, we reduce the width of the Gaussian profile towards zero while keeping $A\sigma$ finite. Doing this in GPT also has the negative side effect of ramping up transverse effects, so we find a trade-off that makes the cavities significantly short, so the cavity behaves as an impulse while making sure the transverse fields do not play a role.

Considering the case of equal frequency cavities, each operating at (2.856 Ghz). In Fig. 3.12, the final energy spreads at the exit of the second cavity from GPT are plotted along with the predictions from Eq. 3.54. In Fig. 3.13 we illustrate initial and final phase spaces at the optimal compensating cavity amplitude. The non-linear phase space correlation begins to dominate the final distribution which presents a preeminent cubic dependence of energy modulation on particle time of arrival. A shorter initial pulse duration and low second cavity frequency lower the optimum energy spread. Recall that this is because the lower initial bunch length sets the initial phase space area, and a lower frequency reduces the effects correlations imparted by the second cavity. It is attractive to seek a lower frequency cavity because it can potentially reduce energy spread more, but this requires more power and much larger cavities given that the optimum is directly proportional to the square root of frequency.

These results highlight the main features of this scheme. Cavities less ideal than those

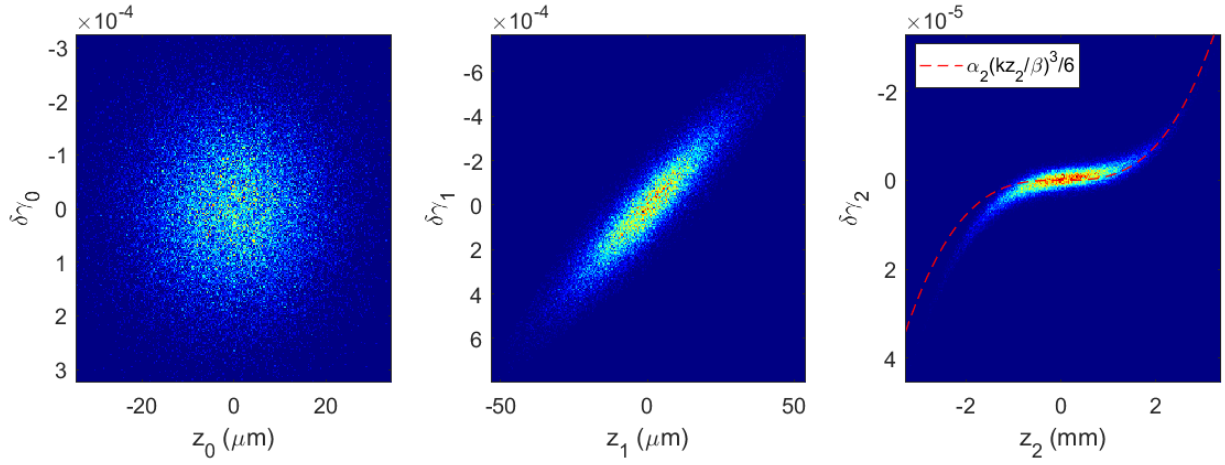


Figure 3.13: Initial, intermediate, and final phase space of mildly over-compensated simulation. Space charge effects are turned off. These phase spaces should be considered an aggregate of many shots.

just considered can be used to achieve more realistic expectations; however, the general scalings presented will still apply. This 1D model sets the limit of the experimental setup and will guide future optimization of the low-energy EELS beamline setup to be implemented at the Pegasus lab.

3.3 Pegasus beamline - longitudinal phase space shaping

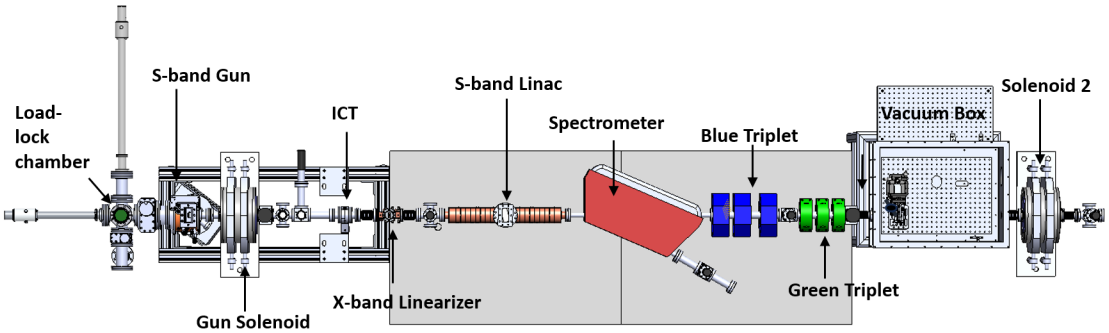


Figure 3.14: Layout of the PEGASUS beamline. The electron beam is accelerated by the S-band gun and travels downstream through additional accelerator components from left to right.

The Pegasus beamline is a compact electron beamline utilized for a variety of ap-

plications, including the study and development of ultrafast imaging, THz generation, dielectric laser accelerators, and high-brightness beam diagnostics. Aerospace companies have also employed the Pegasus beamline in target studies, showcasing its versatility.

Pegasus is housed in a 10 m x 3 m radiation-shielded bunker, complemented by a nearby laser room equipped with a 3-mJ, 45 fs Ti:Sapphire laser. The accelerator consists of an S-band RF photogun, a short S-band 11-cell linac, and a compact 7-cell X-band linearizer.

The most recent configuration of the Pegasus beamline is shown in Fig. 3.14. Electrons originate on the left, inside the RF photogun [37], where they are accelerated to 99% of the speed of light. Following the photogun, the main components are: the gun solenoid for collimating the divergent beam exiting the gun; an ICT [97] for shot-to-shot charge measurements; the X-band linearizer, which is used to linearize the longitudinal phase space; a linac [98] that can either boost the energy to a maximum of 7 MeV or impart a negative chirp to initiate velocity bunching; a dipole spectrometer for measuring the beam energy; blue and green quadrupole triplets for focusing into the vacuum box; and a final focusing solenoid that can be utilized for downstream applicationss.

3.3.1 1.6 Cell RF photoinjector

The 1.6 cell high-gradient RF photoinjector was originally prototyped at Pegasus and continues to be a pivotal element of the Pegasus laboratory, enabling the production of relativistic beams with high current density and low emittance [81]. Operating in a π mode, this dual capability results from accelerating electrons rapidly to relativistic energies, which mitigates the space-charge forces that would otherwise disturb the beam. Initially developed to provide a high-current source for free electron lasers [37], the innovative application at Pegasus (and similar facilities [75, 99, 100, 101, 78]) leverages the photoinjector's potential for low bunch charge operation ($< 1pC$). This mode offers low emittance ($\epsilon_{nxr} < 25nm$), provided there is a small enough photo-emitting spot, making it an invaluable tool for ultrafast scattering experiments. The actual cavity has a shunt impedance of 1.78MΩ, and given enough RF power, is designed to achieve up 100MV/m gradients after appropriately conditioning.

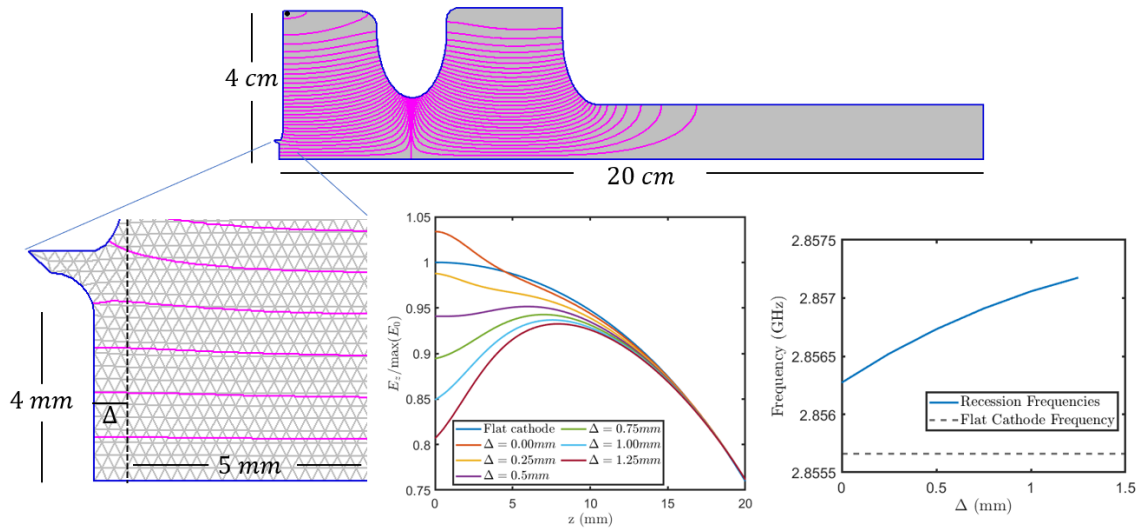


Figure 3.15: Basic cell structure of the 1.6 cell RF photoinjector. The inset shows the on-axis longitudinal field profile of the cavity mode.

The field map and geometry of the 1.6 cell is shown in Fig. 3.15. Most particle tracking simulations through the RF structures performed in this thesis are based on 1D off-axis expansions of the field for comparison with analytical formulas. This approach is suitable for low emittance beams with small envelopes that remain close to the axis. However, for the gun, a full 3D map was developed using HFSS, which shows multipole moments and small dipole kicks on the beam due to minor asymmetries in the geometry during GPT simulations. Emittance evolution agreed between the 1D and 3D maps for the range of initial laser profiles relevant to this work.

The backplane of the gun was modified to allow different cathodes to be loaded. Alkali antimonide photocathodes grown on molybdenum plug surfaces at Cornell University were loaded into the gun. The UHV load lock maintains a vacuum level of 10^{-10} torr using a combination of NEGs and ion pumps. This modification results in the recession of the actual cathode and back plane when inserted, changing the gun frequency and extraction field. The extraction field and cavity frequency for different recessions are shown in the bottom middle and bottom right plots in 3.15. Digital low-level RF enables the gun to achieve the shifted resonance frequency without heating.

After installing the cathodes, the gun frequency shift was measured with a vector network analyzer. The shift in frequency corresponded to a 1.25 mm recession in Superfish

simulations.

Using this recession and using the gun solenoid to perform a solenoid scan. The MTE in the GPT simulation model can be fit to the experimentally measured beam size. The resulting fit is shown in Fig. 3.16(a). Given the measured laser spot on the cathode was 190 μm, laser pulse duration of 100fs rms, 15 degrees RF phase, 500 fC, and 42.5 gunloop mV (corresponding to a 3.2 MeV beam), the fit resulted in 1 eV at 266 nm and 0.45 eV at 400 nm. These MTE values are comparable to those of copper. Given the observed QE degradation in transit from Cornell to UCLA, it is not surprising that the MTE is higher than expected for what are supposed to be high-quality cathodes.

From this recession study, we found that Na K Sb photocathodes at high gradients maintain comparable emittance to copper cathodes at UV wavelengths for multiple weeks of operation after initial degradation during transport.

3.3.1.1 Thermal emittance dominated beams

For single-shot imaging applications, the ideal mode of operating is primarily in a low-charge ($< 1pC$) cigar regime, where the incident laser pulse is long ($> 1ps$) to ensure the pulse duration is preserved. At the same time, the illumination spot on the cathode is small ($< 50\mu$m) for minimal emittance. For UED, it may be favorable to initialize a very short laser pulse and, subsequently, a pancake aspect ratio to achieve sub-ps resolution. The goal is to exemplify that although longitudinal emittance grows substantially for longer pulses at the gun exit due to RF, space charge-driven dynamics at the cathode are sufficiently weak that each slice of the beam is emittance-dominated at the gun exit when operating either in the cigar or pancake regime.

Although the RF forces are linear near the axis, the temporal structure correlates the transverse position and the longitudinal emission time or exit phase. Therefore, it is necessary to assess how the cathode emittance is modified by the RF field's temporal structure. As shown in [95], calculating the transverse impulse on an electron at the exit of the gun leads to the expression $\frac{p_x}{mc} = \frac{p_{x0}}{mc} + \alpha k \sin(\phi) x$ where in this case ϕ is interpreted as the exiting phase. Assuming a symmetrical distribution in z at the optimum phase,

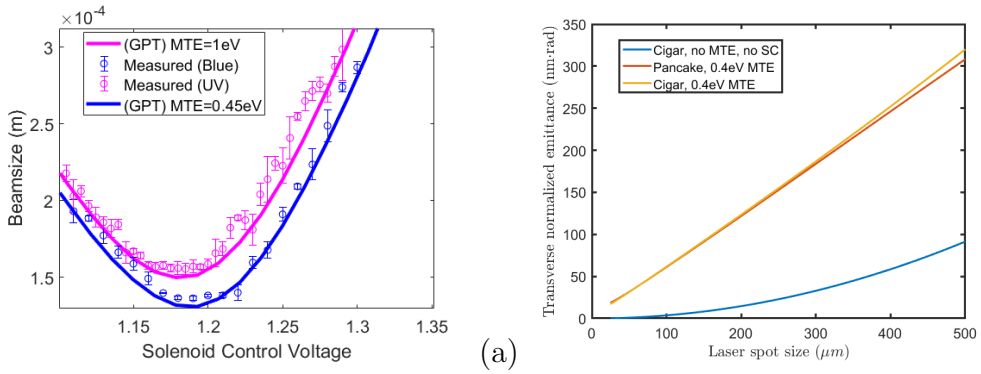


Figure 3.16: (a) Na-K-Sb MTE measurements, solenoid scan performed on a screen downstream of the gun solenoid, using a 500fC, emitted from a 190μm spot, with a 40 fs duration. (b) Transverse emittance dependence on initial laser spot size at the gun exit for pancake and cigar beams, as well as the case of a cigar beam with space charge off and no MTE, showing the beam is emittance dominated in both the cigar and pancake regime at the gun exit.

the transverse emittance scales as:

$$\epsilon_{x,gun} \approx \alpha k^2 \sigma_x^2 \sigma_z \tag{3.58}$$

depending on the longitudinal distribution. This contribution is assumed to be added in quadrature with the cathode emittance. Typically, we can substitute σ_x with the laser spot size and σ_z with the laser pulse duration. Either configuring the initial pulse to exhibit a cigar-like aspect ratio (with the longitudinal dimension significantly exceeding the transverse dimension) or a pancake, the emittance at the exit of the gun is thermal, for achievable laser spots on the cahtode [60]. Given the challenges in reducing the initial laser spot size; in our experimental scenario, the smallest the laser spot can feasibly be made is nearly 25μm. In Fig. 3.16(b), results from GPT simulations of the gun are shown where the beam is intercepted on a screen 15 cm from the cathode. The emittance is sampled as the laser spot size is varied for an initially parabolic laser profile with a 4.5 ps rms duration. In blue, a beam initialized with no MTE and space charge off develops RF-induced emittance growth for large spots. In orange and yellow, the space charge is set to 500fC with an MTE of 0.4 eV; the long pulse case had a 4.5ps rms, while the short pulse case was 100fs. Given that in both the pancake and cigar regimes, the emittance is linearly dependent in the same way on the initial spot size, we can conclude the RF

forces do not appreciably disrupt beam quality and normalized transverse emittance is thermal at the gun exit in the experimental scenario. Thus, the main issue to address is the longitudinal phase space correlations for cigar beams.

3.3.2 Dual slot resonance linac electromagnetic description

The Dual Slot Resonance Linac (DSRL) on the Pegasus beamline can serve to either velocity bunch or accelerate the beam. It was designed to achieve an energy gain of up to 10 MeV [102]. This 10+2/2 cell standing-wave structure, operating in a phase and amplitude stabilized $\pi/2$ mode, is integral to the experimental and simulation work conducted throughout this thesis.

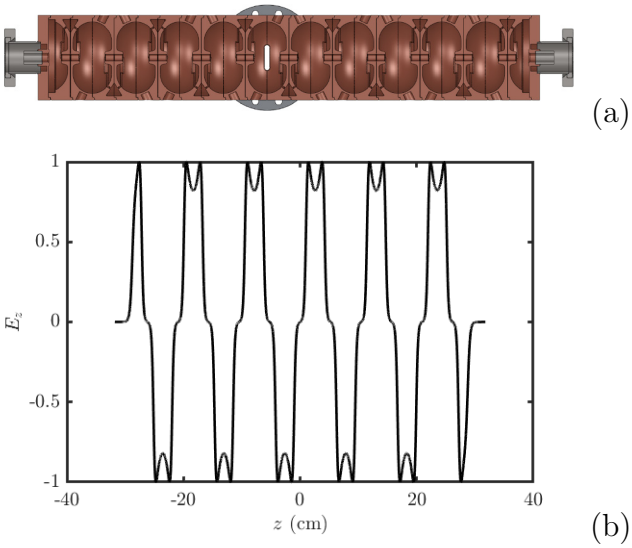


Figure 3.17: (a) Geometry of the DSRL accelerating structure. Figure taken from Ref. [33]. (b) On-axis accelerating field distribution for the DSRL accelerating structure.

The DSRL features an innovative dual-slot coupling mechanism, where two resonant slots separated by a non-resonant void region are located within the walls between adjacent cells. This configuration provides strong cell-to-cell coupling, enhancing the structure's efficiency and compactness. The design overcomes the low-velocity limitations of single-slot systems by introducing an additional sign change in the electric field between slots, resulting in a 180 degree phase advance in the on-axis accelerating field when driven in a $\pi/2$ mode. This mode ensures that the slots remain unexcited, thereby minimizing RF losses and enhancing the shunt impedance. The shunt impedance of the

DSRL installed on the PEGASUS beamline is approximately 85 MΩ/m.

3.3.2.1 Comment on RF transverse focusing

Fast, reliable, optimizable MATLAB-based transfer maps aid beamline operation. First-order transfer map descriptions are unique to each beamline, while the method to achieve accurate evaluations is generalizable in terms of axial field data[103, 104]. Ensuring compatibility with the higher-order particle tracking codes and the actual beamline requires precision and accurate axial field data. This section comments on future work developing RF transfer maps based on axial derivatives. In MATLAB, with the potential of composing RF maps with the other beamline optics (i.e., solenoids, drifts and quadrupoles) accurate beamline models become quickly optimizable functions of the RF cavity's phase and amplitude, and potentially other important experimental parameters, which target ideal operational settings of the accelerator.

To develop the map, work in lab coordinates, where z references an axial point along the beamline. Whether or not the near speed of light approximation is valid, generally writing $t(z) = \int_{z_0}^{z} d\tilde{z}/c\beta(\tilde{z})$, we can still derive the first order paraxial equations of motion assuming that the on-axis energy gain has been obtained. The relativistic factor $\gamma(z)$ fully defines the on-axis velocity $c\beta(z)$ and $t(z)$. The transverse force is purely in the radial direction and given by $\dot{p_r} = e(E_r - c\beta B_\theta)$ Hence in the x-direction: $\dot{p_x} = -\frac{e}{2}\left(\frac{\partial E_z}{\partial z} + \frac{\beta}{c}\frac{\partial E_z}{\partial t}\right)x$. Thus, applying the result from chapter 2 regarding general linear forces, we can obtain the transverse equation of motion; $x'' + \frac{p_c'}{p_c}x' + \frac{e}{2p_0c\beta}\left(\frac{\partial E_z}{\partial z} + \frac{\beta}{c}\frac{\partial E_z}{\partial t}\right)x = 0$ [89],[105]. In reduced form, the ray equation is:

$$X'' + \left[\frac{\gamma^2+2}{(\gamma\beta)^4}\left(\frac{e}{2mc^2}\right)^2 E_z^2 - \frac{e}{2mc^2(\gamma\beta)^3}\frac{\partial E_z}{\partial t}\right]X = 0 \tag{3.59}$$

This form makes apparent the relevant oscillator strength to solve the hill's equation for the problem, and the transfer matrix can be constructed. When appended with a matrix form representing the longitudinal coordinates, full 6D transport maps of the linac structures can be made.

3.3.3 Installation of X-band harmonic linearizer

In collaboration with UCLA, Radiabeam technologies developed the X-band linearizer for future use in PEGASUS as an energy spread suppressor for the UTEM instrument.

The X-band linearizer consists of 7 cells of an X-band (9.6 GHz) accelerating section. Each cell features a side-coupled geometry with noses, and the relative phase velocity is $\beta = 0.9922$. The beam pipe radius is 4 mm. The noses on the accelerating cells enhance the accelerating voltage. Coupling cells are included to ensure the structure operates in the $\pi/2$ mode, which has a larger frequency separation from neighboring modes than the π-mode, thereby increasing the fabrication tolerances of the structure. The accelerating structure is designed to achieve an accelerating voltage of 300 kV with an input power of less than 10 kW. Power is delivered to the accelerating section through a WR90 waveguide and coupling slot. Circulator losses are estimated to be between 0.3 and 0.5 dB.

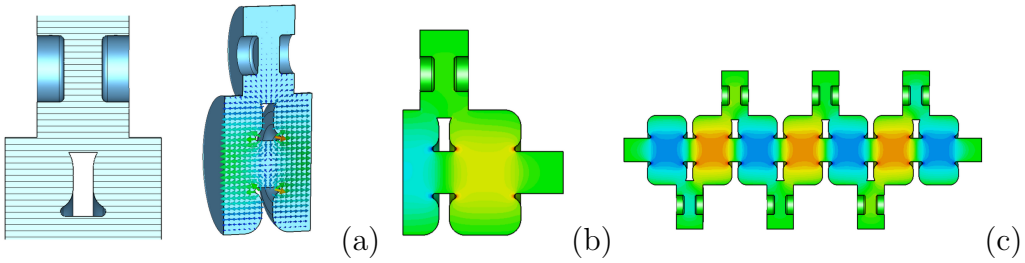


Figure 3.18: (a) One period of the side-coupled accelerating structure. (b) End cell of the structure. (c) Full 7-cell model of the X-band cavity.

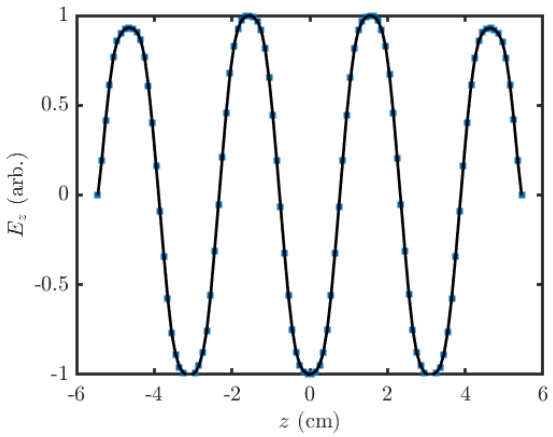


Figure 3.19: On-axis accelerating field distribution for the 7-cell X-band cavity model.

One period of the side-coupled accelerating structure is shown in Fig. 3.18(a). Both

the accelerating and coupling cells are tuned to the same frequency of 9.6 GHz. The main RF parameters are listed in Table 3.2. The coupling between cells was tuned to 2.6% to achieve a 30 MHz mode separation between the operating $\pi/2$ mode and neighboring modes, thereby increasing the geometry tolerances. The shunt impedance is optimized to 90 MΩ/m to achieve a 300 kV accelerating gradient with a source power below 10 kW.

To ensure the proper design of the accelerating cells and estimate the accelerating gradient, a full 7-cell model was simulated (Fig. 3.18(c)). The operating frequency was confirmed as 9.6 GHz, and the on-axis accelerating field distribution is shown to be flat (Fig. 3.19).

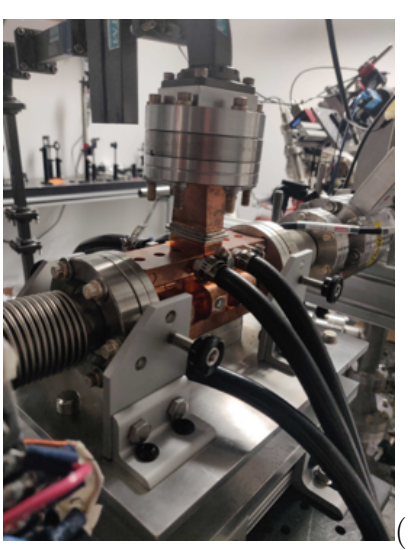
By integrating the accelerating field along the structure length, including the transient time factor corresponding to $\beta = 0.9922$ and normalizing the voltage to 300 kV, the required power should be around 8.4 kW.

Table 3.2: Accelerating Cell RF Parameters

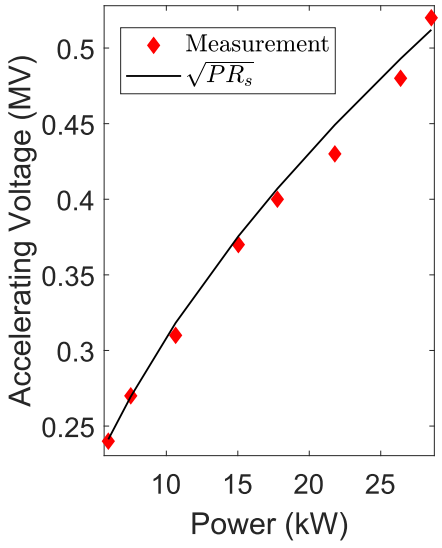
Parameter	**Value**
Resonance frequency (GHz)	9.6
Q-factor	9000
Shunt impedance (MΩ/m)	93
Mode separation (MHz)	30

The installation of the X-band cavity is shown in Fig. 3.20. The synchronization of the X-band linearizer with the S-band gun and linac system is essential for ensuring phase coherence across the RF components. Part of the low-level 2.856 GHz seed used to drive the S-band klystron is frequency divided by 36 to 79.33 MHz, then split into two paths. One fraction powers the synchrolock system to synchronize the laser with the RF system. The other portion is frequency multiplied by 121, pre-amplified to 1 W, and further amplified up to 40 kW with a small X-band klystron, phase-locking the X-band system to the laser at the operating frequency of 9.599 GHz.

The Xband linearizer's shunt impedance was measured using the dipole positioned immediately after the linac; genric low-charge operating conditions were used and centroid tracking was sufficient. Specifically, the accelerating voltage was determined from the



(a)



(b)

Figure 3.20: (a) X-band cavity installation at PEGASUS. (b) Shunt Impedance measurements of X-band cavity.

energy gain of the beam at the maximum accelerating phase as the input power was varied. The results, illustrated in Fig. 3.20, show a shunt impedance of 8.69 MΩ, which aligns well with the design specifications of shun impedance per unit length and cell length presented earlier. Based on the klystron power budget, the maximum achievable field in the cavity is 10 MV/m.

Amplitude stability was quantified by setting the X-band to the maximum accelerating phase, where the beam energy is linearly dependent on the cavity amplitude and essentially insensitive to the phase. Over one hundred shots, the rms amplitude fluctuations were 2×10^{-3}. Phase fluctuations were quantified at the zero crossing, where the beam energy changes linearly with phase offset. After removing long-term drift, rms phase fluctuations were estimated to be 1°. For the intended applications of the X-band cavity, the phase fluctuations of this scale will have a greater impact on stable compensation than amplitude fluctuations. Therefore the development of a phase feedback system is paramount to the future success of this element on the beamline.

3.3.4 Demonstration of X-band harmonic linearization

For the demonstration, a NaKSb photocathode was used and illuminated by a 266nm drive laser. Note that photoemission from NaKSb cathodes using a UV is well above the

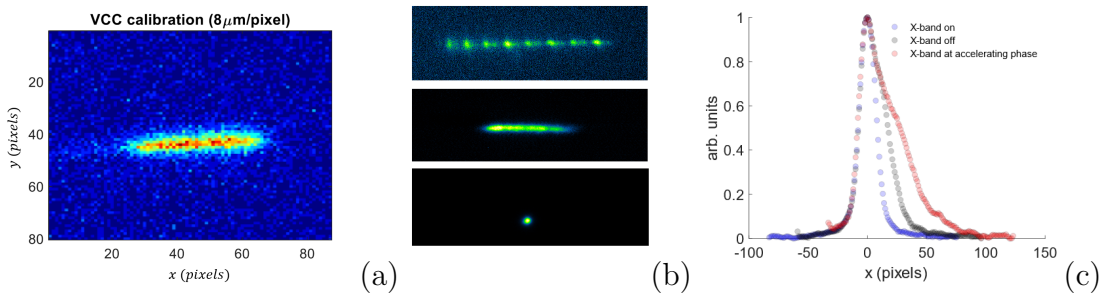


Figure 3.21: (a) The laser profile measured on the virtual cathode camera for the energy spread minimization setup. (b) in the top figure, spectrometer 1 screen measurement of 8 beamlets formed after inserting all 3 α-BBO crystals. In the middle panel, spectrometer 2 measurements are taken of the beam without optimal energy spread compensation, and the beamlets have been smeared together by using the prism and MgF_2; in the bottom panel, the optimum beamsize on spectrometer 2 is shown with compensation to be a round spot suggesting the beam is emittance dominated. (c) Shows a comparison of horizontal projections taken at spectrometer 1, with x-band on, off, and at accelerating phase while linac was turned on to compensate the linear chirp.

photoemission threshold, which ultimately increases the MTE to around 1eV. The drive laser was focused by a 175 mm focal length lens through a 72 degrees oblique incidence vacuum port. The oblique port allows the final lens to be brought closer to the cathode surface [92], but the illumination of the cathode is asymmetric. Fig. 3.21(a) shows the laser profile from the virtual cathode in the setup.

To stretch the temporal profile of the laser from the initial 40 fs RMS temporal duration to a nearly uniform flat-top $>$ 10 ps long pulse, we use 3 α-BBO crystals with respective lengths 8.75mm, 4.375mm, 2.1875mm, which splits the single 40fs beam into 8 beamlets spanning 13.3 ps [106]. In Fig. 3.21(b), the top panel shows the effect on the beam when measured on a spectrometer while having all 3 of these crystals shaping the drive laser. The pulses are then sent through a MgF_2 crystal and a prism to further stretch each pulse and facilitate temporal overlap. The middle panel of Fig. 3.21(b) shows an example of this smoothing on the energy spectra; this was measured on the beamline's second spectrometer. The bottom panel of Fig. 3.21(b) shows the energy spectrum on the beamline's second spectrometer. Given that the beam is round after the dispersion arm, the horizontal extent of the beam is at a resolution limit, which in this

case is limited by the beam's transverse emittance, not the energy spread.

The presence of second-order curvature in the longitudinal phase space manifests on the spectrometer projection as a skewed distribution. The suppression of skewness correlates to removal of the curvature. To take advantage of this fact, the Linac phase was tuned to remove linear chirp, then we made a comparison between three cases; a compensation phase of the xband, the accelerating phase of the xband to intentionally give a small second order correlation, and finally the xband was turned off. By comparing the projections on the screen, we could visualize the removal of second-order correlations and reduction in energy spread by using the xband. In Fig. 3.21(c), the projections on the spectrometer 1 screen for each case are shown, where the removal of the second-order correlation is apparent. Unfortunately, due to the emittance of the beam, the spectrometer could, at best, resolve the relative energy spread down to 3×10^{-4}. Without a smaller beam emittance, the resolution on both beamline spectrometers would be limited above 10^{-4}.

In our second spectrometer line located nearly 4 m downstream of the PEGASUS linac, we placed a pinhole with radius $r = 50$ µm , 0.5 m upstream of a round-pole dipole magnet. The pinhole reduces the spot size and subsequently the emittance over an order of magnitude at the cost of charge. There are two quadrupoles between the pinhole and dipole which enable the full system to image the pinhole onto the detector.

Theoretically, for a final dispersion length L, dipole bend radius ρ and bend angle θ the dispersion is denoted by $R_{16} = \rho(1-\cos(\theta))+L\sin(\theta)$. In the thin lens approximation, if a defocusing lens with focal length f is positioned in the dispersion arm between the dipole and the screen such that it maximizes dispersion, then the expression characterizing dispersion becomes:

$$\tilde{R}_{16} = R_{16}\left(1 + \frac{L + \rho\tan(\theta/2)}{4f}\right) \tag{3.60}$$

So we positioned the final quadrupole at this optimal in the final 0.5 m dispersion arm. The entire layout of the modified second spectrometer is shown in Fig. 3.22(a). With these additions, spectrometer 2's dispersion reaches $1.08m$ when the final quadrupole is fully powered.

With the dispersion magnifier quadrupole fully powered, the beam sizes measured

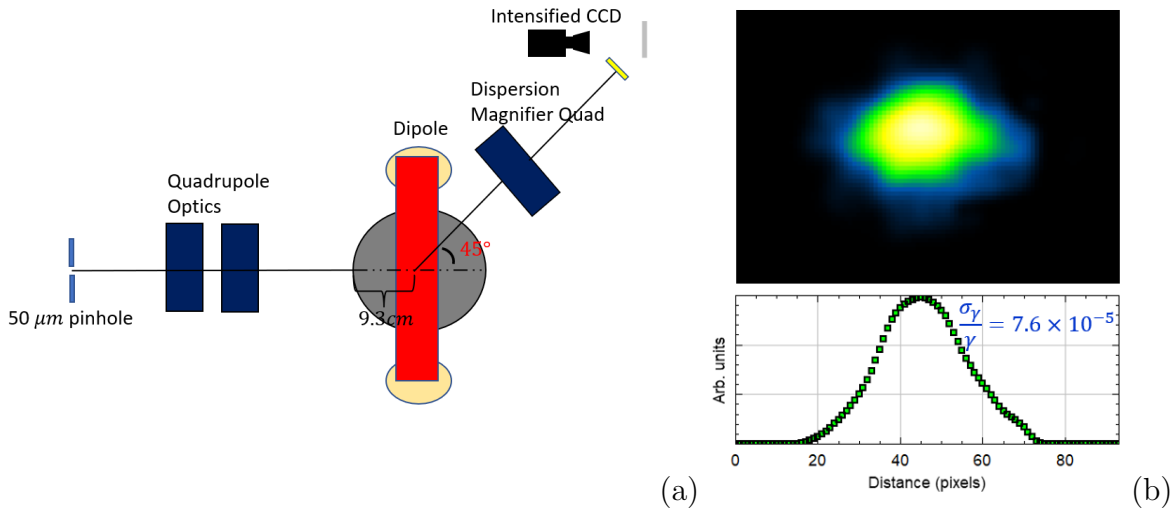


Figure 3.22: (a) Shows a top-down cartoon view of the second spectrometer layout with a pinhole and quadrupole optics. (b) Shows the best shot having an RMS relative energy spread of 7.6×10^{-5}

typically had an RMS size of 10 pixels. Given that the calibration of the screen was $8.4\mu m/pixel$ and the dispersion value $R_{16} = 1.08m$, we can estimate the energy spread as $\sigma_\gamma/\gamma \approx \sigma_x/R_{16} = 7.6\times10^{-5}$. In reality, there is the additional contribution to the spot size from the finite size of the pinhole, provided the system is imaging, and the magnification of the system is known, then the spot size is determined by $\sigma_x = \sqrt{(Mr/2)^2 + (R_{16}\sigma_\gamma/\gamma)^2}$, where M is the magnification, and r is the pinhole radius. With $r = 50\ \mu m$ and a demagnification factor of 0.64, the correction to the relative energy spread is of order 10^{-6} and is neglected.

To fully characterize the longitudinal phase space and any residual correlations, it would have been beneficial to have an X-band deflecting cavity operational upstream of the dispersion arm. This addition would enable longitudinal phase space tomography and overall better characterization of higher-order phase space correlations. However, at the time, we could not successfully split klystron power with independent phase control over both cavities. Future work will enable this feature and facilitate a thorough analysis and characterization of beams having nominal longitudinal phase spaces.

Table 3.3: Simulation of experimental setup

Parameter	**Value**
Beam Kinetic Energy	4.2 MeV
Relative Energy Spread	2×10^{-5}
Normalized Emittance	130 nm·rad

3.3.5 Start-to-end simulations of X-band harmonic linearization

Employing GPT, An injector simulation was conducted, using `spacecharge3Dtree` option [85], to capture not only smooth space charge effects but also intrabeam scattering. The beam was initialized with a mean transverse energy (MTE) of 1 eV to account for the lower work function of the NaKSb photocathode photoexcited still by UV; a bunch charge of 250 fC, a transverse rms size of 50 µm, and a duration of 10 ps (FWHM). The initial laser profile was created as a superposition of Gaussians, replicating the expected profile after passing the drive laser through α-BBO crystals for pulse stacking.

In Table 3.3, we list the results of the simulation at a screen 3 m downstream of the cathode. The beam's LPS at this location is shown in Fig. 3.23. Notably, all RF correlations other than 0th order have been removed from the LPS. The remaining spread in energy is mainly due to smooth space charge effects. To minimize these effects, the initial temporal profile should be free of temporal modulations, and the rise and fall times of the laser pulse should be sharp. Transverse uniformity of the electron beam is difficult to maintain because the beams inherit a Gaussian angular spread from the photocathode at emission. The outlying particles in the LPS are due to Coulomb collisions, which overall are found to contribute negligibly to the energy spread in this case.

3.3.6 Start-to-end Pegasus beamline compression simulations

Building on the theory and ideal simulations of beam compression focused on validating those scalings, we now apply these principles to the Pegasus beamline in a realistic start-to-end simulation, shaping the beam from the source towards optimums indicated by the envelope analysis.

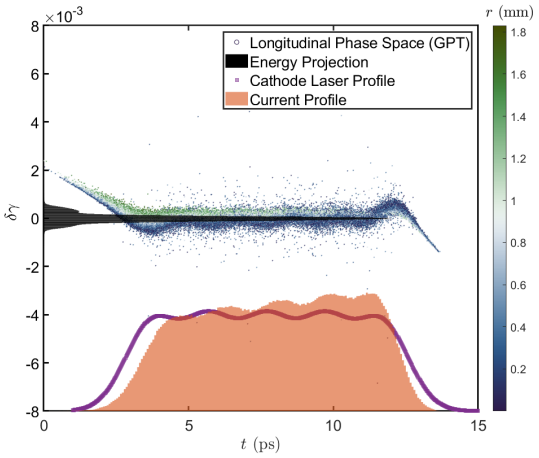


Figure 3.23: Longitudinal phase space 3 m downstream of the gun, after the linearizer and linac cavities, where the beam coasts with 2×10^{-5} relative energy spread. The LPS is color-coded with particle radius from the core. Energy and temporal projections are shown along the corresponding axes. The input laser profile (initial temporal distribution) is also shown.

A spot size of 10 μm is initialized at the cathode with an MTE of 1 eV. An initially quadratic pulse with a 4.5 ps RMS duration with 500fC of charge. The gun phase is set to 25 degrees with peak accelerating gradient of nearly 70 MV/m which brings the beam up to 3.2 MeV. The LINAC and Xband are optimized to achieve a focus inside the vacuum box where there is ample room for potential diagnostics.

The dynamics of the bunch lengths (transverse and longitudinal) and beam energy are plotted in Fig. 3.24. The X-band phase is set to the decelerating phase ($\phi_2 = 0$) with an amplitude of 8.6 MV/m, compensates to third order the LPS distortions, attaining a final bunch length of 1.4fs, which in this case turns out to be limited from achieving a sub-fs duration by some space charge driven emittance growth. By going to lower charge sub 200fC, the emittance growth is relaxed sub fs is achievable with the same RF setpoints.

The transverse and longitudinal dynamics of the beam are largely independent, except near a simultaneous waist. This independence allows the beam to be focused tightly transversely to several microns without significantly affecting the longitudinal focus, provided the transverse beam quality is high. However, there are some practical parasitic limitations. For instance, stray quadrupole moments, which may not be aligned with the laboratory axes, can degrade the transverse emittance, impacting both transverse

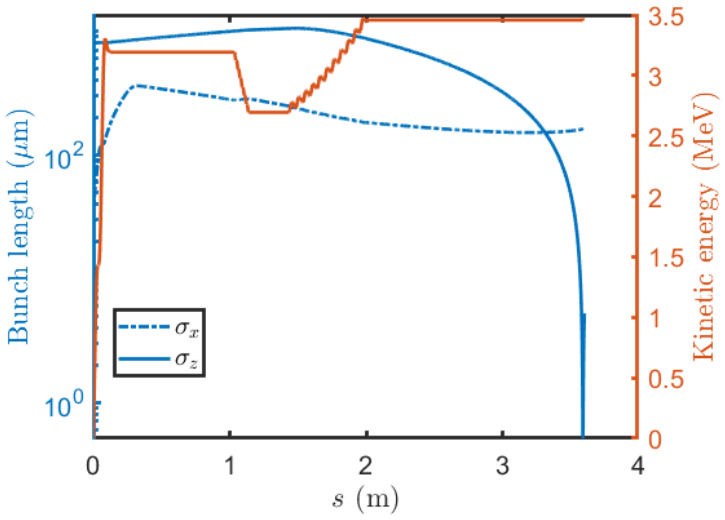


Figure 3.24: Evolution of the bunch length and beam energy from the cathode up to the focus.

and longitudinal focusing quality. In principle, a skew quadrupole can be utilized at the gun's exit to counteract this emittance growth, but the precise details are not known a priori. Additionally, minor asymmetries in the RF elements, such as those from the RF coupling ports, can lead to transverse centroid kicks that must be corrected with steering magnets. These effects are not accounted for in the ideal case but are anticipated to have a somewhat significant impact on the optimum bunch length.

If the beam size at the final focus lens is too large, time-of-flight effects can limit the achievable bunch length. Specifically, the longitudinal focus positions of the particles become correlated with their initial transverse positions at the lens, as detailed in [107]. To address this, the injection solenoid is employed to collimate the beam and maintain a narrow waist, keeping the beam size below 0.5 mm up to the longitudinal focus. This minimizes time-of-flight correlations, which is crucial when both transverse and longitudinal focuses are needed simultaneously.

With these considerations in mind, we explore the limitations to achieving a fully compressed beam with a unit aspect ratio in a realistic beamline setting. Initially, we utilize the green quadrupole triplet, which has an effective focal length of about 25 cm. The envelope evolution is depicted in Fig. 3.25(a), along with the final longitudinal phase space (LPS) in Fig. 3.25(b). The uncorrelated energy spread limits the final phase space, with some additional temporal broadening due to time-of-flight correlations from

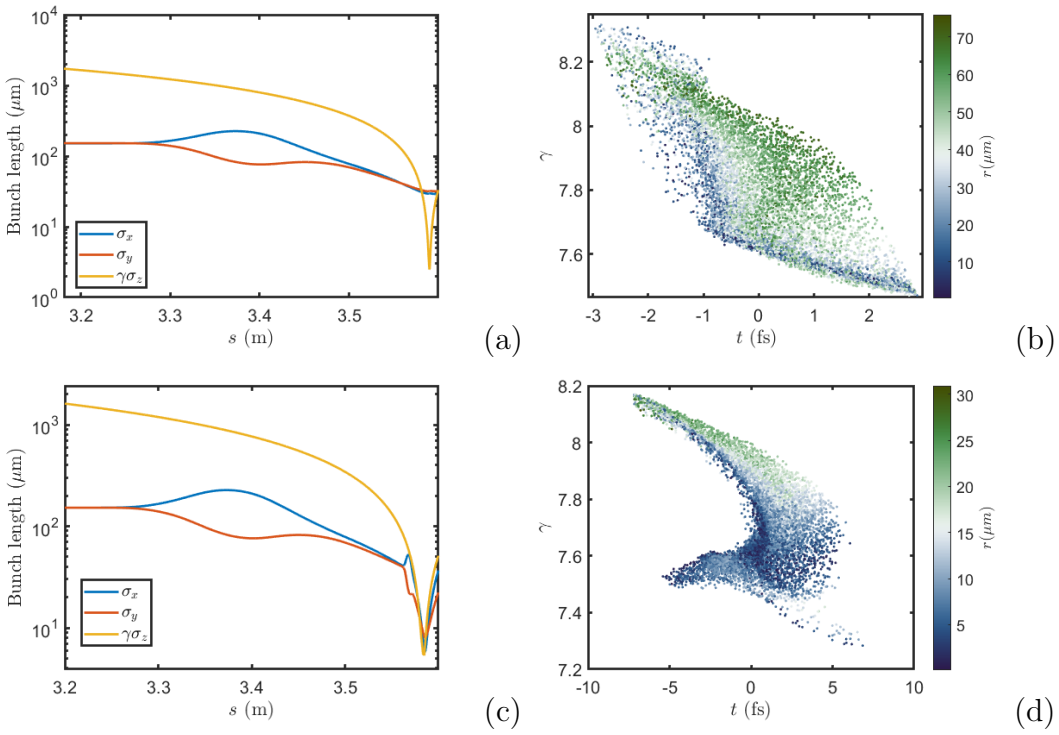


Figure 3.25: (a) The green electromagnetic quadrupole triplet focus the beam into the box to 25 μm while achieving nearly the same 1.4 fs bunch length. (b) The longitudinal phase space in the vacuum box. (c) An additional PMQ triplet is added to nearly achieve a unit aspect ratio. (d) The LPS at the interaction point with a unit aspect ratio.

the electromagnetic triplet lens. In this case, the transverse focusing is not enough to become fully space charge dominant. The transverse size of 25 µm is still limited by the transverse emittance.

Finally, we can employ a permanent magnet quadrupole triplet to achieve the desired degree of focusing into the space charge dominated regime. The triplet utilized here is based on the design presented in Chapter 7, with minor modifications from the recent 8.5 MeV diffraction experiment at Pegasus. This triplet comprises a 6 mm quadrupole placed between two 3 mm quadrupoles to obtain a symmetric focus.

In this scenario, the envelope evolution is depicted in Fig. 3.25(c), and the final longitudinal phase space (LPS) in Fig. 3.25(d). The LPS exhibits correlations stemming from residual non-linear space charge forces and some time-of-flight distortions, largely aligning with the previous analytical discussion. However, in this case, the second-order RF emittance growth is compensated. Therefore, the optimum starting bunch length is

determined by equating the space charge dominant regime with the uncorrelated energy spread result, suggesting that the final bunch length should be limited by the uncorrelated energy spread.

Of course, this is not exactly what we achieve because the analysis neglected transverse correlations and the potential for non-linear space charge forces because we assumed the fields are linear, so the model is not fully kinetic. Underfocusing transversely makes it easier to achieve the thermal energy spread limit. These realistic effects offset the analytically ideal case, but the approach remains effective when RF non-linearities are not fully compensated, for instance, if the beamline lacks access to a cavity of a higher harmonic than the buncher. It is easier to achieve the uncorrelated energy spread limit when not utilizing transverse focusing, which makes sense.

While this work shows promise in producing short bunch lengths, the challenge of temporally resolving sub-femtosecond (sub-fs) bunch lengths in practice remains. Achieving a tight transverse focus is critical for most techniques that would be used to time-stamp the bunch, such as coherent transition radiation (CTR) or THz streaking [108, 109]. These methods are currently the most viable candidates for measuring the bunch length at Pegasus.

In the case of CTR, the bunch length is comparable to optical wavelengths. However, the transverse extent of the beam suppresses the CTR's spectral content far into the infrared (IR) range unless the beam can be focused to achieve a unit aspect ratio at the point of impingement [110]. A potential solution to reestablish the spectral content involves using a tilted foil structure, which can remove the dependence on the transverse form of the beam [111].

Another proposed technique is coherent Thomson scattering, which provides feedback on the quality of beam compression, as the degree of coherence is a direct measure of the bunching quality [112].

Alternatively, a THz streak camera offers a promising solution. By using a portion of the drive laser to generate the THz field that excites the streaking mode, synchronization is simplified. These streak cameras have proven effective in measuring sub-10 fs bunches and hold the potential for achieving sub-fs resolution [113].

CHAPTER 4

Space-Charge aberrations in axially symmetrical lenses for time-resolved electron microscopy

4.1 Introduction

While static snapshots of samples at nanometer and subnanometer resolutions offer substantial insights, introducing time as a fourth dimension to electron microscopy promises groundbreaking advancements. This significant enhancement would enable researchers to observe how sample structures, compositions, and properties evolve in response to external stimuli [8]. This chapter delves into the specific challenges associated with this advancement, focusing on quantifying the trade-offs between temporal and spatial resolution that arise from collective space charge effects in modern microscopes.

When a sample's dynamics are irreversible and cannot be continuously replicated, it becomes necessary to utilize single-shot illumination—deploying all electrons in one bunch, with the bunch's duration determining the microscope's exposure time[114, 11, 115]. Currently, these techniques are confined to nanosecond-scale studies. However, extending resolution into the 10-100 ps range would enable the examination of phenomena such as defect motion and dislocation dynamics in materials science, as well as the effects of material heterogeneity in ultrafast phase transitions. In biology, advanced ultrafast high-energy TEMs could observe cells and viruses in aqueous environments. Single-shot illumination is particularly vital in scenarios where electron dosage might preclude repeated exposures on the detector [116]. In these single-shot scenarios, the peak current is dramatically higher than in conventional TEM modes. For instance, 10^8 electrons in a 10 ps bunch equates to peak currents exceeding 1 A, under which space-charge effects can severely impair image quality. Recent advancements suggest that elevating the electron beam's energy to the megaelectron-volt range can leverage relativistic effects to mitigate

space-charge forces, thereby enhancing spatial resolution in the microscope [43, 117].

Two primary space-charge interactions—smooth mean field effects described in chapter 2, and binary collisions—crucially affect electron beam distributions in transmission electron microscopes (TEMs). Notably, stochastic Coulomb interactions scale with the square root of the charge density [118], while smooth space-charge effects increase proportionally with beam current. Therefore, in high-current scenarios typical of single-shot TEM applications, smooth space-charge effects tend to dominate beam dynamics. Although existing studies [118, 119, 43] primarily use numerical methods to investigate how stochastic binary collisions affect image quality in standard TEM configurations, their focus on specific setups limits their broader applicability. Developing analytical formulas and scaling laws would enable the rapid assessment of spatiotemporal resolution limits across diverse microscope configurations, differing markedly in electron energies and peak currents.

This chapter further develops the analytical framework to calculate space-charge-induced aberrations in the relativistic regime. We validate our approach with particle tracking simulations [120] that include smooth and stochastic binary effects. We use a perturbation method to assess how a smooth space-charge field, acting as a defocusing lens, affects electron trajectories and introduces aberrations similar to those in magnetic lenses, as noted in studies by Hawkes and others [121, 122]. This study quantifies these effects and examines the spatial and temporal resolution trade-offs at various beam energies. Additionally, we discuss compensation techniques for optimizing beam distribution and the potential use of multipole electron optics for correction, although their complexity may limit practical implementation.

The chapter is structured as follows: We examine the single particle dynamics of charged particles in axially symmetric magnetic focusing systems up to third order. Employing Green's function method discussed in Chapter 2, we evaluate aberrations analytically. We compare our results with particle tracking simulations for scenarios without space-charge effects, specifically addressing spherical and chromatic aberrations in magnetic round lenses. We then explore the impact of space-charge-induced nonlinearities on particle transport, developing analytical expressions for space-charge aberration coefficients. While dependent on the shape of the initial beam phase space, these formulas

provide useful estimates for aberrations in straightforward cases such as uniform and Gaussian distributions, enabling quantitative analysis of the trade-offs between spatial and temporal resolution in various electron column designs. The insights from this study aim to guide the development and performance expectations of single-shot time-resolved TEM and their range of scientific applications.

4.2 Single Particle Dynamics in a Solenoid Field

The goal of this research was to develop accurate analytical framework for evaluation of space charge aberration coefficients in axially symmetric systems. Before addressing the self-consistent effects, we review the dynamics of a single particle in a solenoid field, considering a first-order expansion of the magnetic field, showcasing that even if the field is linear in r, the dynamics are still of higher order. Close to the axis, the magnetic field can be Taylor expanded, characterized by:

$$\boldsymbol{B} = B(s)\hat{\mathbf{z}} - \frac{B'(s)r}{2}\hat{\mathbf{r}} \tag{4.1}$$

where $B(s)$ represents the axial magnetic field, typically measured using an axial Hall probe [123].

The particle's velocity in cylindrical coordinates is expressed as:

$$\mathbf{v} = \dot{r}\hat{\boldsymbol{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}} + \dot{z}\hat{\boldsymbol{z}} \tag{4.2}$$

Applying the Lorentz force law in the absence of electric fields, $\frac{d\mathbf{p}}{dt} = q\mathbf{v}\times\mathbf{B}$, we derive the following coupled differential equations:

$$\gamma m(\ddot{r} - r\dot{\theta}^2) = qr\dot{\theta}B \tag{4.3}$$

$$\gamma m(2\dot{r}\dot{\theta} + r\ddot{\theta}) = -\frac{q\dot{z}rB'}{2} - q\dot{r}B \tag{4.4}$$

$$\gamma m\ddot{z} = \frac{qr^2\dot{\theta}B'}{2} \tag{4.5}$$

By manipulating Equation 4.4 and multiplying by r, we convert it into a total time derivative, thereby revealing a conserved canonical angular momentum:

$$\frac{d}{dt}\left(\gamma mr^2\dot{\theta} + \frac{qr^2B}{2}\right) = 0 \implies P_\theta = \gamma mr^2\dot{\theta} + \frac{qr^2B}{2} \tag{4.6}$$

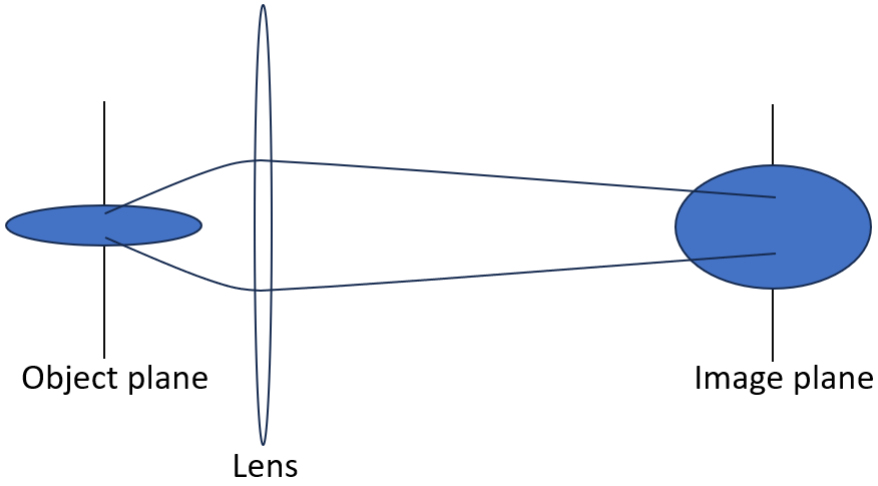


Figure 4.1: Planes (screens) and time: Aberrations are assessed by analyzing the transport between planes. In scenarios involving finite beams where collective effects are significant, it is crucial to differentiate the methods used to develop the equations of motion. One approach references particle positions relative to the beam centroid, while the other tracks a single particle, using its z-coordinate as a plane reference.

Here, P_θ represents the canonical angular momentum. Consequently, the angular velocity $\dot{\theta}$ is given by:

$$\dot{\theta} = \frac{P_\theta}{\gamma m r^2} + \omega_L(s) \tag{4.7}$$

where $\omega_L(s) = -\frac{qB}{2\gamma m}$ is the Larmor frequency, which, for particles having no initial angular momentum, represents the particle's precession rate around the magnetic field axis.

In principle, once we know the evolution of r, we also know the angular frequency of revolution of the azimuthal angle. We can eliminate $\dot{\theta}$ from the longitudinal and radial equations to obtain:

$$\ddot{z} = -\frac{P_\theta \omega_L'}{\gamma m} - \omega_L \omega_L' r^2 \tag{4.8}$$

$$\ddot{r} + \omega_L^2 r = \frac{P_\theta^2}{\gamma^2 m^2 r^3} \tag{4.9}$$

In its current form, the radial equation is independent of z and θ, which allows us to derive all subsequent equations once it is solved. However, the equation still depends

on time. To consistently compare the initial and final planes for all particles, we must shift the independent variable from t to s. This change is crucial because particles, depending on their initial conditions, traverse from the object to the image plane in varying durations. Standardizing the reference planes is essential for accurate aberration assessment. This approach contrasts subtly with our method of parameterizing particle coordinates relative to the beam centroid in Chapter 2, where one perspective focuses on a fixed screen while the other considers time, the latter being valuable for analyzing beam compression as illustrated in Fig. 4.1. Moving forward, while we develop the third-order ray equation under the assumption of a continuous beam, it is important to note that the equations of motion would differ if we continued to reference the beam centroid.

To achieve this, we normalized Equation (4.9) by $\dot{z}^2$, allowing us to write

$$\frac{\ddot{r}}{\dot{z}^2} = r'' + \frac{\ddot{z}}{\dot{z}^2}r' = -\frac{\omega_L^2 r}{\dot{z}^2} + \frac{P_\theta^2}{\gamma^2 m^2 \dot{z}^2 r^3} \tag{4.10}$$

The total velocity is fixed, and for a monochromatic beam, all the total velocities of all the particles within the beam are equal to $c\beta_0$. So, we can write

$$\frac{\dot{z}^2}{c^2\beta_0^2} = \frac{1 - r^2\dot{\theta}^2/c^2\beta^2}{1 + r'^2} \tag{4.11}$$

. we can expand to second order in r and r':

$$\frac{c^2\beta_0^2}{\dot{z}^2} \approx 1 + r'^2 + r^2\dot{\theta}^2/c^2\beta_0^2$$

This allows us to write the third order ray equation:

$$\begin{aligned} r'' &= \frac{c^2\beta_0^2}{\dot{z}^2}\left(\frac{P_\theta^2}{p_0^2 r^3} - \kappa r - \frac{\ddot{z}}{c^2\beta_0^2}r'\right) \\ &= \left(1 + r'^2 + r^2\left(\frac{P_\theta}{p_0 r^2} + \frac{\omega_L}{c\beta_0}\right)^2\right)\left(\frac{P_\theta^2}{p_0^2 r^3} - \kappa r + \left(\frac{P_\theta \omega_L'}{p_0 c\beta_0} + \frac{\kappa'}{2}r^2\right)r'\right) \end{aligned}$$

where $\kappa = \omega_L^2/c^2\beta_0$.

If we consider meridional rays, which are those that begin with no initial angular momentum, then $P_\theta = 0$, and the third-order ray equation for rays of this type can be rearranged as:

$$r'' + \kappa r = -\kappa r'^2 r + \frac{\kappa'}{2} r' r^2 - \left(\kappa^2 - \frac{1}{2}\kappa \left(\frac{B_0''}{B_0} \right) \right) r^3, \tag{4.12}$$

which agrees with the result in [**?**] for a first-order field expansion. The left-hand side of this equation represents Hill's equation, while the right-hand side consists of third-order components, which are assumed to be small for the trajectories. Therefore, they are considered perturbations, and the methodology discussed in Chapter 2 for evaluation of aberrations from these higher-order terms is applicable. Note that we have included the next highest term in the magnetic field expansion, appearing as the final term on the right ($\propto B_0''$).

4.3 Calculating lens aberrations of meridional rays

The unperturbed meridional ray equation governing the linear trajectory of an electron in a solenoid field characterized by the focusing strength $\kappa(z)$ is:

$$r'' + \kappa r = 0 \tag{4.13}$$

Here, $\kappa(z) = \left(\frac{B_0(z)}{2[B\rho]} \right)^2$ defines the focusing strength, where $B_0(z)$ represents the profile of the axial magnetic field of the lens, and $B\rho = m_0 c\beta\gamma/e$ denotes the magnetic rigidity of the beam. If the axial width of the solenoid field is sufficiently short so that r is effectively unchanged in the lens, then for a ray entering the lens having no slope, the slope at the exit is related to the lens focal length as $r'/r = -\int \kappa dz = 1/f$, where f is the focal length, r' is the exit angle, and r is the radial offset at the exit of the lens.

In Chapter 2, we derived the general solution to this equation using a symplectic map:

$$\begin{pmatrix} r \\ r' \end{pmatrix} = \begin{pmatrix} C(z) & S(z) \\ C'(z) & S'(z) \end{pmatrix} \begin{pmatrix} r_0 \\ r_0' \end{pmatrix} \tag{4.14}$$

In this formulation, $C(z)$ and $S(z)$ represent the two linearly independent solutions to Eq. 4.13, which satisfy the initial conditions $C(0) = 1$, $C'(0) = 0$, $S(0) = 0$, and $S'(0) = 1$. The behaviors of the cos-like and sine-like rays, as they are transported through the solenoid detailed in Table. 4.1, are depicted schematically in Fig. 4.2.

We now outline the numerical method employed to derive the matrix transport for arbitrary axial field distributions. The function $\kappa(z)$, representing an arbitrary bump function, can be fit with analytical functions [1]

The matrix propagator technique is employed to compute the transport map after determining the focusing strength function $\kappa(z)$. Within a narrowly defined region $(z_i - \Delta z/2, z_i + \Delta z/2)$, where $\kappa(z_i) \equiv \kappa_i$ is essentially constant, a small step size $\Delta z = L/N$ facilitates the construction of the differential mapping ΔM_i. Here, L is the total propagation distance and N is the number of steps. The differential mapping is given by:

$$\Delta M_i = \begin{pmatrix} \cos\left(\sqrt{\kappa_i}\Delta z\right) & \sin\left(\sqrt{\kappa_i}\Delta z\right)/\sqrt{\kappa_i} \\ -\sqrt{\kappa_i}\sin\left(\sqrt{\kappa_i}\Delta z\right) & \cos\left(\sqrt{\kappa_i}\Delta z\right) \end{pmatrix}.$$

The overall transport matrix M for the entire stage of length L is then calculated as:

$$\begin{pmatrix} C(z) & S(z) \\ C'(z) & S'(z) \end{pmatrix} = \lim_{N\to\infty} \prod_{i=0}^{N-1} \Delta M_{N-i}.$$

This formulation ensures that the cumulative effects of focusing across successive small intervals are integrated to yield the complete transport map for the specified transport interval.

For optimal focusing and imaging over a distance L, the conditions $C(L) = M$ and $S(L) = 0$ must be satisfied, where M signifies the magnification factor. These conditions are critical to ensure that the transport map facilitates imaging, effectively rendering the final position a magnified, angle-independent replica of the initial position.

Throughout the discussion, we will allow the space-charge forces whenever present to modify the linear transport, but notably assume that the nonlinearities associated with deviation from design energy, large angles, space-charge, etc., induce only small image plane deviations and the transport can be well approximated by the first-order linear optics [124].

To analyze aberrations, we apply the Green's function method outlined in Chapter 2 to evaluate aberrations and solve the driven Hill equation governing electron transverse motion through an objective lens. At the imaging plane, where $S(L) = 0$ and $C(L) = -M$

[1] For example, as a superposition of Gaussians: $\kappa(z) = \sum_{n=1}^{N} a_n \exp\left(-\frac{(z-b_n)^2}{c_n^2}\right)$, where a_n, b_n, and c_n respectively denote the amplitude, center, and width of each Gaussian component.

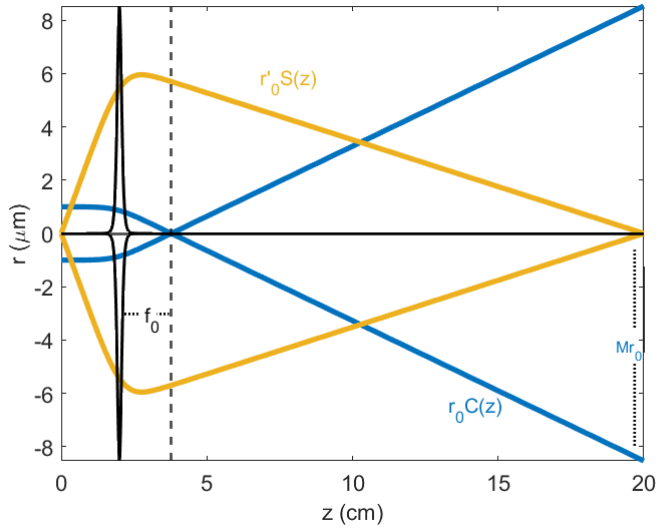


Figure 4.2: Sin-like and cos-like trajectories within a solenoid imaging transport system, adjusted to be displayed on a unified set of axes.

(with M representing the system's magnification), the Green's function for the image plane takes the form $\mathcal{G}(L,s) = MS(s)$. This formulation allows us to express image plane deviations as:

$$\delta r(L) = M \int_0^L S(s) f(s) ds \tag{4.15}$$

4.3.1 Chromatic aberration example

The first instance of image plane deviation we explore arises from chromatic aberration. If a particle's energy is slightly higher or lower than the design value for the optical column, it will experience a focusing kick slightly weaker or stronger than the reference particle. For a small relative momentum deviation $\delta p/p$, the corresponding focusing strength on this particle can be approximated as:

$$\kappa \left(\frac{1}{1 + \frac{\delta p}{p}} \right)^2 \approx \kappa \left(1 - 2\frac{\delta p}{p} \right) \tag{4.16}$$

At first order, the equation for the deviation from the reference trajectory becomes:

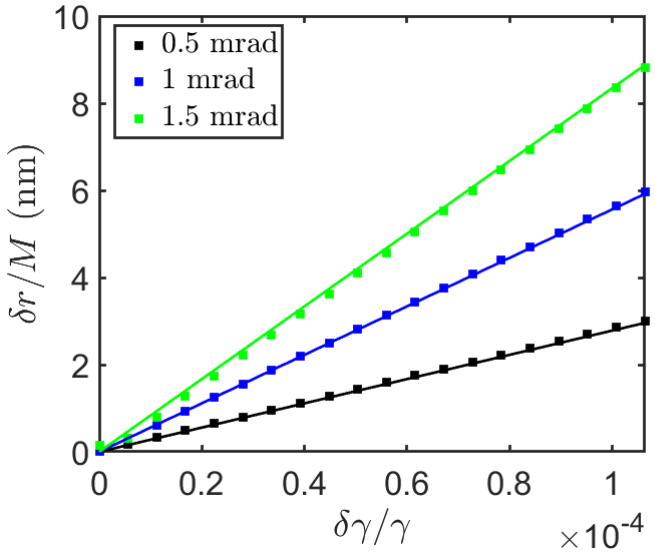


Figure 4.3: GPT output of image plane deviations normalized by magnification for three different deflection angles, plotted as a function of rms energy spread.

$$\delta r'' + \kappa \delta r = 2\kappa \frac{\delta p}{p} r_c \tag{4.17}$$

This equation can be solved using the Green's function method described earlier. For sine-like reference trajectories $r_c(z) = r_0' S(z)$, the image plane deviation is given by:

$$\delta r(L) = M r_0' \frac{\delta p}{p} \int_0^L \frac{\kappa S(s)^2}{2} ds = M r_0' \frac{\delta p}{p} C_c \tag{4.18}$$

This expression allows us to evaluate the chromatic aberration coefficient C_c. In Table 4.1, we list the nominal optical and beam parameters for the simulation results presented throughout this paper, unless indicated otherwise. In this analysis, the model equation for the on-axis field of the solenoid lens is given by:

$$B_0(z) = \frac{\mu_0}{2} \frac{\mathrm{NI}}{d} \left(\frac{z + \frac{d}{2}}{\sqrt{\left(z + \frac{d}{2}\right)^2 + R^2}} - \frac{z - \frac{d}{2}}{\sqrt{\left(z - \frac{d}{2}\right)^2 + R^2}} \right) \tag{4.19}$$

Here, $d = 0.015$ m, $R = 0.008$ m, and $N = 1720$. The resulting lens has an effective length of 1.4 cm and, at around 20 amps, a focal length of 1.5 cm for a 4.3 MeV beam. The lens images over a distance of 0.2 m with a magnification of $\times 8.5$.

In Fig. 4.3, the chromatic aberration coefficient $C_c = 3$ cm, calculated using Eq. 4.18,

aligns remarkably well with the imaging plane deviation derived from General Particle Tracer (GPT) simulations [120]. These simulations were conducted with space-charge effects disabled, across different beam divergence angles and input energy spread levels. It's important to note that we normalize the image plane deviations by the magnification factor in all subsequent plots to establish their relationship with object plane distances.

Table 4.1: Nominal parameters for single solenoid lens stage GPT simulations.

Parameter	**Value**
Full Width Pulse Length	10 ps
E-beam kinetic energy	4.3 MeV
E-beam charge	250 fC
Peak Dose at sample	0.5 e/nm^2
Spotsize/Edge Radius	1 μm
Beam Divergence	3 mrad
RMS Energy Spread	$< 10^{-5}$
Lens Focal Length	1.5 cm
Object to Image plane distance	20 cm
Magnification	8.5

When energy spreads approach zero, a slight inconsistency between the analytical prediction and the simulation results becomes evident. This disparity arises from the imperfect imaging, even in the absence of chromatic aberrations, attributable to the finite beam divergence and the presence of the spherical aberration term. We delve into this analysis in the subsequent section.

4.3.2 Spherical aberration example

Whenever the energy spread of the beam can be kept sufficiently low to minimize the chromatic effects, the main contributions to the trajectory deviation from the ideal imaging condition will be associated with the radial dependence of the focusing field in magnetic round lenses (spherical aberrations). Non linear effects arise due to the longitudinal velocity variation through the lens and higher order terms in the magnetic field components.

Following the description in Reiser [46] (and again assuming that these terms can be treated as perturbation), we can write for the driven Hill equation:

$$\begin{aligned}\delta r'' + \kappa \delta r &= -\kappa r_c'^2 r_c + \frac{\kappa'}{2} r_c' r_c^2 - \left(\kappa^2 - \frac{1}{2}\kappa\left(\frac{B_0''}{B_0}\right)\right) r_c^3 \\ &= f(r_c, r_c', B_0)\end{aligned}$$

For a general trajectory having initial position and angle offset $r_c = r_0 C + r_0' S$ the deviation at the image plane due to the radial dependence of the lens field can be written as:

$$\delta r(L) = M\left(r_0^3 C_p + 3 r_0^2 r_0' C_q + 3 r_0 r_0'^2 C_r + r_0'^3 C_s\right) \tag{4.20}$$

where the coefficients with subscripts p, q, r and s are related to the classical distortion, coma, image curvature and spherical aberration terms respectively [121].

The spherical aberration coefficient C_s can be extracted by setting $r_0 = 0$. Then convolution with the Green function at the imaging plane yields:

$$\delta r(L) = M r_0'^3 \int_0^L S^4 \left(\frac{\kappa'}{2}\left(\frac{S'}{S}\right) + \frac{1}{2}\kappa\left(\frac{B_0''}{B_0}\right) - \kappa^2 - \kappa\left(\frac{S'}{S}\right)^2\right) ds = M r_0'^3 C_s \tag{4.21}$$

In Fig. 4.4 we show a comparison for the image plane deviations obtained using the analytical results from Eq. 4.21 and the numerical GPT simulation for a monochromatic beam with no space-charge (and other parameters as listed in Table 4.1). The beam distribution at the object plane is assumed uniform within a 1 μm hard-edge radius and a 2 mrad rms gaussian angular spread. The excursions from the ideal reference trajectory plotted as a function of the initial ray deflection show an excellent agreement between the calculated $C_s = 3$ cm cubic dependence and the particle tracking output. As an example, for incidence angles of 3 mrad, the particular lens employed in our simulation (i.e. Eq. 4.19) contributes to a blurring in the image plane of 0.81 nm. Similarly to the chromatic aberration coefficient C_c, it is also typical for C_s to be comparable with the lens focal length. Appropriately crafting the longitudinal on axis magnetic field profile can reduce the spherical aberration.

An astute observer would notice that while the cubic dashed line accurately captures the overall trend of the perturbed trajectories, there's an additional widening of the image

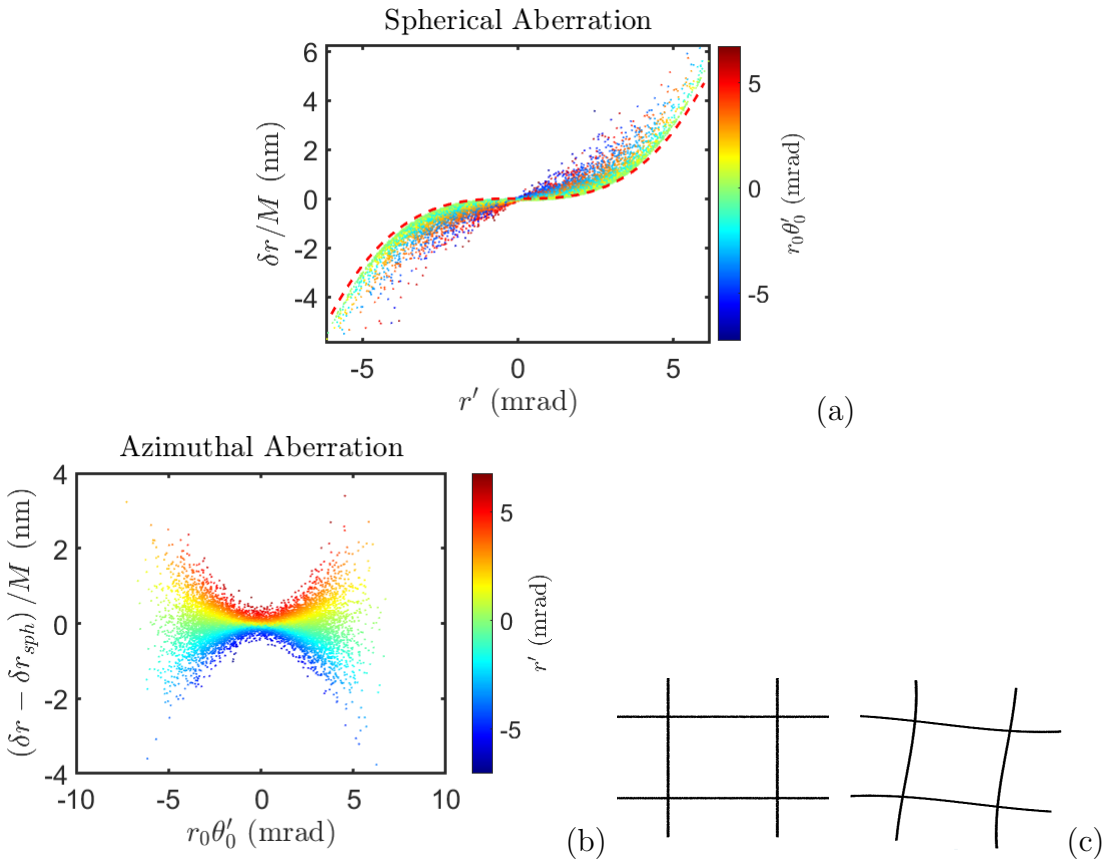


Figure 4.4: (a) Image plane deviations in a GPT simulation output without energy spread and space-charge effects. The analytical cubic function, illustrating spherical aberration, is shown with a dashed line. Particles are color-coded by initial angular velocity to demonstrate the effect of handkerchief aberration. (b) Image plane output plotted against $r_0\theta'$ after compensating for the cubic dependence of spherical aberration. (c) GPT simulation depicting the handkerchief distortion on a 10mm hash symbol with a bar width of 125 μm.

deviations noted in the particle tracking simulation. This broadening arises from another form of lens aberration and warrants a separate discussion. In our simplified treatment of electron dynamics within the cylindrically symmetric column, we've disregarded azimuthal motion. However, the combined effect of solenoid Larmor rotation and radial aberrations in Eq. 4.20 can induce a deviation in the image plane proportional to the square of the particle's initial angular velocity $r_0\theta_0'$, or alternatively the square of initial angular momentum (as indicated by the color coding in the left figure and more explicitly depicted in Fig. 4.4(b), where the image plane deviation with the cubic spherical aberration subtracted is illustrated). Additionally, we present the results of a GPT simulation

aimed at elucidating this effect, commonly referred to as "handkerchief aberration," in Fig. 4.4(c). A significantly large field of view is employed to visualize the handkerchief distortions in the image plane. Generally, this effect can be disregarded when assessing the imaging performance limit at very minute transverse offsets from the optical axis.

4.4 Space-Charge Aberration

In this section, we apply the theoretical framework developed above to estimate the third-order space-charge-induced deviations from the unperturbed trajectories.

As discussed in Chapter 2, consider an electron bunch with a very long length L_b, characterized as having a cigar-shaped aspect ratio in the beam's rest frame. In scenarios where the input charge density satisfies $\frac{1}{\rho}\frac{\partial \rho}{\partial z} \ll \frac{1}{L_b}$, the space-charge field is predominantly two-dimensional, and longitudinal components can be effectively neglected.

In such cases, for a known charge density distribution within a cylindrically symmetric optical system, the transverse electric field can be derived using the cylindrically symmetric form of Gauss' Law in integral form, as shown in Chapter 2; $E_r(r;s) = \frac{1}{r}\int_0^r \frac{\rho(\xi;s)}{\epsilon_0}\xi d\xi$, which can be off-axis expanded in axial derivatives of charge density. Here, ξ serves as a dummy variable for radial integration, while s specifies a plane that parametrizes the evolution of the transverse density. This parametrization directly correlates a plane with the radial space charge field when the beam is centered on that plane, effectively representing the average position of the charge distribution in the TEM column.

Our approach to calculating third-order space-charge aberrations is as follows:

1. Compute the evolution of the transverse charge density using the method of characteristics.

2. Once the evolved transverse density is determined, calculate the first and third-order space-charge fields using Gauss' law.

3. Integrate the third-order field, weighted by the Green's function, over the column. This integration yields the deviations from the linear trajectories at the image plane and the corresponding aberration coefficients.

Assuming that nonlinear space-charge forces represent only a minor perturbation to particle motion, the particles in the beam will evolve along characteristic orbits described by the linear transport map $w = Rw_0$, where $w = (x, x', y, y')^T$. The dynamics of this system are governed by the 4×4 symplectic matrix R, which encapsulates the linear uncoupled dynamics in the Larmor frame:

$$R = \begin{bmatrix} C & S & 0 & 0 \\ C' & S' & 0 & 0 \\ 0 & 0 & C & S \\ 0 & 0 & C' & S' \end{bmatrix} \tag{4.22}$$

Given that the dynamics are Hamiltonian, the initial distribution remains stationary in phase space. Consequently, at any point along the optical column, the distribution function can be defined as $f(R^{-1}w) = f(w_0)$, where $f(w_0)$ is the initial distribution.

Extending the kinetic approach described in Chapter 2, while disregarding longitudinal emittance, the space-charge density at any position in the optical column is calculated by integrating over the transverse momentum space:

$$\rho_{sc}(x, y; z) = \iint \frac{dudv}{C(z)^2} f\left(\frac{x - S(z)u}{C(z)}, u, \frac{y - S(z)v}{C(z)}, v\right) \tag{4.23}$$

Here, the substitutions $u = C(z)x' - C'(z)x$ and $v = C(z)y' - C'(z)y$ facilitate the integration by transforming coordinates to the initial momentum space.

In cylindrical symmetry, the lowest-order space-charge-induced corrections are derived up to the third-order. Substituting the on-axis derivatives into Gauss's law, the electric field is calculated up to the third order. Similar to the approach in the previous section, space-charge aberrations are computed from the convolution of the nonlinear field evaluated along the reference trajectory with the Green's function of the driven Hill equation:

$$\delta r(L) = \frac{eM}{\gamma^3 mc^2 \beta^2} \int_0^L S(z) \rho^{(2)}(z) \frac{r_c(z)^3}{8\epsilon_0} dz \tag{4.24}$$

Here, the relativistic factor γ^3 in the denominator accounts for the effects of relativistic mass increase and beam magnetic field forces. The electron's longitudinal velocity, expressed as $c\beta$, transforms the radial acceleration time derivatives into spatial derivatives. The linear space-charge defocusing field's effect is precisely incorporated by modifying

the linear transport trajectories $C(z)$ and $S(z)$, alongside adjustments in lens strength to maintain the imaging condition.

Substituting $r_c(z) = r_0 C(z) + r_0' S(z)$ provides the space-charge-induced image plane deviation terms, akin to Eq. 4.20:

$$\frac{\delta r_{sc}(L)}{M} = r_0^3 C_e^{(p)} + 3r_0^2 r_0' C_e^{(q)} + 3r_0 {r_0'}^2 C_e^{(r)} + {r_0'}^3 C_e^{(s)} \tag{4.25}$$

The aberration coefficients are given explicitly as:

$$C_e^{(p)} = \frac{e}{8\epsilon_0 \gamma^3 mc^2 \beta_0^2} \int_0^L \rho^{(2)}(z) C(z)^3 S(z) dz \tag{4.26}$$

$$C_e^{(q)} = \frac{e}{8\epsilon_0 \gamma^3 mc^2 \beta_0^2} \int_0^L \rho^{(2)}(z) C(z)^2 S(z)^2 dz \tag{4.27}$$

$$C_e^{(r)} = \frac{e}{8\epsilon_0 \gamma^3 mc^2 \beta_0^2} \int_0^L \rho^{(2)}(z) C(z) S(z)^3 dz \tag{4.28}$$

$$C_e^{(s)} = \frac{e}{8\epsilon_0 \gamma^3 mc^2 \beta_0^2} \int_0^L \rho^{(2)}(z) S(z)^4 dz \tag{4.29}$$

4.4.1 Space-charge aberrations in the uniform illumination case.

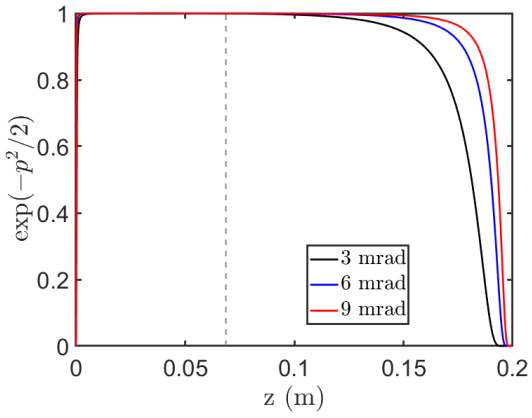
We will first examine the behavior of the aberration coefficients when the sample is uniformly illuminated by a beam with a Gaussian angular distribution. Assuming the beam is focused to a waist at the specimen location, the initial phase space distribution is represented as:

$$f(x, x', y, y') = \begin{cases} \frac{Q}{\pi R_0^2 L_b} \frac{1}{2\pi\sigma_\theta^2} \exp\left(-\frac{(x_0'^2 + y_0'^2)}{2\sigma_\theta^2}\right) & \text{if } r_0^2 < R_0^2 \\ 0 & \text{if } r_0^2 > R_0^2 \end{cases} \tag{4.30}$$

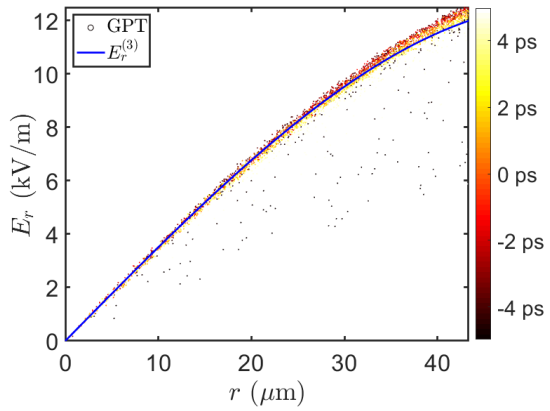
where Q is the total charge of the beam, R_0 is the initial edge radius, σ_θ is the root-mean-square (rms) beam divergence at the object plane, and L_b is the bunch length.

Using the results from the previous section, we invert the transport map to express the initial coordinates in terms of the final coordinates within the distribution function. We then integrate over the momentum space to ascertain the evolution of charge density along the optical transport. The charge density and its second derivative evaluated on-axis are given by:

$$\rho^{(0)}(z) = \frac{Q}{\pi \sigma_\theta^2 S^2 L_b} \frac{1 - \exp\left(-\frac{p^2}{2}\right)}{p^2} \tag{4.31}$$



(a)



(b)

Figure 4.5: (a) Plot of the non-linearity coefficient of the charge density distribution from the object plane to the image plane, with the red vertical dashed line indicating the sampled position along the column. (b) Comparison of the radial electric field experienced by each particle at its radial position in the beam against the analytical estimate used to derive the space-charge aberration coefficients, with particles color-coded based on their longitudinal position within the beam.

$$\rho^{(2)}(z) = -\frac{Q \exp\left(-\frac{p^2}{2}\right)}{2\pi\sigma_\theta^4 S^4 L_b} \tag{4.32}$$

The variable $p = \frac{R_0 C(z)}{\sigma_\theta S(z)}$ serves as an indicator, revealing whether the transverse spatial distribution transitions to Gaussian or remains uniform. A high p value suggests dominance of the initial spatial distribution, manifesting as a uniform beam profile. Conversely, a low p value indicates that the initial distribution is primarily angular, which leads to a more Gaussian spatial profile. Such a Gaussian profile typically exhibits a more pronounced non-linear characteristic in the space-charge field.

We begin by observing that as $p \to 0$ (i.e., when $C \to 0$), the expression $\frac{1-\exp\left(-\frac{p^2}{2}\right)}{p^2} \to$

$\frac{1}{2}$. Consequently, the zero-order density $\rho^{(0)}$ approaches $\frac{Q}{2\pi S^2\sigma_\theta^2 L_b}$, which corresponds to the on-axis density of a Gaussian profile with an rms spot size defined by $S\sigma_\theta$. Conversely, when $S \to 0$, such as at an object or image plane, $\rho^{(0)}$ converges to $\frac{Q}{\pi R_0^2 C^2 L_b}$, representing the density of a uniformly charged cylinder.

With the second derivative of the on-axis space-charge density known, we utilize the formulas from the previous section to calculate the aberration coefficients:

$$C_e^{(p)} = -\frac{K}{8\sigma_\theta R_0^3}\int_0^L p^3 \exp\left(-\frac{p^2}{2}\right) dz \tag{4.33}$$

$$C_e^{(q)} = -\frac{K}{8\sigma_\theta^2 R_0^2}\int_0^L p^2 \exp\left(-\frac{p^2}{2}\right) dz \tag{4.34}$$

$$C_e^{(r)} = -\frac{K}{8\sigma_\theta^3 R_0}\int_0^L p \exp\left(-\frac{p^2}{2}\right) dz \tag{4.35}$$

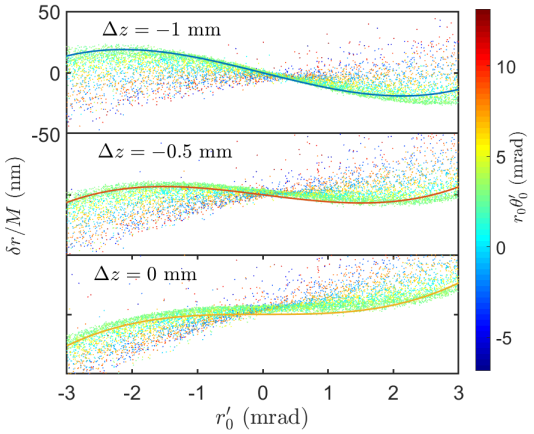
$$C_e^{(s)} = -\frac{K}{8\sigma_\theta^4}\int_0^L \exp\left(-\frac{p^2}{2}\right) dz \tag{4.36}$$

where the perveance factor $K = \frac{2I}{I_A\gamma^3\beta^3}$, with $I = \frac{Qc\beta}{L_b}$ representing the beam current, and $I_A \approx 17kA$ denoting the Alfven current.

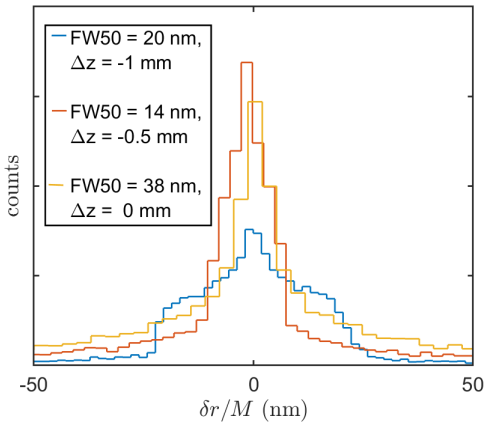
For minor deviations of trajectories from the axis, or when the field of view is confined to a small area around the axis, the primary contributor to image plane deviations will be the term associated with Eq. 4.36.

In scenarios where $\frac{R_0}{\sigma_\theta f} \ll 1$, the beam predominantly maintains a Gaussian profile, leading to $\int_0^L \exp\left(-\frac{p^2}{2}\right) dz \approx L$. For larger initial spot sizes or smaller focal waists, the Gaussian profile becomes localized primarily around the back focal plane of the lens.

According to the electron beam and lens parameters listed in Table 4.1, the validity of this approximation is illustrated in Fig. 4.5, where $\exp\left(-\frac{p^2}{2}\right)$ is plotted for three different beam divergences. In these instances, we have a beam current $I = 25$ mA, a perveance $K = 3.6 \times 10^{-9}$, and $C_e^{(s)} = 1$ m, which is significantly larger than contributions from any other aberrations. Thus, this effect is expected to be dominant at the image plane. Furthermore, linear space-charge forces alter $C(z)$ and $S(z)$, consequently amplifying the lens spherical aberration. However, as subsequent analysis will demonstrate, third-order space-charge effects overwhelmingly surpass third-order lens effects by nearly two orders of magnitude for a beam current of 25 mA.



(a)



(b)

Figure 4.6: (a) GPT output of image plane deviations and $\frac{\delta r}{M} = C_e^{(s)} r_0'^3 + \frac{\Delta z r_0'}{M^2}$, with M$\approx$ 8.5 and $C_e^{(s)} \approx$ 1m, plotted with respect to the initial angle for three different defocus conditions $\Delta z = -1$ mm, -0.5 mm and 0 mm. The latter corresponds to the linear transport image plane. (b) Corresponding histograms for each defocus position.

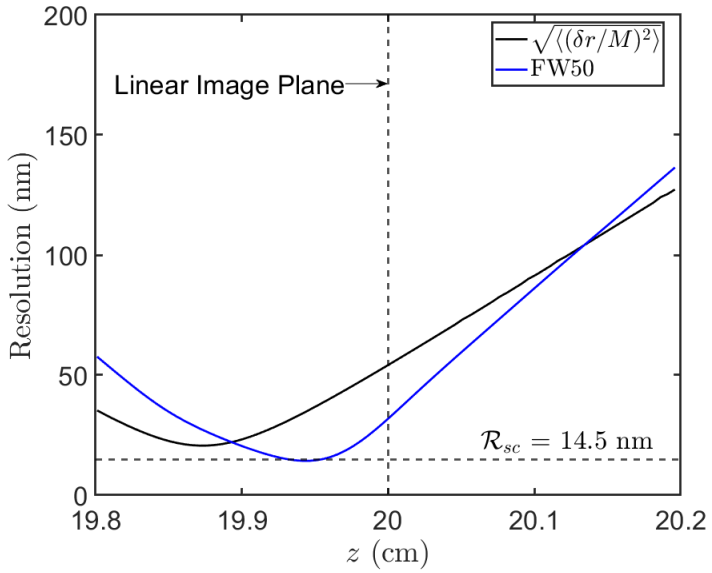


Figure 4.7: FW50 and standard deviation of image plane excursions plotted with respect to image plane placement. Linear transport is tuned so that the screen at 20 cm exactly satisfies the imaging condition ($\Delta z = 0$).

We continue to validate our model and estimates for $\delta r_{sc}/M$ using the General Particle Tracer (GPT). Specifically, we employ the "spacecharge3Dmesh" algorithm, which solves Poisson's equation on a mesh in the beam's rest frame [125]. This method utilizes a Lorentz transformation to the co-moving frame, capturing the fields interacting with particles at each timestep. While this approach offers computational efficiency—scaling linearly with the number of particles—it does not fully capture the granularity of the particle distribution, which could impact the precision of the simulation.

Further validation will be presented later in the paper, where we compare these smooth field simulation results with those obtained from "spacecharge3D." This alternative GPT algorithm models all Coulomb interactions among particles, providing a more detailed examination of inter-particle effects.

In Fig. 4.5b, we illustrate the transverse electric field experienced by particles at a specific location along the beamline (marked by the dashed line in Fig. 4.5a), and compare it with the third-order polynomial derived from our analytical model. The concordance between the two is excellent, with notable deviations occurring only at the beam's head and tail, where the assumption of an infinitely long beam does not hold.

The GPT simulation results, depicted in Fig. 4.6, show image plane deviations that largely conform to the expected cubic dependence. For instance, using the parameters listed in Table 4.1, particles impinging on the sample at 3 mrad are predicted to deviate by 30 nm from the ideal imaging condition. A comparison of the y-scale with that in Fig. 4.4 highlights the substantial impact of space-charge-induced aberrations on the instrument's spatial resolution.

In Fig. 4.6(a), color coding reveals that, despite space-charge effects, handkerchief distortions remain pronounced. These distortions are significantly magnified by space-charge forces, as evidenced by comparing this figure with Fig. 4.4(a). This amplification stems from our analysis being conducted in the Larmor frame, which inherently captures only radial distortions or meridional rays in the image plane. In the laboratory frame, the beam's rotation due to the solenoid lens field results in handkerchief distortions of a magnitude comparable to the radial deviations.

The spatial resolution of the instrument can be assessed by examining the width of the histogram of image deviations. It is crucial to acknowledge that by optimizing the linear defocus of the lens, we can significantly reduce the width of the projected distribution—by nearly a factor of two compared to its width at the image plane, as demonstrated in Fig. 4.7.

This approach aligns with the Scherzer defocus theorem [126]. To quantitatively understand the impact on image plane deviations, consider adjusting the output plane slightly by Δz. Taking into account the spherical aberration term, the output plane deviations are described by:

$$\delta r = \frac{\Delta z}{M} r_0' + M C_e^{(s)} r_0'^3 \tag{4.37}$$

Without defocus ($\Delta z = 0$) and assuming a Gaussian angular distribution, the rms spread of the image plane deviations is:

$$\sqrt{\langle (\delta r/M)^2 \rangle} = \sqrt{15} C_e^{(s)} \langle r_0'^2 \rangle^{\frac{3}{2}} \tag{4.38}$$

However, by introducing a small defocus $\Delta z = -3M^2 C_e^{(s)} \langle r_0'^2 \rangle$, we minimize the rms spread of the image deviations by a factor of $\sqrt{\frac{6}{15}} = 0.6325$:

$$\sqrt{\langle (\delta r/M)^2 \rangle}_{min} = \sqrt{6} C_e^{(s)} \langle r_0'^2 \rangle^{\frac{3}{2}} \tag{4.39}$$

While rms is commonly used to estimate beam width in accelerator physics, it tends to overweight outliers. Therefore, in TEM literature, the FW50 (full width containing 50% of the beam distribution) is often preferred. In our case, the minimum FW50 is 14 nm, significantly smaller than the rms width at the image plane as depicted in the defocus scan in Fig. 4.7.

Given these insights, we approximate the instrument's spatial resolution for the following section as:

$$\mathcal{R}_{sc}(I, \gamma, L, \sigma_\theta) \approx \frac{1}{2} C_e^{(s)} \sigma_\theta^3 \cong \frac{KL}{16\sigma_\theta}. \tag{4.40}$$

This approximation produces a value of 14.5 nm in our example, closely aligning with the GPT simulation results. If the lens is sufficiently thin, and the space-charge modified magnification approaches the zero-charge magnification, then $L \approx M f_0$, where f_0 is the lens focal length, suggesting that smaller focal length optics could reduce space-charge aberrations.

The perveance, scaling as γ^{-3}, and the lens focal length, scaling as γ^2 for a fixed magnetic field profile, dictate that the space-charge contribution to resolution will scale as $IM/\gamma\sigma_\theta$. This assumes that focal length reductions proportional to γ^2 are feasible, which may be limited by existing magnet technology capabilities.

4.5 Trade-offs between spatial and temporal resolution

This analytical estimate of space-charge aberration allows us to approximate the spatial resolution of a time-resolved single-shot transmission electron microscope, considering factors such as beam energy, beam current, spot size at the sample, and maximum opening angle. Given that the cumulative effect of spherical, chromatic, space-charge, and dose resolution limits surpasses the diffraction limit $R_d = 1.22\frac{\lambda}{\theta}$, where λ is the electron De Broglie wavelength, we do not include this term in our estimates. Notably, for beam energies exceeding 700 keV, where λ is less than 1 pm, even a semicollection angle of up to 5 mrad yields a diffraction limit smaller than a Bohr radius.

Assuming the contributions from different sources are independent, the overall resolution is estimated by summing the squares of all deviations from the ideal imaging

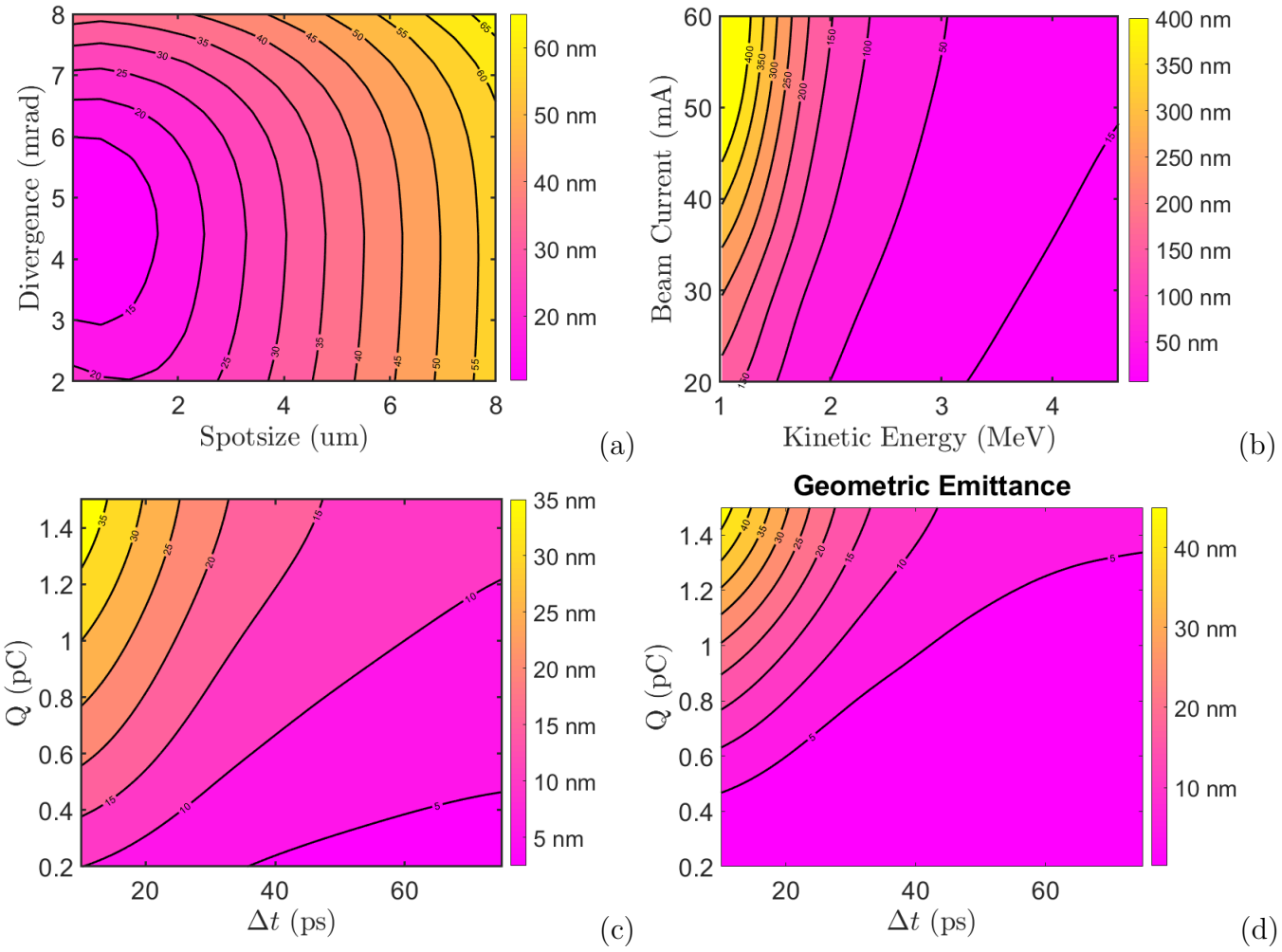


Figure 4.8: Analytical/numerical matrix propagator calculations of resolution. (a)Scanned dependence of resolution on illumination geometry for 4.3 MeV kinetic energy and 25 mA peak current. (b) Optimal resolution as a function of beam current and energy. The lens, image and object plane position are held constant in this plot. For each energy and beam current, the illumination parameters are chosen to optimize the resolution. (c) Optimized resolution for varying beam charge and bunch length for a 4.3 MeV kinetic energy beam. The illumination geometry is optimized at each point and the corresponding required geometric emittance is shown in (d).

condition and the resolution limit set by specimen illumination:

$$d = \sqrt{\left(C_c \frac{\delta\gamma}{\gamma}\sigma_\theta\right)^2 + (C_s\sigma_\theta^3)^2 + \mathcal{R}_{sc}^2 + \frac{SNR^2}{Dose}} \tag{4.41}$$

Here, we neglect the diffraction limit, $SNR = 5$ is the minimum signal-to-noise ratio required by the Rose criterion to distinguish features visually [127], and $Dose$ represents the time-integrated electron particle density at the object plane (e/area). This additional final term underscores that insufficient electron beam charge results in inadequate electrons per resolution pixel to resolve low contrast (20%) features statistically.

The formula for d in Eq. 4.41 serves as a preliminary approximation to gauge the resolution in a single-shot time-resolved TEM. It aids in understanding the trade-offs among various parameters but does not account for correlations between deviations from the imaging condition, which are not strictly independent. Detailed numerical simulations are still necessary to accurately determine the ultimate resolution limit. We treat Eq. 4.41 as a multivariate cost function to optimize within a reasonable range of beam and lens parameters for optimal imaging performance.

Optimizing the spot size and beam divergence at the sample plane using the condenser lens can enhance resolution. The scaling of space-charge aberration in Eq. 4.36 and its inverse dependency on beam divergence illustrate that increasing beam divergence reduces $\mathcal{R}_{sc}$ and thereby improves resolution until spherical aberrations predominate. Likewise, enlarging the illumination spot size boosts space-charge resolution but reduces the illumination dose, potentially degrading image quality due to the Rose criterion. An optimal balance can be struck based on given peak current and beam energy.

In practice, spot size and divergence at the sample are interdependent given a specific beam emittance from the electron gun. It's important to note that the typical contrast mechanism used in imaging intercepts scattered electrons with an aperture. Thus, excessive angular divergence at the sample diminishes contrast by mixing the scattered electron distribution with the transmitted distribution, obscuring object features. A lower beam emittance would enhance contrast by reducing intrinsic beam divergence.

In Fig. 4.8(a), we plot the estimated resolution for a 4.3 MeV energy and 25 mA peak current TEM as a function of spot size and divergence at the object plane. Divergence is varied between 2-8 mrad and spot size between 10 nm and 8 μm. The resulting contour

map, generated using Eq. 4.41, shows that nearly 10 nm resolution is achievable if the beam is focused to a 1 μm spot size with a divergence of 4 mrad. GPT simulations confirm this, showing a FW50 resolution of 12 nm for the optimized illumination geometry.

In Fig. 4.8(b), we illustrate the optimal resolution as a function of beam energy and current. This analysis is pertinent as various electron sources are being considered for ultrafast TEM operations [128, 129, 130, 131, 132]. Higher beam energies are generally preferable, although achieving a focal length as short as 1.5 cm becomes increasingly challenging at higher energies due to magnet technology constraints.

The optimization of illumination geometry is also feasible for different pulse lengths and charges, provided the beam aspect ratio is sufficient to satisfy the 2D limit approximation. In Fig. 4.8(c), the illumination geometry is optimized across various bunch lengths and charges, with the beam kinetic energy held constant at 4.3 MeV. Lower beam charges and longer pulse lengths generally improve resolution. However, each point in the plot reflects a re-optimization of beam divergence and illuminated area, demonstrating that higher beam charges allow a larger field of view but at a reduced spatial resolution. Sub-5 nm resolution is achievable with bunch lengths exceeding 100 ps and emittances below 5 nm.

4.6 Mitigation effects. Reshaping the Distribution.

In many time-resolved TEM applications, trading off temporal resolution for spatial resolution may not be desirable. In this section, we explore the potential resolution enhancements achievable by modifying the functional dependence between momentum space and real space. This can be implemented practically by utilizing the condenser lens to refocus the beam post an overfilled aperture, thereby creating a uniform illuminating momentum distribution while forming a Gaussian spatial distribution at the sample. By inverting the spatial and momentum space configurations in this manner, the second derivative of the space-charge density is altered to:

$$\rho^{(2)}(z) = -\frac{Q\exp(-p^{-2}/2)}{2\pi\sigma_r^4 C^4 L_b} \tag{4.42}$$

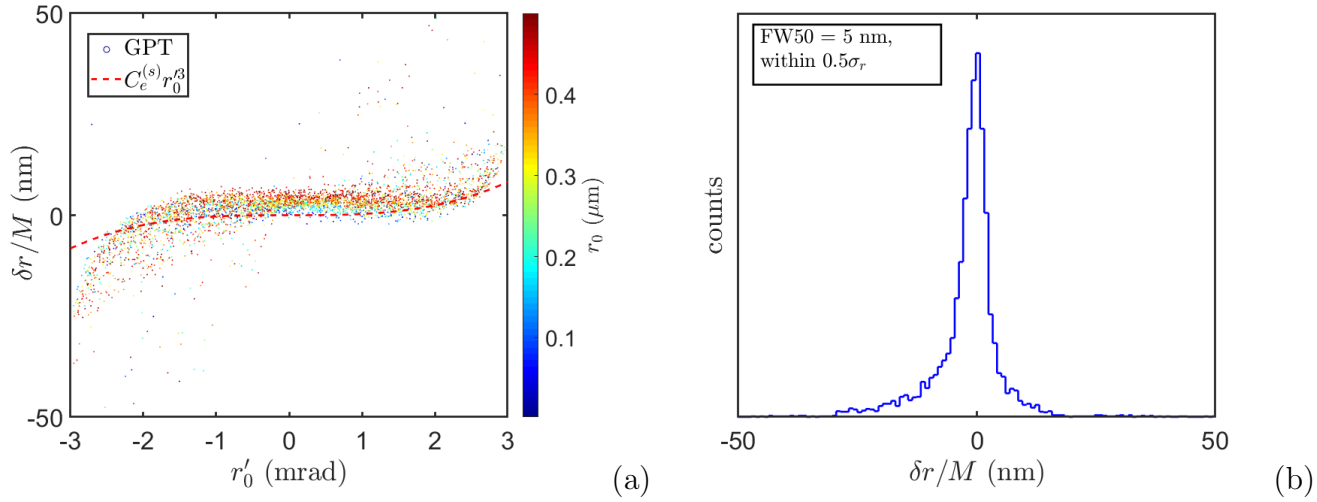


Figure 4.9: (a) Deviations at the image plane as a function of initial beam angles for a Gaussian spatially illuminated beam with a uniform angular distribution. (b) Histogram of the beam's projected distribution within the central 0.5 μm of the object space.

where $p = \frac{\sigma_r C(z)}{\theta_0 S(z)}$, with σ_r and θ_0 representing the Gaussian rms spot size and the hard-edge angular deflection at the object plane, respectively. Consequently, the aberration coefficients are recalculated as follows:

$$C_e^{(p)} = -\frac{K}{8\sigma_r^3\theta_0}\int_0^L \frac{\exp\left(-p^{-2}/2\right)}{p}dz \tag{4.43}$$

$$C_e^{(q)} = -\frac{K}{8\sigma_r^2\theta_0^2}\int_0^L \frac{\exp\left(-p^{-2}/2\right)}{p^2}dz \tag{4.44}$$

$$C_e^{(r)} = -\frac{K}{8\sigma_r\theta_0^3}\int_0^L \frac{\exp\left(-p^{-2}/2\right)}{p^3}dz \tag{4.45}$$

$$C_e^{(s)} = -\frac{K}{8\theta_0^4}\int_0^L \frac{\exp\left(-p^{-2}/2\right)}{p^4}dz \tag{4.46}$$

This specific object plane distribution yields a stronger non-linear dependence on r_0 and a suppressed dependence on r_0'.

Fig. 4.9 showcases GPT simulation results of the linear image plane deviations. The simulation parameters, detailed in Table 4.1, use a uniform angular distribution up to 3 mrad at the object plane instead of a Gaussian beam divergence. The perveance and peak dose remain consistent with those used in the beam simulated in Fig. 4.6. In Fig. 4.9(a), output image plane deviations are color-coded according to the initial position, r_0. Particles farther from the optic axis, moving toward the core, exhibit the largest

deviations. Sparse residual smearing, uncorrelated with r_0, instead correlates with $r_0\theta_0'$. Alongside the output deviations, the predicted spherical aberration curve is plotted. The notable discrepancy from the cubic trend in Fig. 8, compared to Fig. 5, stems from the pronounced influence of other space-charge aberration coefficients such as $C_e^{(p)}$, $C_e^{(q)}$, and $C_e^{(r)}$. The theoretical curve accurately predicts the behavior of particles as $r_0 \to 0$, but higher-order contributions in the Taylor series expansion of the charge density account for the residual discrepancies at larger r_0'. In Fig. 4.9(b), a histogram of the core outputs within a 0.5 μm offset from the optical axis shows an FW50 spatial resolution of 5 nm, nearly three times better than the corresponding uniform/Gaussian phase space distribution case. This indicates that significant improvements can be achieved by appropriately shaping the illumination in single-shot time-resolved TEMs.

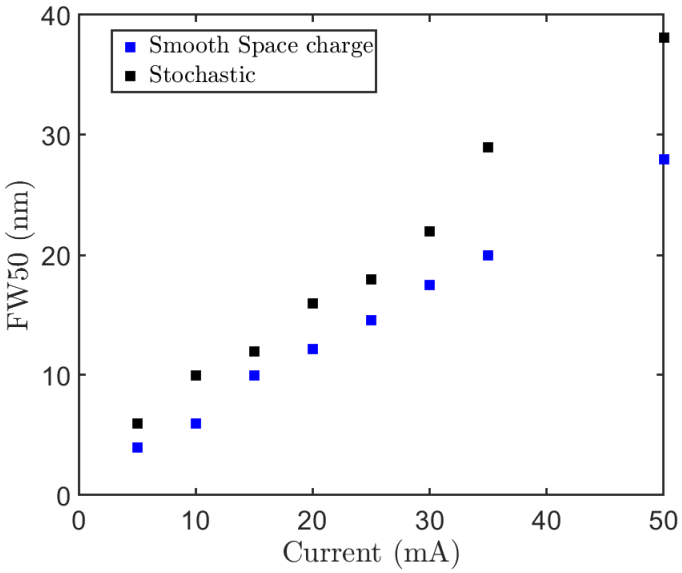


Figure 4.10: Comparison of FW50 measures between the ‘spacecharge3Dmesh‘ and ‘spacecharge3D‘ simulation models in General Particle Tracer (GPT).

4.7 Stochastic Scattering

To enhance the understanding of the analytical estimates we've discussed, we executed a final series of simulations using the GPT's "spacecharge3D" algorithm. This algorithm, which models binary interactions between particles, allows us to compare the deviations at the image plane observed in our smooth space-charge model with those in a more

general scenario. Due to the computational intensity, which scales quadratically with the number of particles ($\mathcal{O}(N^2)$), we limited our simulations to a small segment of the beam. Utilizing the continuous beam approximation, where fields remain relatively constant along the z-coordinate—especially true in the center of flat-top current distributions—we focused on just a 100 fs slice of the beam's central section. While this reduced sample size does not satisfy the Rose criterion, it still permits an assessment of spatial resolution by analyzing the FW50 of the image plane deviations. Despite the relativistic effects stretching the spatial dimensions in the beam's rest frame, we effectively approximated the space-charge field on average.

The findings are illustrated in Fig. 4.10, where we assess the FW50 across different beam currents using two distinct space-charge algorithms from GPT. We adjusted the number of particles for each simulation so that each macroparticle represents a single electron. Aside from charge, the parameters listed in Table 4.1 were employed to configure the beams for this study. Notably, the FW50 results are primarily influenced by the linear impacts of smooth space-charge effects. Including binary interactions only slightly increases the FW50 within the simulated current range. However, at lower beam currents, binary interactions become relatively more significant and ultimately determine spatial resolution limits. This observation underscores that mitigating smooth space-charge effects remains a considerable challenge in advancing single-shot time-resolved TEMs.

4.8 Multiple Magnification Stages

Previously, our discussion centered on aberrations associated with a single magnification stage. We now extend our analysis to consider the cumulative effect of space charge aberrations across multiple magnification stages. This analysis leverages a realistic optical setup incorporating MeV solenoids currently under development at SLAC by Electron Optica. Although technical details and schematics are not included here because they are proprietary, the results presented are derived from applying our formalism to the axial field data from actual lenses in the setup.

The Electron Optica optical column consists of an objective lens positioned at z=0mm, with the sample located at z=-23 mm. A weaker lens is positioned at z=500 mm, forms an

intermediate image, and transfers it to the first projector lens, providing a magnification of 2-3. and three projector lenses further downstream at z=1010 mm, 1290 mm, and 1570 mm, respectively, with the final screen at z=2000 mm. The system achieves a maximum magnification of about 2500x when all lenses run at full power.

The system under consideration is designed for electrons with a kinetic energy of 3.9 MeV. Consistent with earlier discussions, the initial electron distribution at the object plane assumes a Gaussian transverse profile, with a uniform angular spread. This distribution is characterized by an rms spot size of 0.5 μm, a maximum angular divergence of 3.5 mrad, and a total charge of 250 fC distributed over a duration of 10 ps.

The cumulative aberration (normalized by magnification) at any desired plane is given by:

$$\delta r(z) = -\frac{e}{\gamma^3 mc^2\beta^2}\int_0^z S(s)\rho^{(2)}(s)\frac{r_c(s)^3}{8\epsilon_0}\,ds \tag{4.47}$$

Here, δr represents a critical component of the trajectory correction in imaging transport systems relative to the object plane. Although this expression omits another term typically included, it vanishes near image planes because it is proportional to $S(z)$, thus not relevant to the assessment of aberrations. To evaluate the quality of an imaging transport system, δr is computed throughout the transport, together with the charge density evolution. This methodology facilitates pinpointing areas within the microscope where image quality is most adversely affected.

Figure 4.11(a) illustrates the rms spot size evolution through the system, overlaid with the axial field distribution indicated by a dashed black line. Additionally, a solid black line depicts an equivalent defocusing distribution resulting from the linear space charge force. The positions of each lens are marked by the vertical dashed lines. Figure 4.11(b) details the increasing resolution along the column primarily at the end of the second magnification stage. A solid black line represents the function $\exp(-p^{-2}/2)$ plotted to indicate whether the charge density distribution is Gaussian or transitions to uniform.

Significantly, the resolution deterioration is most pronounced near the first and second image planes, with 10 nm and 30 nm increases, respectively. This degradation could potentially be mitigated by reconfiguring the weaker second lens to match the specifications of the objective lens or projector lenses, decreasing both the focal length of the

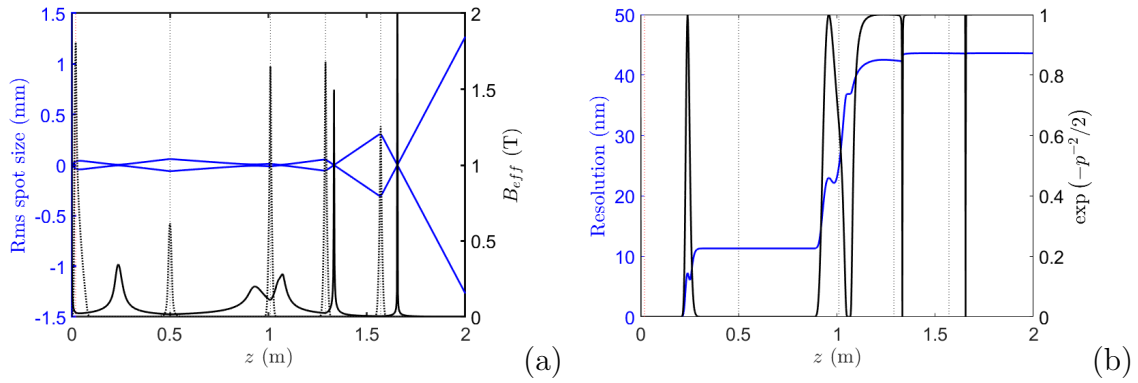


Figure 4.11: (a) Evolution of the rms spot size through the Electron Optica magnification stages, overlaid with the axial field distribution (dashed line) and the defocusing effect from linear space charge forces (solid black line). (b) Resolution along the column (blue), with the solid black line marking the transverse spatial distribution, showcasing the transition from Gaussian to potentially uniform charge density. These plots emphasize the difficulties in preserving resolution under the system's mechanical design constraints.

intermediate lens and the distance between objective and projection optics. However, the mechanical constraints of the current design prevent this modification. It is crucial to note that the column layout, not initially optimized for single-shot imaging, had already progressed to the fabrication phase when this numerical assessment was conducted.

4.9 Summary

In summary, by utilizing the Green's function technique to solve the driven Hill equation, we have evaluated the smooth space-charge-induced aberrations in single-shot time-resolved TEMs and derived analytical expressions that estimate spatial resolution based on relevant beam parameters.

Our analysis holds primarily in the perturbative regime, where space-charge nonlinear effects are minor compared to the linear forces acting on the beam. Despite this limitation, our findings have been validated through benchmarking against comprehensive particle tracking simulations. These formulas facilitate rapid assessments of the trade-offs between spatial and temporal resolution in single-shot time-resolved TEMs, as detailed in Section IV. Our results demonstrate that single nanometer spatial resolution

is achievable with relativistic beam energies, pulse durations in the tens of picoseconds, and precisely engineered spatial distributions.

To fully leverage the advantages of relativistic beam energies, there is a pressing need to innovate in electron optics, particularly through the development of elements with extremely short focal lengths. It's important to note that this analysis primarily addresses deviations from ideal imaging conditions. Additional factors must be considered when setting the optimal energy level for the instrument. At higher energies, the electron interaction cross-section decreases, which reduces contrast but permits the use of thicker samples due to greater penetration depths. In terms of diffraction contrast, it is crucial to fine-tune the beam divergence to minimize space-charge forces without exceeding the Bragg scattering angles. Moreover, although electron microscope detectors have been traditionally optimized for 100-300 keV, there are promising developments in direct electron detection for MeV energies [133].

Complex optical setups, including quadrupole and multipole lenses, could also be explored to further mitigate smooth space-charge aberration effects. Ultimately, the resolution limit of the instrument will be determined by stochastic blurring due to coulomb collisions between electrons in the beam post-sample.

This work aims to highlight the capabilities of single-shot time-resolved TEMs and to inform the development of future instruments for time-resolved microscopy applications.

CHAPTER 5

High energy UED instruments with tunable camera lengths

5.1 Introduction

Ultrafast Electron Diffraction (UED) provides a powerful means to observe the dynamic changes in materials at the atomic level, both spatially and temporally. This technique starts by exciting a sample with an ultrafast laser pulse and follows by probing its structure with a highly concentrated electron burst. This process captures detailed snapshots of the material's reciprocal space, reflected as diffraction patterns on a distant screen [134, 135, 136]. UED's capability to elucidate atomic-scale structural dynamics has been proven in various applications, from studying ultrafast phase transitions and warm-dense matter to examining gas and liquid phase phenomena [137, 138, 139, 140].

The initial applications of UED operated with non-relativistic electron energies [141, 142, 74]. However, it was soon realized that using relativistic electrons could substantially improve experimental outcomes, particularly in managing space charge effects, which either limit the electron count or the duration of electron bursts [75, 143, 144]. Modern UED systems have seen significant advancements, thanks to the integration of accelerator physics techniques like RF compression [145], cutting-edge cathode designs [146], phase space manipulations [147, 148], and enhanced detectors [149, 150].

A critical area of innovation in UED has been the development of magnetic optics. Prior to the sample interaction, robust lenses focus the high-brightness electron beam to sub-micron sizes, which is essential for analyzing minute or diverse samples [151, 78]. Following the sample interaction, preliminary tests have shown improvements in reciprocal space resolution through the use of magnetic lenses [152, 153], although many systems still rely on simple propagation drift to translate scattering angles into spatial

offsets at the detector.

Enhancing the energy range of UED instrumentation is also a focus. Traditionally, relativistic UED systems were capped at kinetic energies below 4 MeV due to constraints like available RF power and the accelerating gradient in RF guns. High energy beams face challenges such as reduced scattering angles requiring longer drifts for separating scattered electrons on detector screens [136]. Yet, increasing the beam energy has several advantages, such as deeper penetration through thicker samples, which could significantly broaden the experimental scope of UED, including applications in liquid cells and shock dynamics studies.

This chapter, through simulations and experimental demonstrations at the UCLA Pegasus beamline, the integration of post-sample optics to extend UED energies to 8 MeV [81, 136]. By employing a dual-lens telescope system, where a strong permanent magnet-based quadrupole (PMQ) triplet is used to enhance diffraction contrast and angular magnification, we demonstrate the potential for tunable diffraction camera lengths and improved reciprocal space resolution. This system opts for quadrupole triplets over round solenoidal lenses, as achieving short focal lengths at energies exceeding 5 MeV with solenoids would require extremely high magnetic fields and the use of bulky, often superconducting, coils [154].

Beginning with an examination of quadrupole dynamics, then we discuss specific optimization strategies of quadrupole transport for angular mangification, followed by detailed descriptions of experimental alignment techniques for the PMQ triplet, ensuring minimal beam steering effects while achieving optimal optical focusing. Experimental results with the high brightness beam from the Pegasus beamline, which showcase tunable magnification of a diffraction pattern from a crystalline gold sample at 8.2 MeV over a 1-meter distance, further substantiate our approach. Additionally, by studying diffraction pattern distortions as the optical system's alignment varies, we evaluate the imaging performance of the quadrupole triplets. Cross-referencing these observations with 3D magnetostatic simulations reveals the factors contributing to lens quality, offering insights into the underlying mechanisms affecting performance.

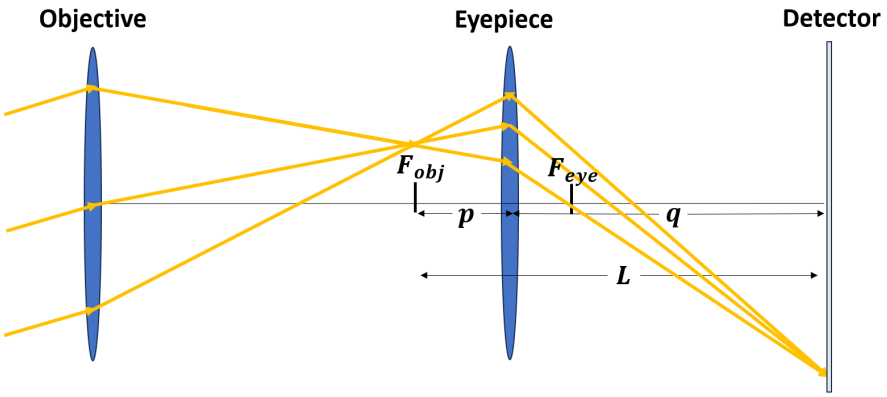


Figure 5.1: Ray diagram showing a angular magnification telescope. Initially, parallel rays converge to the same position at the detector with a magnified offset compared to the objective's back focal plane.

5.2 Thin lens transfer matrix description of an angular magnification transport

The magnification of the optical system after the UED sample can be easily understood using thin lens transport matrices. Consider two stages; the first is an objective lens, with focal length f_o, followed by a drift distance equal to its focal length, which transports to the back focal plane. The second is an eyepiece lens, with focal length f_e, which images the back focal plane of the objective lens to a downstream detector plane with magnification m. The matrix transport relating initial coordinates (x_0, x_0') to the final coordinates (x_f, x_f') is given by:

$$\begin{pmatrix} x_f \\ x_f' \end{pmatrix} = \begin{pmatrix} m & 0 \\ m' & \frac{1}{m} \end{pmatrix} \begin{pmatrix} 1 & f_o \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f_o} & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix} \tag{5.1}$$

where m' is the axial derivative of m at the image plane [155]. A ray diagram of this transport is shown in Fig. 5.1.

If the eyepiece lens is imaging over a total distance L, the object and image distances p and q can be expressed in terms of L and f_e as $\frac{L}{2}\left(1 \pm \sqrt{1 - 4f_e/L}\right)$ respectively. For a strong eyepiece lens (i.e. $f_e << L$), we can approximate the magnification $m \approx -L/f_e$.

In this case, the relationship between the final position on the detector screen to the

initial diffraction angle can be written:

$$x_f = m f_o x_0' \approx -L\,(f_o/f_e)\,x_0'. \tag{5.2}$$

clearly showing how a tunable diffraction camera can be obtained by adjusting the focal lengths for the eyepiece and objective in the angular magnification telescope.

This analysis, though focused on motion within a single transverse plane, highlights the broader objective. The challenge is to optimize a complex, quadrupole-based optical system to achieve consistent transport configurations across both planes. To ensure focusing in all directions, it is crucial to arrange quadrupole lenses into multiplets, such as triplets. This arrangement is necessary because while an individual quadrupole lens focuses in one plane, it simultaneously defocuses in the perpendicular plane.

5.3 Enhancing diffraction resolution by increasing angular magnification through the use of post-sample optics

To assess the impact of magnification on the resolution of reciprocal space in diffraction patterns, it is beneficial to introduce the concept of a diffraction contrast ratio, denoted as $R = \sigma/q$. Here, σ represents the root mean square (RMS) size of the beam at the detector, and q is the distance between a Bragg peak and the direct beam. The separation of the Bragg peak, q, is related to the Bragg angle θ_B and the angular magnification S via the equation $q = S\theta_B$. Since θ_B varies with the specific plane of reciprocal space being examined, normalizing the diffraction contrast ratio by the lattice periodicity provides the momentum transfer or q-resolution:

$$\sigma_Q = \frac{\sigma}{S\lambda} \tag{5.3}$$

where λ is the De Broglie wavelength of the electrons used to illuminate the sample.

Assuming that the transverse phase space of the beam at the sample is uncorrelated (this is typically the case with nearly parallel beam illumination), the width of a Bragg peak at the detector is influenced by the initial beam conditions at the sample, expressed as:

$$\sigma = \sqrt{C^2\sigma_0^2 + S^2\sigma_\theta^2 + PSF^2} \tag{5.4}$$

where σ_0 and σ_θ represent the initial RMS spot size and angular divergence of the beam at the sample, respectively, and PSF denotes the point spread function of the detector.

Inserting the above into the formula for relative q-resolution yields:

$$\sigma_Q = \frac{\sigma_\theta}{\lambda}\sqrt{1 + \left(\frac{C\sigma_0}{S\sigma_\theta}\right)^2 + \left(\frac{PSF}{S\sigma_\theta}\right)^2} \tag{5.5}$$

In the absence of post-sample optics, S equates to the drift length and $C = 1$. To enhance resolution, increasing the drift distance is advantageous as it reduces the contributions in the quadrature sum that multiply σ_θ/λ. Longer drifts are necessary for high-energy beams, where the intrinsic beam divergence scales as $\beta\gamma^{-1}$ to achieve near-ideal resolution.

The significant benefit of using post-sample optics is the ability to configure them with $C = 0$ at the detector plane. In this configuration, especially at higher beam energies, increasing the angular magnification S effectively nullifies the impact of the PSF.

As shown in Fig. 5.2, adding an objective triplet in the middle cartoon sharpens the diffraction pattern by effectively nullifying C. The inclusion of a PMQ eyepiece in the bottom cartoon magnifies the pattern expanding the separation of Bragg peaks potentially far beyond the PSF.

5.4 PMQ Triplet Specifications

In this implementation, we utilize two quadrupole triplets, serving as the objective and the eyepiece, respectively. For the eyepiece, it is crucial to employ the strongest lenses possible to maximize magnification over a given distance. Therefore, we use high-gradient permanent magnet quadrupole (PMQ) lenses, known for their short focal lengths of just a few centimeters at MeV electron energies [156].

A significant advantage of quadrupole-based lenses is their ability to manipulate rays in the horizontal and vertical planes independently. This feature allows for various imaging solutions with different magnifications, achievable through minor translations of the components within the eyepiece triplet. This adaptability is vital due to the fixed gradients of the PMQ lenses.

Each PMQ used in this experiment is a 16-sector Halbach-style array featuring an

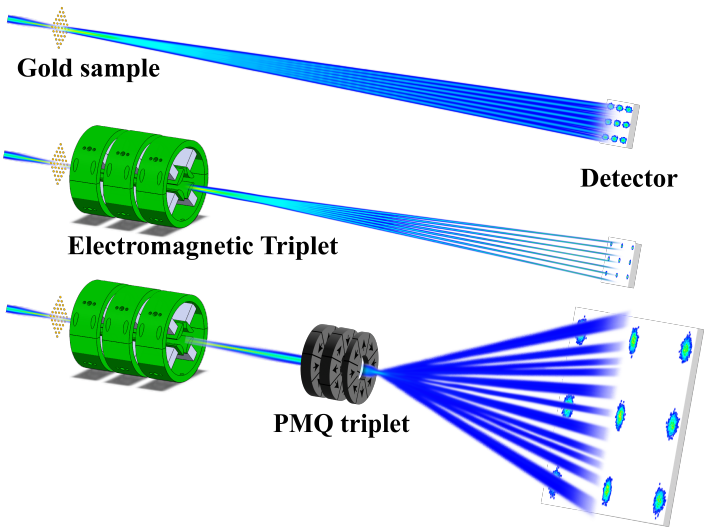


Figure 5.2: Illustration of the angular magnification telescope concept implemented using two quadrupole triplets at the Pegasus UED beamline. The cartoon is based on a start-to-end simulation.

inner radius of 1.75 mm and an outer radius of 3.5 mm, with each magnetized sector having a remnant magnetization of $M_0 = 1.21\,\mathrm{T}$ [45]. These magnets are constructed by assembling segmented permanent magnet pieces, where the magnetization vector rotates through 4π as it travels around the azimuth. Specifically, the magnetization segmentation is designed to approximate the continuous distribution $\mathbf{M} = M_0(\hat{\boldsymbol{\rho}}\sin(2\theta) + \hat{\boldsymbol{\theta}}\cos(2\theta))$, where θ is the azimuthal angle, and $\hat{\boldsymbol{\rho}}$ and $\hat{\boldsymbol{\theta}}$ are the orthogonal unit vectors in the radial and azimuthal directions, respectively, of the polar coordinate system.

The magnets are based on the design introduced in [157], which provides a convenient

Table 5.1: PMQ parameters measured by Hall probe

	Gradient (T/m)	**Effective Length (mm)**
First Quadrupole	510	6.16
Second Quadrupole	518	6.16
Third Quadrupole	417	3.9

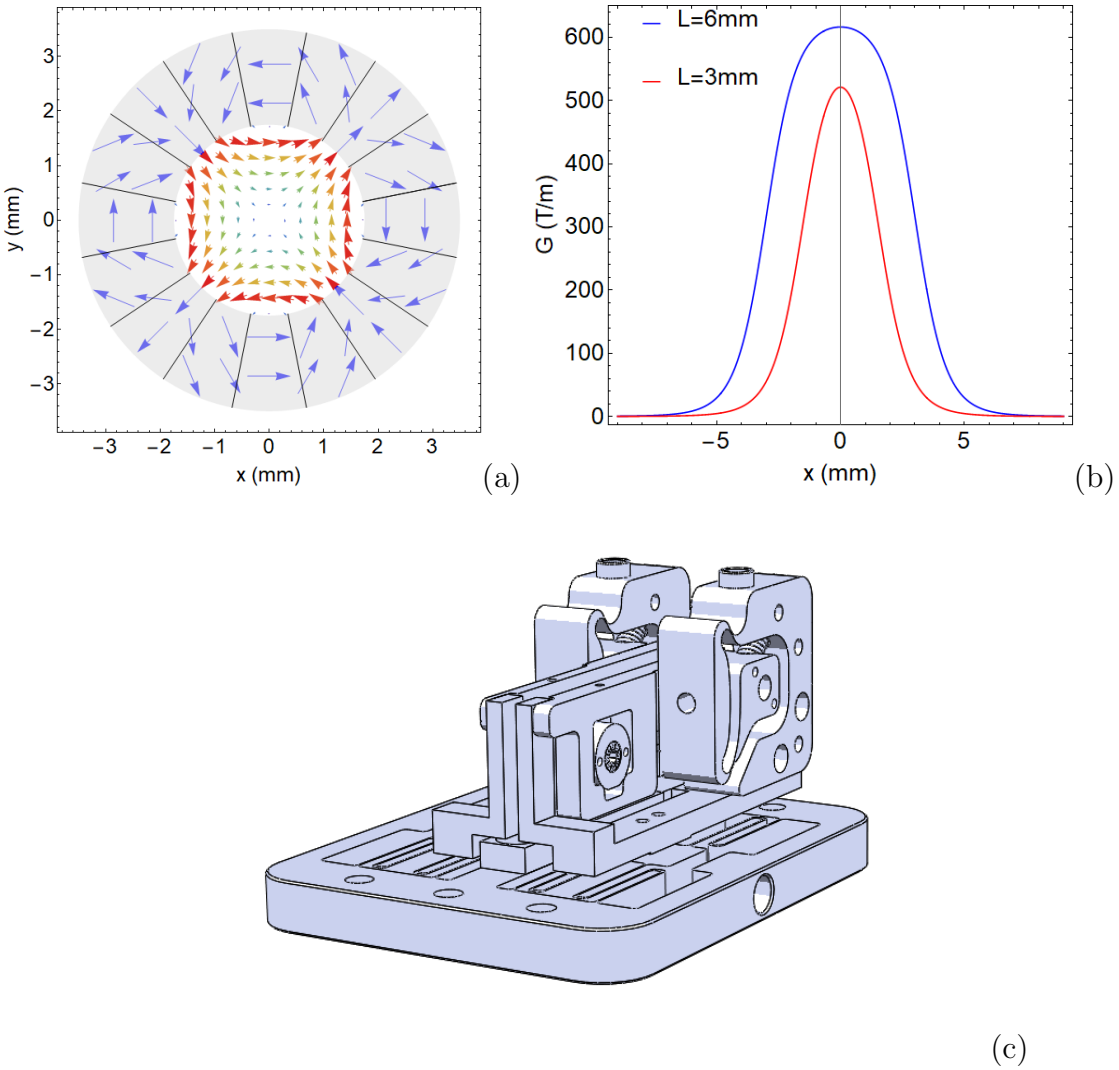


Figure 5.3: (a) Halbach style quadrupole with 16 slices; the magnetization is shown in blue and is the source for the quadrupole field within the aperture. (b) Expected gradient profiles from RADIA for the long and short PMQs. (c) Technical drawing of the PMQs as mounted on the flexure stage.

estimate for the gradient, assuming the Halbach quadrupole has a continuously changing magnetization (not segmented) and an infinite axial extent. The estimated peak gradient is $G_0 \sim 2M_0 \left(\frac{1}{r_{in}} - \frac{1}{r_{out}}\right) \approx 690\,\mathrm{T/m}$. However, since the lengths of the quadrupoles in the triplet are finite and comparable to the aperture size, the actual gradients are weaker.

Specifically, the triplet consists of two thick PMQs, each 6 mm in length, and one thinner PMQ, 3 mm in length. Figure 5.3(a) shows the magnetization profile of the 16-slice design and the resulting aperture field obtained from magnetostatic *RADIA* simulations [158]. The predicted gradient profiles for both PMQs are shown in Figure 5.3(b).

The actual magnets were formed using wire electrical discharge machined N35SH NbFeB. The PMQ triplet is positioned on a flexure-based mounting stage, which keeps the central quadrupole fixed in position relative to the stage but allows fine-tuning of both upstream and downstream PMQ spacings (refer to Figure 5.3(c)). This design theoretically allows us to achieve optimal imaging conditions by precisely adjusting the longitudinal distances between the quadrupoles.

5.5 Residual multipole components

The permanent magnet quadrupoles in the experiment exhibited significant aberrations due to a residual octupole component. To provide a comprehensive understanding, we will briefly delve into the theoretical origins of these aberrations.

At higher energy, quadrupoles are preferred because they achieve comparable focusing in a compact footprint. The individual quadrupole magnets are designed based on fundamental solutions to the Laplace equation, $\nabla^2\psi = 0$. These solutions conventionally assume that the elements extend infinitely along the axial (z) direction. In practical applications, however, these elements are of finite length. This is especially critical in quadrupoles where the theoretical expectation of producing linear fields is compromised. The finite dimensions of the quadrupole alter the expected linear relationship between the fields and the transverse coordinates, leading to the emergence of nonlinear forces and aberrations, which we seek to discuss in detail in this section.

In the design of charged particle magnetic lenses, it is common practice to approximate the magnetic field to the transverse plane, simplifying the complex 3D behavior to 2D.

Initially, solutions are sought for the 2D Poisson equation, which are then extended to be compatible with Maxwell's equations in a 3D context.

This can be achieved by converting the 2D Laplace equation to complex variables, letting $z = x+iy$ and $\bar{z} = x-iy$, where the Laplace equation becomes $\nabla_\perp^2 \psi = \left(\partial_x^2 + \partial_y^2\right) \psi = 4\partial_{\bar{z}}\partial_z \psi = 0$. This implies the scalar potential is the sum of a function of z and another of $\bar{z}$, i.e., $\psi = \mathcal{A}(z) + \mathcal{B}(\bar{z})$. For simplicity, if we consider a scalar potential that is only a function of z, then it can be expanded in a power series, $\psi = \sum_{n=1}^{\infty} iC_n z^n$, where either real or imaginary parts are solutions.

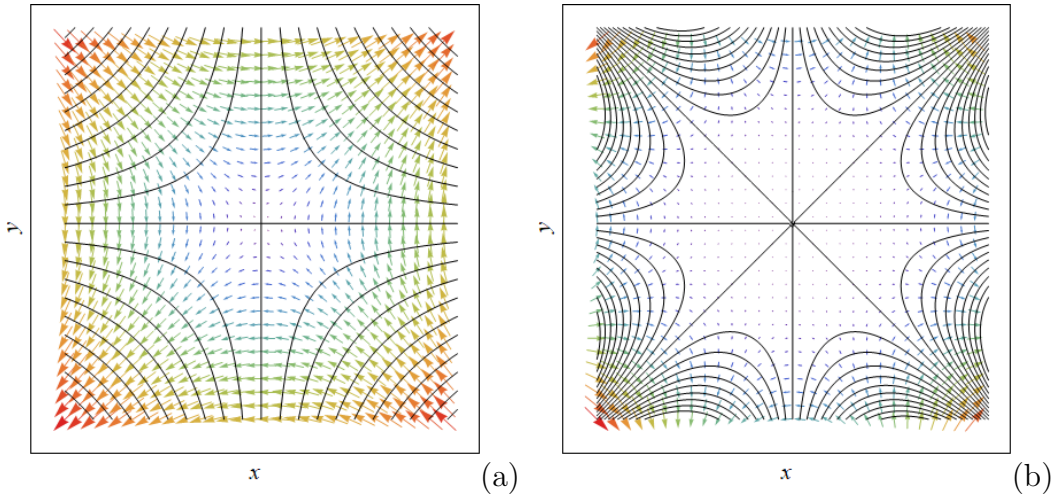


Figure 5.4: The magnetic scalar potential and vector field of a regularly oriented (a) quadrupole and (b) octupole.

A regularly oriented quadrupole field is derived from the second harmonic term in the expansion. We can identify the constant as being related to the field gradient $C_2 = G/2$ so that $\psi_2 = iC_2 z^2 = \frac{iG(x^2-y^2)}{2} - Gxy$. Taking the real part, we obtain a normal quadrupole field $\mathrm{Re}(\psi_2) = -Gxy$, which is depicted in Fig. 5.4(a) and yields desirable magnetic field components $\mathbf{B}_\perp = -\nabla_\perp \mathrm{Re}(\psi_2) = Gy\hat{\mathbf{x}} + Gx\hat{\mathbf{y}}$, and hence linear transverse forces for a charged particle moving in the $\hat{\mathbf{z}}$ direction.

Typically, when a quadrupole magnet is manufactured to meet design criteria, its properties are verified in the laboratory using a Hall probe. This probe measures the magnetic field along lines passing through the quadrupole parallel to its symmetry axis, with the probe oriented to detect either horizontal or vertical field components. If measurements reveal a linear trend across each plane, a gradient function $G(z)$ can be as-

signed.

Assuming that to leading order as $(x, y) \to (0, 0)$, the quadrupole potential approximates to $\psi_{Quad} \approx -G(z)xy$, we extend the 2D solution to 3D, ensuring Laplace's equation is satisfied, given the linear field gradient $G(z)$ from measurements. We assume a separable series solution of the form $\psi_{Quad} = -Gxy + \sum_{n=1}^{\infty} \mathcal{W}_n(x, y)\mathcal{V}_n(z)$. Applying the 3D Laplacian to ψ yields a suggestive equation:

$$\nabla^2\psi_{Quad} = -G''xy + \mathcal{V}_1\nabla_\perp^2\mathcal{W}_1 + \sum_{n=2}^{\infty}\left(V''_{n-1}W_{n-1} + \mathcal{V}_n\nabla_\perp^2\mathcal{W}_n\right) = 0 \tag{5.6}$$

So ψ is a solution, provided that we set $\mathcal{V}_1 = G''$, solve $\nabla_\perp^2\mathcal{W}_1 = xy$, then subsequently have for $n > 1$ that $\mathcal{V}_n = \mathcal{V}''_{n-1}$ and $\nabla_\perp^2\mathcal{W}_n = -\mathcal{W}_{n-1}$. So the series is determined by a sequence of successive solutions to the 2D Poisson equation. Again, by exploiting complex variables, the subsequent $\mathcal{W}_n$, satisfy $\partial_{\bar{z}}\partial_z W_n = -\frac{W_{n-1}}{4}$, and in particular $\partial_{\bar{z}}\partial_z W_1 = (z^2 - \bar{z}^2)/16i$. We obtain $\mathcal{W}_1 = (z^3\bar{z} - z\bar{z}^3)/48i = (x^3y + xy^3)/12$, then subsequently all later terms in the series can be generated to yield:

$$\psi_{Quad} = -Gxy + G''(x^3y + xy^3)/12 - G''''(x^5y + 2x^3y^3 + xy^5)/384 + \cdots \tag{5.7}$$

Design imperfections often lead to higher-order multipole moments within the focusing aperture. The short Halbach PMQs exhibit strong fringe fields and a significant fourth-order octupole moment. A regularly oriented octupole is derived from the real part of the potential, $\mathrm{Re}(\psi_4) = \mathrm{Re}(iC_4z^4) = 4C_4(xy^3 - x^3y)$, as shown in Fig. 5.4(b). Here, $H = 4C_4$ represents the octupole gradient function.

Extracting an octupole gradient function from a Hall probe measurement can be challenging, especially if the field resolution is poor. The field resolution is determined by the Hall probe's active area and its positioning's precision. However, it is straightforward to measure these moments in the post-processing of magnetostatic simulations using Fourier analysis at each axial plane. This procedure was actually utilized to validate the octupole moment post-experiment and is discussed in more detail at the end of this chapter.

The first term in the series expansion for the full solution approximates $\psi_{Oct} \approx H(xy^3 - x^3y)$ near the axis. Extending this procedure, the complete octupole solution is:

$$\psi_{Oct} = H(xy^3 - x^3y) - \frac{H''(x^5y - xy^5)}{5} + \frac{H''''(x^2 - y^2)(x^2 + y^2)^2}{240} + \cdots \tag{5.8}$$

When magnetic boundaries closely follow the potential contours of the designated 2D scenario, non-designated multipoles are suppressed. However, due to their finite axial extent, PMQs exhibit octupole moments in the aperture. Thus, it is useful to combine the 3D axial expansion solutions of the quadrupole and octupole moments for superposition, $\psi = \psi_{Quad} + \psi_{Oct}$.

5.6 Linear transport through quadrupole optics and expressions to calculate relevant aberrations

The optimization of the quadrupole strengths in the objective lens and the positioning of the eyepiece PMQs is informed by a matrix representation of the transport process from the sample to the detector. The focusing matrices account for quadrupole fringe fields, modeled by an Enge function [159] that describes the gradient for each quadrupole as follows:

$$G(z) = \frac{G_0}{2}\left[\tanh\left(\frac{b}{2}\left(\frac{l}{2} - z\right)\right) + \tanh\left(\frac{b}{2}\left(\frac{l}{2} + z\right)\right)\right], \tag{5.9}$$

where l is the effective length of the quadrupole, b represents the steepness parameter, approximately half the quadrupole gap radius, and G_0 is the nominal peak magnetic field gradient [160].

Within the paraxial approximation, the linearized transverse equations derived from the results in Chapter 2 yield the quadrupole focusing strength $\kappa(z) = \frac{G(z)}{[B\rho]}$, where $[B\rho] = \frac{p}{e}$ represents the axial momentum normalized by the electron charge [124]. Interestingly, this focusing strength scales more favorably compared to a round solenoid, which is inversely proportional to the square of $[B\rho]$.

The general solution to the system of ray equations in linearized quadrupole fields is represented as:

$$\begin{pmatrix} x_c \\ x_c' \\ y_c \\ y_c' \end{pmatrix} = \begin{pmatrix} C_x & S_x & 0 & 0 \\ C_x' & S_x' & 0 & 0 \\ 0 & 0 & C_y & S_y \\ 0 & 0 & C_y' & S_y' \end{pmatrix} \begin{pmatrix} x_0 \\ x_0' \\ y_0 \\ y_0' \end{pmatrix} \tag{5.10}$$

In this matrix, the functions C_x, C_y, S_x, and S_y denote the cosine and sine-like trajectories through the optical system. The principal planes of the quadrupole optics are identified

Table 5.2: Octupole Aberration Coefficients

U_{1111}	$\int_{z_0}^{z} \mathcal{G}_x h C_x^3 \, ds$	U_{3113}	$-3\int_{z_0}^{z} \mathcal{G}_y h C_x^2 C_y \, ds$
U_{1112}	$3\int_{z_0}^{z} \mathcal{G}_x h C_x^2 S_x \, ds$	U_{3123}	$-6\int_{z_0}^{z} \mathcal{G}_y h C_x S_x C_y \, ds$
U_{1122}	$3\int_{z_0}^{z} \mathcal{G}_x h C_x S_x^2 \, ds$	U_{3223}	$-3\int_{z_0}^{z} \mathcal{G}_y h S_x^2 C_y \, ds$
U_{1222}	$\int_{z_0}^{z} \mathcal{G}_x h S_x^3 \, ds$	U_{3333}	$\int_{z_0}^{z} \mathcal{G}_y h C_y^3 \, ds$
U_{1133}	$-3\int_{z_0}^{z} \mathcal{G}_x h C_x C_y^2 \, ds$	U_{3114}	$-3\int_{z_0}^{z} \mathcal{G}_y h C_x^2 S_y \, ds$
U_{1233}	$-3\int_{z_0}^{z} \mathcal{G}_x h S_x C_y^2 \, ds$	U_{3124}	$-6\int_{z_0}^{z} \mathcal{G}_y h C_x S_x S_y \, ds$
U_{1134}	$-6\int_{z_0}^{z} \mathcal{G}_x h C_x C_y S_y \, ds$	U_{3224}	$-3\int_{z_0}^{z} \mathcal{G}_y h S_x^2 S_y \, ds$
U_{1234}	$-6\int_{z_0}^{z} \mathcal{G}_x h S_x C_y S_y \, ds$	U_{3334}	$3\int_{z_0}^{z} \mathcal{G}_y h C_y^2 S_y \, ds$
U_{1144}	$-3\int_{z_0}^{z} \mathcal{G}_x h C_x S_y^2 \, ds$	U_{3344}	$3\int_{z_0}^{z} \mathcal{G}_y h C_y S_y^2 \, ds$
U_{1244}	$-3\int_{z_0}^{z} \mathcal{G}_x h S_x S_y^2 \, ds$	U_{3444}	$\int_{z_0}^{z} \mathcal{G}_y h S_y^3 \, ds$

where these functions equal zero; for instance, a focal plane corresponds to $C = 0$, and an imaging plane arises wherever $S = 0$. The optimal optical conditions for diffraction are achieved when $C = 0$ and S is maximized, enhancing reciprocal space resolution.

If we include the leading order octupole term in the magnetic scalar potential, then (neglecting axial derivatives) we can obtain the next leading order in the equations of motion from the potential:

$$\psi \approx -Gxy + H(x^3y - xy^3). \tag{5.11}$$

The equations of motion in the combined potential are given by:

$$x'' + \kappa x = h(x^3 - 3xy^2) \tag{5.12}$$

$$y'' - \kappa y = h(y^3 - 3x^2y). \tag{5.13}$$

$h = H/\left[B\rho\right]$ and H is the octupole gradient function.

Utilizing the Green's function method corrections to the linear transport are computed by integrating third-order driving terms in linear order, weighted by the green's function. The corrections corresponding to the octupole are as follows:

$$\delta x = \int_{z_0}^{z} (x_c^3 - 3x_c y_c^2) h \mathcal{G}_x ds \tag{5.14}$$

$$\delta y = \int_{z_0}^{z} (y_c^3 - 3x_c^2 y_c) h \mathcal{G}_y ds \tag{5.15}$$

In this context, z_0 denotes the objective's back focal plane, while z signifies the position of the detector or any point in between. By substituting and expanding the third-order polynomial, we obtain third-order terms that are proportional to combinations of product trios involving x_0, y_0, x_0', and y_0'. The coefficients of each term are also proportionally related to integrals involving trios of principal rays associated with their respective initial conditions. These coefficients are labeled as U_{ijkl} where i takes on either 1 or 3. The indices j, k, and l range from 1 through 4; for example, 1 and 2 points to C_x and S_x, while 3 and 4 points to C_y and S_y. The superscript corresponds to the plane on which the corrections are applied. Examples of all the relevant octupole aberration coefficients are provided in Table. 5.13

5.7 Diffraction imaging conditions: numerical optimization

The linear optical transport's matrix representation was employed to numerically fine-tune the quadrupole gradients/currents and PMQ spacings, aiming for equal angular magnifications at the detector plane. The parameters obtained from the first phase optimization of the linear transport were subsequently used in comprehensive GPT simulations, which incorporated actual 3D field maps from magnetostatic simulations of the quadrupoles. The simulation outcomes aligned to linear order with the predictions of the idealized fringe field matrix model.

Figures 5.5(a) and (b) illustrate the optical layout, positions, and physical lengths of the magnets, showcasing the optimized configurations for cosine-like and sine-like trajectories, respectively. These shapes determine the matrix-based transport. During optimization, the back focal planes were strategically positioned just ahead of the eyepiece, located 0.8 m downstream from the sample (defined as $z = 0$ in this context), while the detector screen was set at $z = 1.6\ m$. Constraints were imposed on the currents in the objective triplet to prevent overheating and magnetic saturation in the yoke. Similarly, adjustments to the spacing of the PMQs were limited to millimeter ranges due to their flexure-based mounting stage.

At the final screen, the system achieved an S value of -20 m, significantly surpassing the magnification possible without post-sample optics. At this magnification, a Bragg

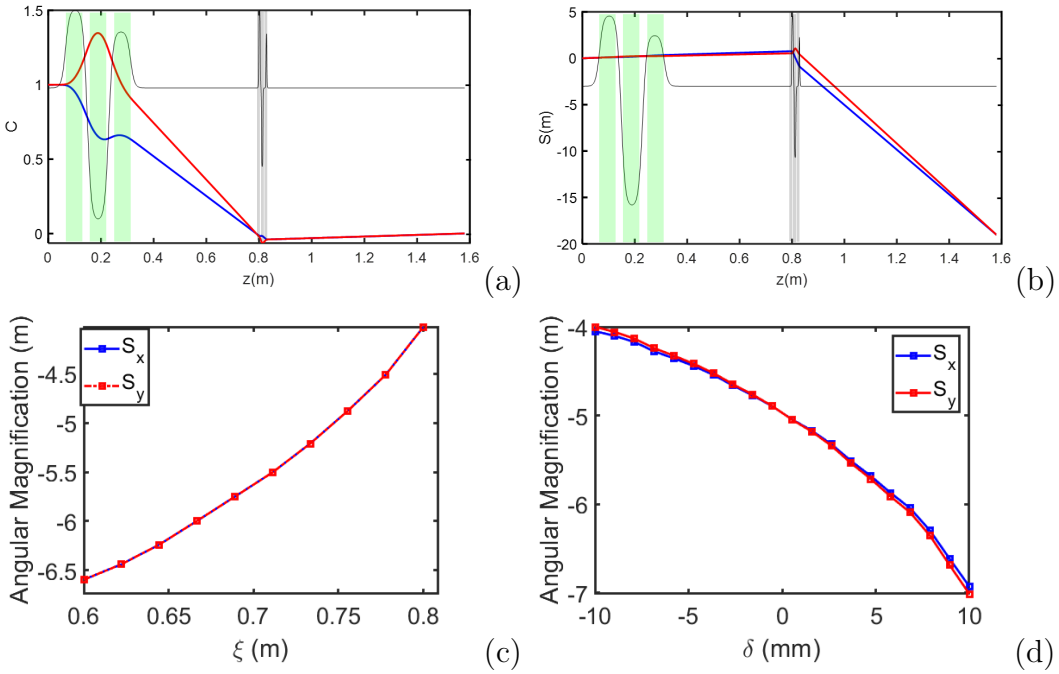


Figure 5.5: (a) The cosine-like ray traverses an optimized optics setup, with the green quadrupole triplet positioning back focal planes just in front of the PMQ triplet, which then forms images at the downstream detector at $z = 1.6\ m$. (b) The sine-like ray undergoes significant magnification at the detector. (c) An optimization scan for angular magnification was performed for a nearby screen, with symmetric back focal planes set at varied positions while maintaining the PMQ triplet's imaging capabilities at $z = 1.02\ m$. (d) The back focal planes are manipulated using the green quadrupole triplet, and the spacings of the PMQ triplet are fine-tuned to ensure symmetric imaging at $z = 1.02\ m$.

angle of 0.5 mrad results in a displacement of 1 cm, necessitating a substantially large detector screen to capture most diffraction peaks. Consequently, the experiment was conducted on a closer screen located at $z = 1.02\ m$, where lower magnification levels allowed for the accommodation of multiple Bragg peaks on the detector screen.

Two approaches were explored to adjust the magnification for a fixed distance from the sample to the detector. One method involved varying the quadrupole currents to align both back focal planes of the green triplet objective lens at the same longitudinal position along the PMQ triplet. The results of this method are illustrated in Figure 5.5(c), where ξ marks the position of the objective back focal planes. In each scenario, the eyepiece was optimized to focus onto the detector, allowing for adjustable angular magnification at the $z = 1.02\ m$ location. However, this method requires long range motion of the PMQ stage, which is impractical within the beamline's vacuum constraints.

Alternatively, by keeping the central position of the PMQ triplet fixed and varying the spacings between the PMQs, it was possible to image asymmetric objective back focal planes onto the detector with equal magnifications. The outcomes of this strategy are depicted in Figure 5.5(d), where δ indicates the separation between the two objective lens back focal planes around a fixed $\xi = 0.75\ m$. The eyepiece spacings were then adjusted to achieve balanced angular magnification.

5.8 PMQ triplet alignment

Attaining the desired imaging conditions required precise pre-alignment of the PMQ triplet on the stage. The mounts for the upstream and downstream PMQs were equipped with horizontal and vertical fine adjustment micrometers. To ensure the relative alignment of the triplet elements, we utilized the pulsed-wire method [161, 162].

As illustrated in Fig.5.6, a 50μm diameter beryllium-enhanced copper wire is threaded through the aperture of the PMQ triplet and tightened. Next, a pulse generator transmits a square-wave electric signal through the wire. The PMQs' magnetic fields cause the wire to experience a force, the magnitude of which depends on its distance from the PMQ centers. This force travels along the wire and is captured by a laser-photodiode system, which translates the movement into a voltage signal displayed on an oscilloscope.

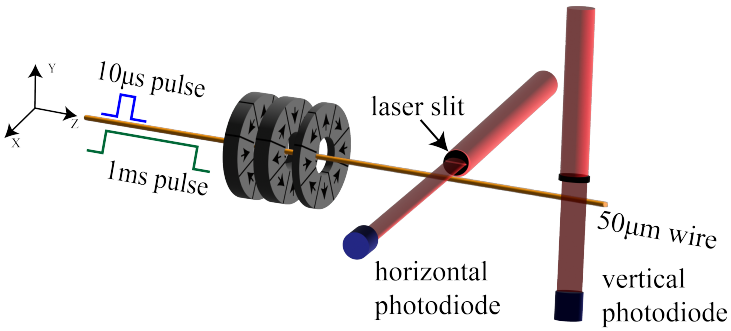


Figure 5.6: Setup of the pulsed-wire alignment technique. A beryllium-enhanced copper wire is threaded through the PMQ triplet aperture, tightened, and subjected to a square-wave electric signal. The wire displacement caused by the PMQ magnetic fields is detected by a laser-photodiode system.

First, a long 1 ms current pulse is applied, so the photodiode trace reflects the second integral of the field, essentially mapping the trajectory of the electrons. Adjusting the PMQ holders' positions results in a linear relationship between the peak signal height of the kick and the PMQ position, as the field increases linearly off-axis. Thus, the PMQs can be aligned by flattening the signal trace. The pulsed-wire method's accuracy is influenced by factors such as photodiode conversion gain, string tension, power supply jitter, and oscilloscope read-out error. As shown in Fig.5.7 (a), there is a linear correlation between signal height and actual PMQ offset. With an oscilloscope read-out jitter of 5 mV after averaging over 128 data points, we estimate the alignment accuracy to be within 25 μm. Post static alignment, we verify the alignment during axial translation of the upstream and downstream PMQs. Fig.5.7 (b) shows that the PMQs remain generally aligned, with a minor horizontal offset detected in the last PMQ. The maximum signal recorded is 34 mV, corresponding to a misalignment limit of 50 μm.

Then, shorter ($<$100μs) pulses are employed, which proxy the first integral of the field. This signal results from the effective velocity kick in the PMQ fields. Fig.5.7 (c) shows the output signals for various transverse offsets. These traces can be numerically analyzed to reconstruct the field profiles at different wire positions. Fig.5.7 (d) presents the fields derived from the photodiode traces using a 10μs pulse. As expected, the peak field changes sign and its magnitude increases with the wire's offset from the alignment axis. Further inspection reveals a linear relationship between the peak fields and the

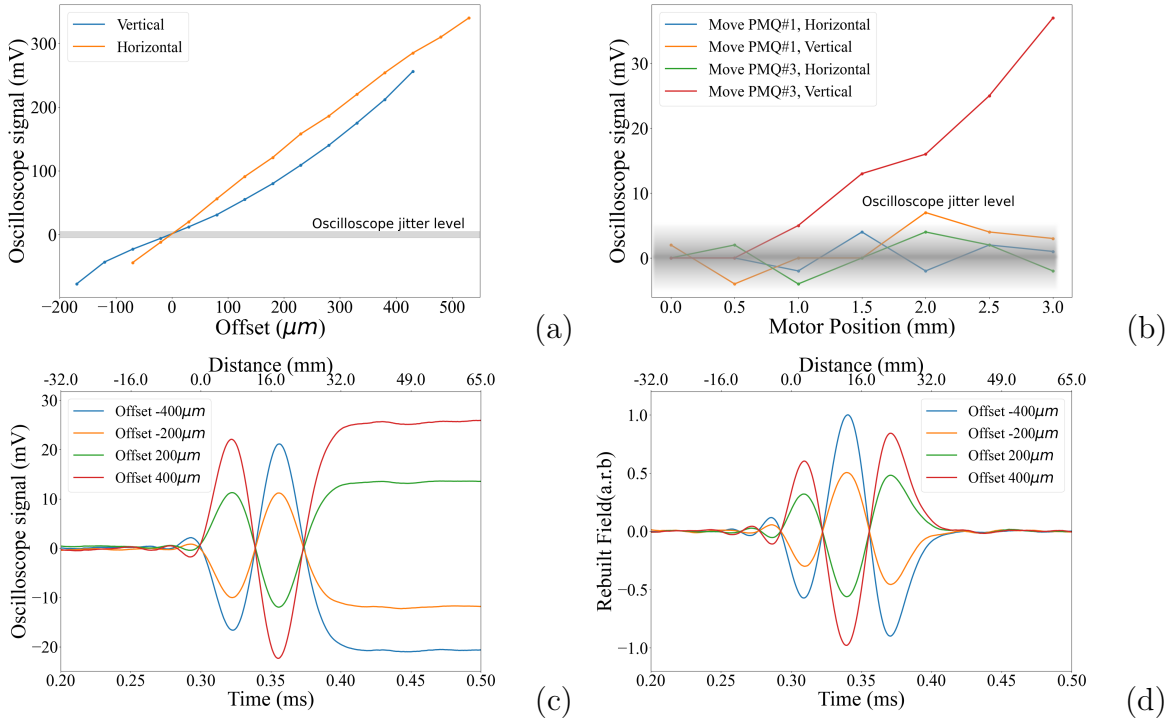


Figure 5.7: Pulsed-wire alignment results: (a) Calibration of the oscilloscope signal's central peak with PMQ displacement, showing a linear relationship. (b) Oscilloscope signal changes when the first and third PMQs are moved. (c) Signal traces for different transverse offsets, with timing on the bottom axis and distance on the top axis. (d) Retrieved PMQ field profiles using a 10μs pulse, showing peak field changes based on wire offset from the alignment axis.

transverse offset in the quadrupoles.

5.9 Results/observations from experimental run

The experimental activities were conducted at UCLA's Pegasus Laboratory, home to a high-brightness beamline powered by a high-gradient 1.6 cell S-band RF photo-injector gun [163]. A detailed, to-scale top-down view of the PEGASUS beamline configuration is depicted in Fig.5.8. The photo-injector, operating with a low charge of 0.5 pC per bunch as verified by an integrated current transformer (ICT) at the gun's exit, focused a laser spot normally on an alkali antimonide photocathode using a 0.75 m focal length lens to achieve a spot size of 150 μm. This setup allowed the gun to rapidly accelerate the electron

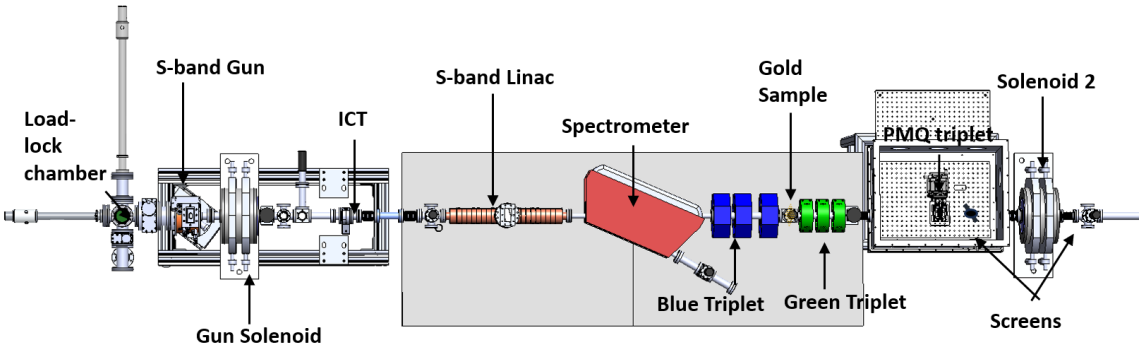


Figure 5.8: Portion of the Pegasus beamline relevant to the experiment. A technical drawing showing the main elements of the beamline used in the experiment.

beam to 3.2 MeV, effectively counteracting the degradation of transverse phase space due to space charge effects. Subsequent acceleration to 8.2 MeV was facilitated by a dual-slot resonantly coupled linac with high shunt impedance. A waveguide switch could be used to cut the RF power supplied to the linac. The beam's RMS angular spread at the sample plane was minimized through the coordinated use of a gun solenoid and a quadrupole triplet. A 30 nm-thick single crystal gold foil sample, mounted on a 3 mm TEM grid holder, was strategically positioned 3.1 m downstream from the cathode. Additionally, a HeNe laser, co-propagating with the electron beam, was utilized for precise alignment of the sample and the optical axes of the objective and eyepiece lenses. As illustrated in Fig. 5.5, the center of the green triplet is located 0.19 m from the sample, and the PMQ triplet, functioning as the eyepiece to enhance the magnification of the objective lens's back focal plane, is situated 0.81 m from the sample plane.

Comprehensive start-to-end simulations of the beamline were performed using the General Particle Tracer (GPT) software [164], which accurately models space charge effects and incorporates actual field maps of the beamline components, thereby accounting for higher-order aberrations. These simulations account for the beam's acceptance through the PMQ triplet aperture. Diffraction modeling at the sample involved the application of diffractive kicks to the simulated phase space distribution of the beam at the sample plane. As depicted in Fig. 5.9, these simulations provide a detailed view of the evolution of the RMS transverse beam sizes and kinetic energy, extending all the way to the final detector plane.

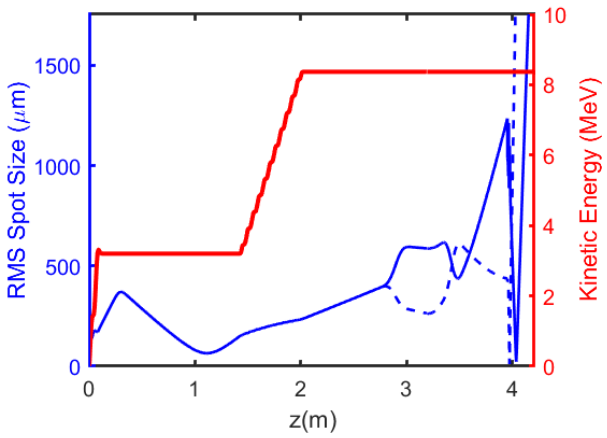


Figure 5.9: Comprehensive start-to-end simulation results of the Pegasus beamline conducted using General Particle Tracer (GPT) software, illustrating the impact of the RF gun, linac, and space charge effects on beam dynamics. Displayed are the root mean square (RMS) horizontal (solid blue line) and vertical (dashed line) beam envelopes alongside the beam's kinetic energy trajectory (shown in red) up to the detector.

During the experimental campaign, we adhered to the optimization strategy formulated during the simulation phase. Initial diffraction images were obtained with the linac deactivated, capturing the beam at a kinetic energy of 3.1 MeV within the vacuum box, as illustrated in Fig.5.10 (a). This served as a baseline reference. Subsequently, the linac was set to operate at the accelerating phase, which elevated the beam energy to 8.2 MeV, concurrently minimizing the energy spread. Fig. 5.10 (b) displays the diffraction pattern achieved under this configuration without post-sample optics.

Further adjustments were made as shown in Fig. 5.10 (c), where the green quadrupole triplet was optimized to align the back focal plane with the first DRZ screen, positioned 22 cm downstream of the PMQ triplet. Continuous monitoring of this screen facilitated further refinements to the green triplet, effectively aligning the back-focal planes directly in front of the eyepiece before the PMQ triplet was introduced into the beam path. Subsequent fine-tuning of the spacings between the PMQs symmetrically adjusted the final image, yielding a significantly magnified diffraction pattern displayed in Fig. 5.10 (d). The notable increase in magnification by the eyepiece is evident. However, it is observed that only six Bragg peaks were transmitted through the constrained PMQ aperture, aligning with the simulation predictions once the actual clearance of the PMQ was accounted

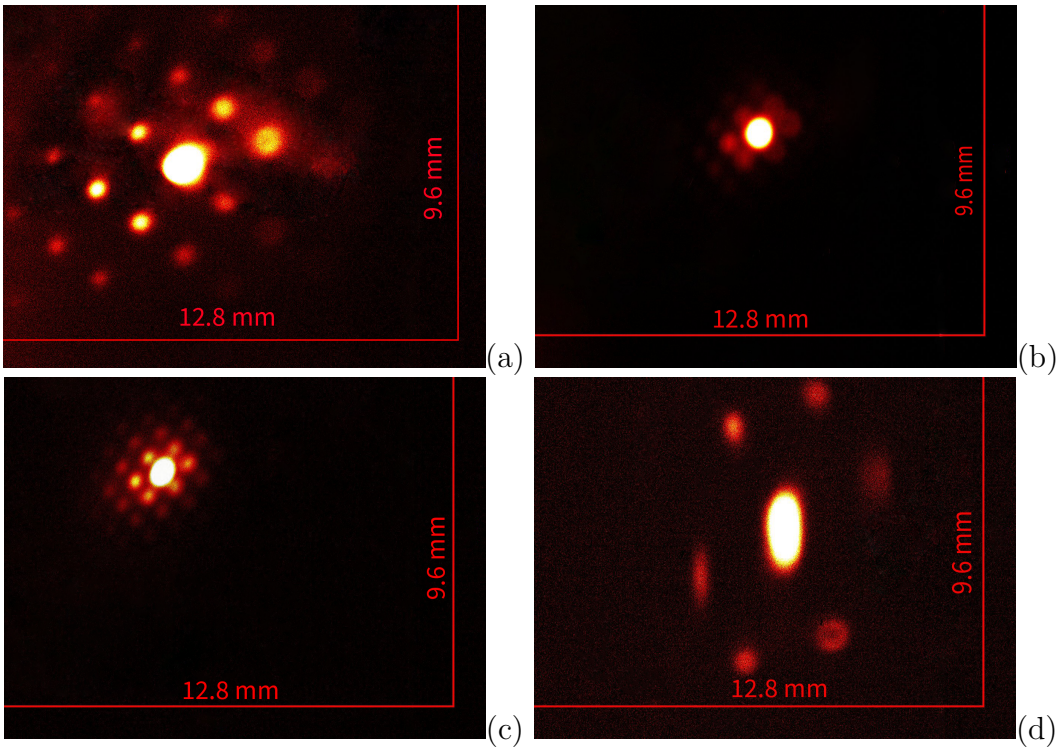


Figure 5.10: Single-shot diffraction patterns obtained at the $z = 1.02$ m fluorescent screen, showcasing varying focusing conditions and beam energies. (a) Captured with the linac deactivated, showing the beam at a kinetic energy of 3.1 MeV, with the green quadrupole triplet finely tuned for optimal focus onto the detector. (b-d) Recorded with the linac activated, elevating the beam energy to 8.2 MeV: (b) illustrates the diffraction pattern without any post sample focusing adjustments, (c) demonstrates the pattern with focusing, and (d) displays enhanced magnification.

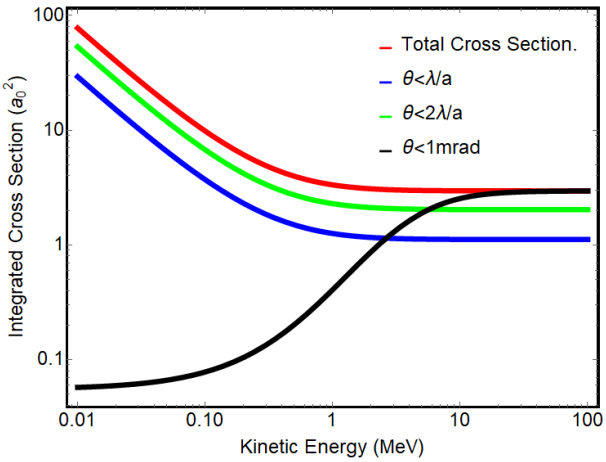


Figure 5.11: Comparative analysis of integrated cross sections for gold (Au) across an energy range from 0.01 to 100 MeV. The total cross-section is depicted by the red curve, while the blue and green curves represent the integrated cross sections up to the first and second Bragg orders, respectively. The black curve illustrates the integration of cross sections over forward-directed angles, specifically limited to 1 mrad, highlighting directional scattering intensities.

for.

In an ideal cylindrically symmetric system, one would expect to detect the eight lowest order Bragg peaks of the cubic gold lattice on the detector screen. However, the astigmatic transport induced by the quadrupole lenses and the specific rotation of the sample resulted in only two out of four 200 Bragg peaks clearing the narrow PMQ aperture, thus not reaching the detector screen. By employing a steering magnet at the entrance of the vacuum box, we manipulated various segments of the diffraction pattern through the PMQ triplet, effectively showcasing different areas of the reciprocal space.

In the analysis of high-energy ultrafast electron diffraction (UED) images, it is essential to first examine the energy dependence of the scattering cross-section, a key parameter for predicting the intensity of diffraction patterns at elevated energies. The differential cross-section for elastic scattering from an atom can be expressed as a function of the atomic number Z and the momentum transfer $s = 4\pi \sin(\theta/2)/\lambda$:

$$\frac{d\sigma}{d\Omega} = \frac{4Z^2}{s^4 a_0^2} \frac{1 - \beta^2 \sin^2(\theta/2)}{1 - \beta^2} \left[1 - F(s)^2\right]^2 \tag{5.16}$$

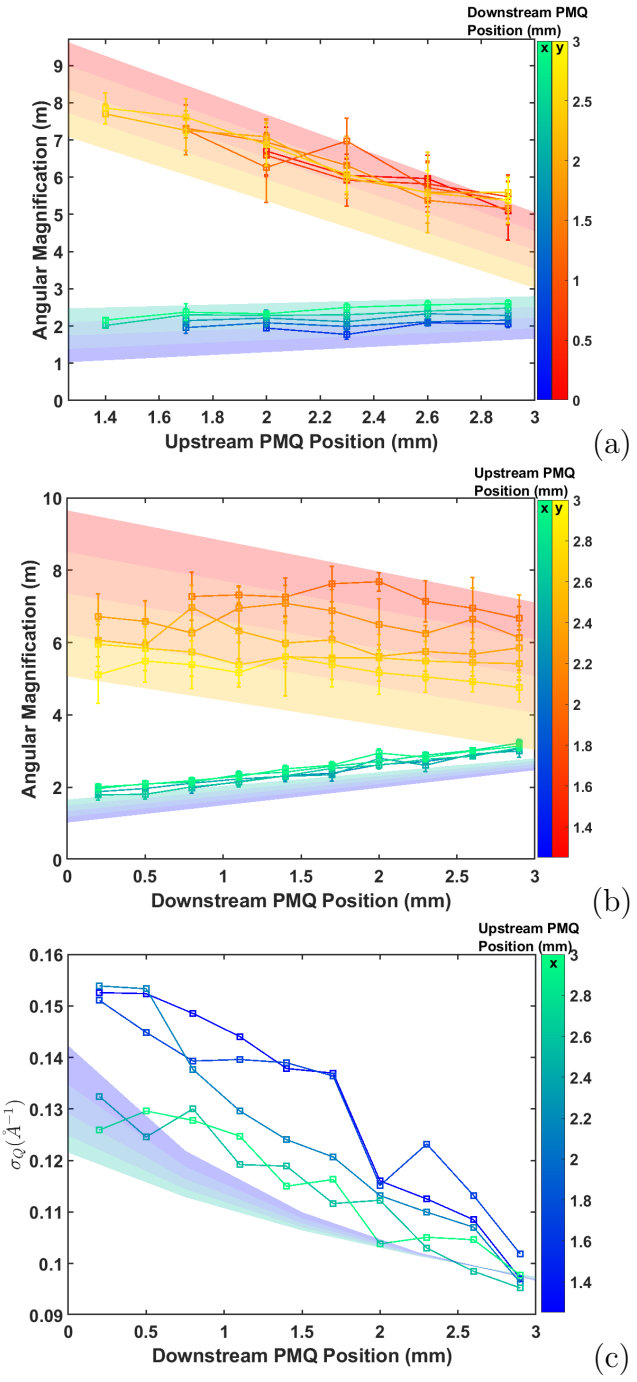


Figure 5.12: Results from spatial adjustments in the PMQ triplet, showing the impact of varying spacing. In (a) and (b), the spacing of the upstream and downstream PMQs is systematically reduced starting from 3mm, illustrating the adjustments' effects on beam dynamics. Panel (c) depicts the horizontal Q-resolution observed throughout these spacing changes.

where a_0 is the atomic Bohr radius, and $F(s) = \sum_{i=1}^{3} \frac{A_i \alpha_i^2}{s^2+\alpha_i^2}$ represents a function dependent on the details of the screened atomic potential. For gold (Z = 79), the constants are $\alpha_i = [22.864, 3.6914, 1.4886]$ Å^{-1} and $A_i = [0.2289, 0.6114, 0.1597]$ [136, 165, 166]. Considering the structure factor of the crystalline lattice, the cross-section is predominantly influenced by forward-directed rays at Bragg angles, as depicted in Fig. 5.11. This figure compares the total cross-section with those integrated up to integer multiples of λ/a, representing the intensity of Bragg peak orders. Additionally, integration up to a specific angle of 1 mrad illustrates how the solid angle, into which most particles are scattered, narrows as the energy increases.

Crucially, cross-sections normalized by the momentum factor $\gamma\beta$ across different annuli remain constant, suggesting minimal variation in the intensities of Bragg orders when the beam energy is increased from 3.1 to 8.2 MeV. This was substantiated by measuring the relative intensity of first-order Bragg peaks compared to the central beam at both energy levels, with pixel counts around these peaks being summed on the detector screen. The first-order peaks at high and low energy levels displayed intensities of 5.2% ± 1.0% and 5.6% ± 0.9%, respectively, aligning with experimental expectations and lying within the uncertainty range of each other.

Adjustments were made to the spacings of the PMQ triplet eyepiece on the flexure stage to explore the effects on magnification and q-resolution in the x and y directions. The axial position of the upstream PMQ was adjusted between 1.5 mm and 3 mm, while the downstream PMQ spacing was varied from 0 to 3 mm. During these adjustments, the centroid of the main beam consistently remained within a 500 μm radius on the detector. Considering the 1.8 cm focal length of the downstream PMQ and its 17.5 cm distance from the detector, the alignment offset was calculated to be approximately 50 μm, corroborating the pulse-wire alignment data expectations. The outcomes of these spacing adjustments are depicted in Fig. 5.12 (a) and (b), illustrating the results for the upstream and downstream PMQs respectively. The data points are color-coded based on the positions of the other quadrupoles, with simulation results superimposed for comparison. Consistent with both simulation and linear transport calculations, the positioning of the upstream PMQ predominantly influences the y-angular magnification, while the downstream PMQ impacts the horizontal magnification more significantly. The optimal

symmetric configuration, achieving an average magnification factor of 4, was established with both quadrupoles set at the 3 mm mark.

The q-resolution corresponding to each magnification setting was deduced by analyzing the ratio of the width of Bragg peaks to their separation from the main beam. This q-resolution, particularly in the horizontal direction along the x-axis, is plotted as a function of the downstream PMQ position in Fig. 5.12 (c). Consistent with the relationship described in (5.5), the data show that q-resolution improves in proportion to the increase in magnification.

The accuracy of the data is further validated through a simulation fit, which incorporated the measured spot size at the sample (530μm) and an estimated rms intrinsic beam angular divergence of 30 μrad. Additionally, a point spread function of 50 μm for the detector screen was assumed in the analysis. This model effectively captures the observed trends and alignments in the data, underscoring the predictability and stability of the optical system's performance.

It's essential to recognize that q-resolution is influenced by factors beyond the beamline optics, such as beam emittance, energy spread, and the screen's point spread function, none of which were optimized in this UED experiment. Therefore, the data mainly demonstrate the potential improvements in resolution enabled by post-sample magnification optics rather than achieving optimal absolute resolution. According to Eq. 5.5, the ultimate q-space resolution heavily depends on beam quality. For example, a tenfold reduction in the sample's spot size would require a corresponding tenfold decrease in source emittance to maintain the same q-resolution.

5.10 Measured PMQ aberrations

We observed noticeable distortions in the diffraction patterns as the beam navigated through the PMQ aperture, steered by an upstream magnet. We quantified the dominant lens aberration coefficients by altering the steering setpoints and capturing images at each adjustment. These were then compared to the influences expected from higher-order multipole moments, as predicted by the Radia model of the PMQs. Fourier analysis of the transverse fields in Radia revealed a residual octupole moment, a result of the finite

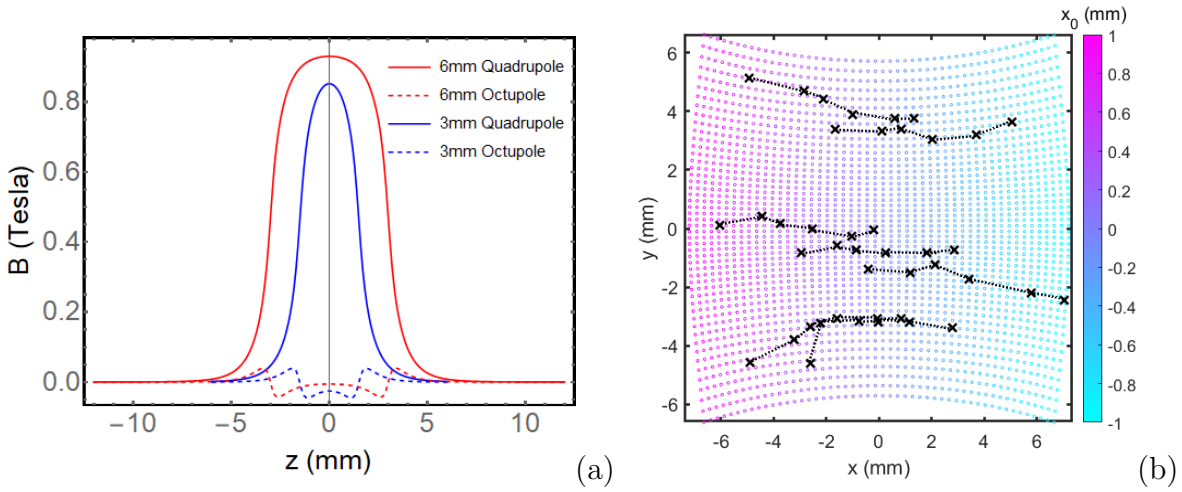


Figure 5.13: (a) Displays the quadrupole and octupole Fourier components and their axial variations for the 3mm and 6mm PMQs, highlighting the axial dependency of these higher-order magnetic field elements. (b) An overlay of horizontally steered diffraction beamlet centroid positions, as recorded on the detector, juxtaposed with a third-order aberrated representation of an initially square grid of rays. This overlay illustrates how the initial positions of the square grid reasonably coincide with the regions expected for the first-order Bragg peaks at the objective lens's back focal plane, effectively mapping the distortion effects on the diffraction pattern.

number of sharply magnetized wedges in the Halbach configuration. Given the standard orientation of the PMQs and their inherent magnetization symmetry, it was anticipated that the x-component of the magnetic field would expand in terms of sine functions, and the y-component in terms of cosine functions. This hypothesis was validated by the projections from the Radia field map:

$$b_{y,n}(r,z) = \frac{1}{\pi}\int_0^{2\pi} B_y(r,\theta,z)\cos(n\theta)\,d\theta \tag{5.17}$$

In particular, the Fourier decomposition at $r = 1.4$ mm across each axial plane, illustrated in Fig. 5.13(a), shows that while the quadrupole term dominates, a non-zero $b_{x,3} = b_{y,3} = b$ octupole moment is also present. These non-linear fields contribute to third-order ray equations and a third-order transport map:

$$x = C_x x_0 + S_x x_0' + U_{1111}x_0^3 + U_{1133}x_0 y_0^2 + \ldots \tag{5.18}$$

$$y = C_y y_0 + S_y y_0' + U_{3333}y_0^3 + U_{3311}x_0^2 y_0 + \ldots \tag{5.19}$$

where the coefficients $U_{1111} = 0.25\ \text{mm}^{-2}$, $U_{1133} = 2.20\ \text{mm}^{-2}$, $U_{3333} = -11.80\ \text{mm}^{-2}$, and $U_{3311} = -1.18\ \text{mm}^{-2}$ represent third-order transport terms, calculable via the Green's function method detailed in the appendix. Given that the beam is collimated at the objective's back focal plane, angle-dependent terms at the eyepiece location are minimized. The appendix Table 5.2 lists all ten unique third-order aberration coefficients for both transverse planes.

Figure 5.13(b) overlays the positions of the Bragg peaks at each steering setpoint with the theoretical transport predictions, including third-order components. This visualization demonstrates how a grid of axial rays distorts in a manner that aligns both qualitatively and quantitatively with the observed Bragg peak distortions. As these peaks are steered horizontally across the PMQ entrance, they delineate a path on the detector screen (marked by dashed lines in Fig. 5.13(b)) that illustrates the pincushion distortion characteristic of U_{3311}, validating our analysis of the PMQ field map.

While it is feasible to account for these aberrations in post-processing, various strategies can be employed to reduce their influence. One approach involves optimizing the PMQ design by enlarging the aperture to diminish the influence of higher-order moments in the magnetic field, which also helps avoid clipping of Bragg peaks and expands the reciprocal space field of view. Alternatively, adjusting the defocus of the imaging stage can mitigate the effects of third-order terms on resolution, a technique often employed to manage spherical aberration in Transmission Electron Microscopes (TEMs) [126, 167]. However, this approach alone cannot fully counteract the non-linearities in the transport; achieving comprehensive aberration correction would require the integration of higher-order optical elements.

5.11 Summary

In conclusion, we have investigated the performance of a compact optical system utilizing permanent magnet quadrupole (PMQ) optics for ultrafast electron diffraction, achieving high angular magnifications within a confined space. The precise alignment and motion control of the PMQs enhance the system's flexibility and tunability, presenting opportunities for innovative engineering solutions. The high focusing gradients introduce challenges,

such as limited beam acceptance due to the optics' small aperture and distortions from higher-order field components. Despite these issues, the system achieves magnification factors up to 8 and improved q-resolution, which are essential for observing dynamics at small scattering vectors.

Although the PMQs were not originally intended as projection lenses, their apertures were not optimized for this application. For a dedicated machine, a better selection of PMQs can be fabricated, demonstrating that optimal multiplets can be produced cost-effectively. However, designing the motion control requires additional care, a challenge that can be effectively addressed by a team of mechanical engineers. This approach enables economical focusing solutions, making high-energy diffractometers more widely accessible.

This system marks a departure from conventional diffraction techniques that typically do not employ post-sample optics. By integrating strong lenses, particularly at higher energy levels, our system addresses the challenge of achieving adequate separation at diminishing Bragg angles over short distances. This is vital for clear diffraction patterning at 8.2 MeV kinetic energy, nearly double that of existing relativistic UED beamlines.

The increased beam kinetic energy enhances diffraction resolution due to shorter electron wavelengths, reduced geometric emittance, and deeper sample penetration, enabling more detailed sample analysis. Additionally, higher energies reduce space charge effects and facilitate advanced phase space manipulations, paving the way for future UED instruments with shorter beams and smaller spot sizes.

References

[1] Maximilian Haider, Stephan Uhlemann, Eugen Schwan, Harald Rose, Bernd Kabius, and Knut Urban. Electron microscopy image enhanced. *Nature*, 392(6678):768–769, 1998.

[2] K. L. Merkle, L. J. Thompson, and Fritz Phillipp. Collective effects in grain boundary migration. *Phys. Rev. Lett.*, 88:225501, May 2002.

[3] Ultrafast transmission electron microscopy report 2012. Technical report, Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, 2012. Accessed: 2024-08-04.

[4] Ahmed H. Zewail. Four-dimensional electron microscopy. *Science*, 328(5975):187–193, 2010.

[5] Gaolong Cao, Shuaishuai Sun, Zhongwen Li, Huanfang Tian, Huaixin Yang, and Jianqi Li. Clocking the anisotropic lattice dynamics of multi-walled carbon nanotubes by four-dimensional ultrafast transmission electron microscopy. *Scientific Reports*, 5(1):8404, 2015.

[6] Omid Zandi, Allan E. Sykes, Ryan D. Cornelius, Francis M. Alcorn, Brandon S. Zerbe, Phillip M. Duxbury, Bryan W. Reed, and Renske M. van der Veen. Transient lensing from a photoemitted electron gas imaged by ultrafast electron microscopy. *Nature Communications*, 11(1):3001, 2020.

[7] Young Min Lee, Young Jae Kim, Ye-Jin Kim, and Oh-Hoon Kwon. Ultrafast electron microscopy integrated with a direct electron detection camera. *Structural Dynamics*, 4(4):044023, 05 2017.

[8] Ahmed H Zewail. Four-dimensional electron microscopy. *science*, 328(5975):187–193, 2010.

[9] Brett Barwick, David J Flannigan, and Ahmed H Zewail. Photon-induced near-field electron microscopy. *Nature*, 462(7275):902–906, 2009.

[10] Giovanni Maria Vanacore, Gabriele Berruto, Ivan Madan, Enrico Pomarico, P Biagioni, RJ Lamb, D McGrouther, O Reinhardt, I Kaminer, B Barwick, et al. Ultrafast generation and control of an electron vortex beam via chiral plasmonic near fields. *Nature materials*, 18(6):573–579, 2019.

[11] Thomas LaGrange, MR Armstrong, Ken Boyden, CG Brown, GH Campbell, JD Colvin, WJ DeHope, AM Frank, DJ Gibson, FV Hartemann, et al. Single-shot dynamic transmission electron microscopy. *Applied Physics Letters*, 89(4):044105, 2006.

[12] Thomas LaGrange, Geoffrey H. Campbell, B. W. Reed, Mitra Taheri, J. Bradley Pesavento, Judy S. Kim, and Nigel D. Browning. Nanosecond time-resolved investigations using the in situ of dynamic transmission electron microscope (dtem). *Ultramicroscopy*, 108(11):1441–1449, Oct 2008. Epub 2008 Jun 26.

[13] Chao Lu, Tao Jiang, Shengguang Liu, Rui Wang, Lingrong Zhao, Pengfei Zhu, Yaqi Liu, Jun Xu, Dapeng Yu, Weishi Wan, Yimei Zhu, Dao Xiang, and Jie Zhang. Imaging nanoscale spatial modulation of a relativistic electron beam with a mev ultrafast electron microscope. *Applied Physics Letters*, 112(11), 3 2018.

[14] Jinfeng Yang, Yoichi Yoshida, and Hidehiro Yasuda. Ultrafast electron microscopy with relativistic femtosecond electron pulses. *Microscopy*, 67(5):291–295, 07 2018.

[15] Ernest Hall, Susanne Stemmer, Haimei Zheng, Yimei Zhu, and George Maracas. Future of electron scattering and diffraction.

[16] Xijie Wang, Pietro Musumeci, Eliane Lessner, and Julia Goldstein. Report of the basic energy sciences workshop on the future of electron sources, september 8-9, 2016. 9 2016.

[17] D Filippetto and H Qian. Design of a high-flux instrument for ultrafast electron diffraction and microscopy. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 49(10):104003, apr 2016.

[18] S. P. Weathersby, G. Brown, M. Centurion, T. F. Chase, R. Coffee, J. Corbett, J. P. Eichner, J. C. Frisch, A. R. Fry, M. Gühr, N. Hartmann, C. Hast, R. Hettel, R. K. Jobe, E. N. Jongewaard, J. R. Lewandowski, R. K. Li, A. M. Lindenberg, I. Makasyuk, J. E. May, D. McCormick, M. N. Nguyen, A. H. Reid, X. Shen, K. Sokolowski-Tinten, T. Vecchione, S. L. Vetter, J. Wu, J. Yang, H. A. Dürr, and X. J. Wang. Mega-electron-volt ultrafast electron diffraction at SLAC National Accelerator Laboratory. *Review of Scientific Instruments*, 86(7):073702, 07 2015.

[19] Y. Murooka, N. Naruse, S. Sakakihara, M. Ishimaru, J. Yang, and K. Tanimura. Transmission-electron diffraction by MeV electron pulses. *Applied Physics Letters*, 98(25):251903, 06 2011.

[20] Feichao Fu, Shengguang Liu, Pengfei Zhu, Dao Xiang, Jie Zhang, and Jianming Cao. High quality single shot ultrafast MeV electron diffraction from a photocathode radio-frequency gun. *Review of Scientific Instruments*, 85(8):083701, 08 2014.

[21] P. Musumeci, J. T. Moody, C. M. Scoby, M. S. Gutierrez, H. A. Bender, and N. S. Wilcox. High quality single shot diffraction patterns using ultrashort megaelectron volt electron beams from a radio frequency photoinjector. *Review of Scientific Instruments*, 81(1):013306, 01 2010.

[22] Renkai Li, Chuanxiang Tang, Yingchao Du, Wenhui Huang, Qiang Du, Jiaru Shi, Lixin Yan, and Xijie Wang. Experimental demonstration of high quality MeV ultrafast electron diffraction. *Review of Scientific Instruments*, 80(8):083303, 08 2009.

[23] T. van Oudheusden, E. F. de Jong, S. B. van der Geer, W. P. E. M. Op 't Root, O. J. Luiten, and B. J. Siwick. Electron source concept for single-shot sub-100 fs electron diffraction in the 100 keV range. *Journal of Applied Physics*, 102(9):093501, 11 2007.

[24] Robert P. Chatelain, Vance R. Morrison, Chris Godbout, and Bradley J. Siwick. Ultrafast electron diffraction with radio-frequency compressed electron pulses. *Applied Physics Letters*, 101(8):081901, 08 2012.

[25] Meng Gao, Hubert Jean-Ruel, Ryan R. Cooney, Jonathan Stampe, Mark de Jong, Maher Harb, German Sciaini, Gustavo Moriena, and R. J. Dwayne Miller. Full characterization of rf compressed femtosecond electron pulses using ponderomotive scattering. *Opt. Express*, 20(11):12048–12058, May 2012.

[26] A Gliserin, A Apolonski, F Krausz, and P Baum. Compression of single-electron pulses with a microwave cavity. *New Journal of Physics*, 14(7):073055, jul 2012.

[27] M. R. Otto, L. P. René de Cotret, M. J. Stern, and B. J. Siwick. Solving the jitter problem in microwave compressed ultrafast electron diffraction instruments: Robust sub-50 fs cavity-laser phase stabilization. *Structural Dynamics*, 4(5):051101, 08 2017.

[28] T. van Oudheusden, P. L. E. M. Pasmans, S. B. van der Geer, M. J. de Loos, M. J. van der Wiel, and O. J. Luiten. Compression of subrelativistic space-charge-dominated electron bunches for single-shot femtosecond electron diffraction. *Phys. Rev. Lett.*, 105:264801, Dec 2010.

[29] W. Verhoeven, J.F.M. van Rens, E.R. Kieft, P.H.A. Mutsaers, and O.J. Luiten. High quality ultrafast transmission electron microscopy using resonant microwave cavities. *Ultramicroscopy*, 188:85–89, 2018.

[30] W Verhoeven, JFM van Rens, MAW Van Ninhuijs, WF Toonen, Erik R Kieft, PHA Mutsaers, and OJ Luiten. Time-of-flight electron energy loss spectroscopy using tm110 deflection cavities. *Structural Dynamics*, 3(5):054303, 2016.

[31] H.W. Kim, N.A. Vinokurov, I.H. Baek, et al. Towards jitter-free ultrafast electron diffraction technology. *Nature Photonics*, 14:245–249, 2020.

[32] Fengfeng Qi, Zhuoran Ma, Lingrong Zhao, Yun Cheng, Wenxiang Jiang, Chao Lu, Tao Jiang, Dong Qian, Zhe Wang, Wentao Zhang, Pengfei Zhu, Xiao Zou, Weishi Wan, Dao Xiang, and Jie Zhang. Breaking 50 femtosecond resolution barrier in mev ultrafast electron diffraction with a double bend achromat compressor. *Phys. Rev. Lett.*, 124:134803, Mar 2020.

[33] D. B. Cesar. *Probing Ultrafast Dynamics With Relativistic Electrons*. PhD thesis, UCLA, 2019.

[34] D.H. Dowell, I. Bazarov, B. Dunham, K. Harkay, C. Hernandez-Garcia, R. Legg, H. Padmore, T. Rao, J. Smedley, and W. Wan. Cathode rd for future light sources. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 622(3):685–697, 2010.

[35] William E. Spicer and Alberto Herrera-Gomez. Modern theory and applications of photocathodes. In Kenneth J. Kaufmann, editor, *Photodetectors and Power Meters*, volume 2022, pages 18 – 35. International Society for Optics and Photonics, SPIE, 1993.

[36] Triveni Rao and D. Dowell. An engineering guide to photoinjectors. 03 2014.

[37] David Alesini, Antonio Battisti, Massimo Ferrario, Luca Foggetta, Valerio Lollo, Luca Ficcadenti, Valerio Pettinacci, Sean Custodio, Eylene Pirez, Pietro Musumeci, et al. New technology based on clamping for high gradient radio frequency photogun. *Physical Review Special Topics-Accelerators and Beams*, 18(9):092001, 2015.

[38] D. Filippetto, P. Musumeci, R. K. Li, B. J. Siwick, M. R. Otto, M. Centurion, and J. P. F. Nunes. Ultrafast electron diffraction: Visualizing dynamic states of matter. *Rev. Mod. Phys.*, 94:045004, Dec 2022.

[39] D. Filippetto, P. Musumeci, M. Zolotorev, and G. Stupakov. Maximum current density and beam brightness achievable by laser-driven electron sources. *Phys. Rev. ST Accel. Beams*, 17:024201, Feb 2014.

[40] D. H. Dowell, S. Joly, A. Loulergue, J. P. de Brion, and G. Haouat. Observation of space-charge driven beam instabilities in a radio frequency photoinjector. *Physics of Plasmas*, 4(9):3369–3379, 09 1997.

[41] O. J. Luiten, S. B. van der Geer, M. J. de Loos, F. B. Kiewiet, and M. J. van der Wiel. How to realize uniform three-dimensional ellipsoidal electron bunches. *Phys. Rev. Lett.*, 93:094802, Aug 2004.

[42] P. Musumeci, J. Giner Navarro, J.B. Rosenzweig, L. Cultrera, I. Bazarov, J. Maxson, S. Karkare, and H. Padmore. Advances in bright electron sources. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 907:209–220, 2018. Advances in Instrumentation and Experimental Methods (Special Issue in Honour of Kai Siegbahn).

[43] RK Li and P Musumeci. Single-shot mev transmission electron microscopy with picosecond temporal resolution. *Physical Review Applied*, 2(2):024003, 2014.

[44] Frederick Cropp. *High-Performance Accelerator Modeling: Toward Improving Controls and Diagnostics for High-Brightness Beams in Experiment.* Phd dissertation, University of California, Los Angeles, Los Angeles, CA, 2023. ProQuest ID: Cropp_ucla_0031D_22210. Merritt ID: ark:/13030/m5hf8pnb.

[45] D Cesar, J Maxson, P Musumeci, Y Sun, J Harrison, P Frigola, FH O'Shea, H To, D Alesini, and RK Li. Demonstration of single-shot picosecond time-resolved mev electron imaging using a compact permanent magnet quadrupole based lens. *Physical review letters*, 117(2):024801, 2016.

[46] Martin Reiser and Patrick O'Shea. *Theory and design of charged particle beams*, volume 312. Wiley Online Library, 1994.

[47] W.A. Strauss. *Partial Differential Equations: An Introduction.* Wiley, 2007.

[48] Helmut Wiedemann. *Particle accelerator physics.* Springer Nature, 2015.

[49] L.D. Landau, E.M. Lifshitz, and J.B. Sykes. *Mechanics: Volume 1.* Course of theoretical physics. Elsevier Science, 1976.

[50] Harald Rose. *Geometrical Charged-Particle Optics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[51] G. W. Hill. On the part of the motion of the lunar perigee which is a function of the mean motions of the sun and moon. *Acta Mathematica*, 8(1):1–36, December 1886.

[52] E.D Courant and H.S Snyder. Theory of the alternating-gradient synchrotron. *Annals of Physics*, 3(1):1–48, 1958.

[53] J. Rosenzweig and Oxford University Press. *Fundamentals of Beam Physics*. Oxford University Press, 2003.

[54] Robert Joel England. *Longitudinal Shaping of Relativistic Bunches of Electrons Generated by an RF Photoinjector*. PhD thesis, University of California, Los Angeles, 2007. PhD Dissertation.

[55] David B Williams and C Barry Carter. The transmission electron microscope. In *Transmission electron microscopy*, pages 3–17. Springer, 1996.

[56] V. Danilov, S. Cousineau, S. Henderson, and J. Holmes. Self-consistent time dependent two dimensional and three dimensional space charge distributions with linear force. *Physical Review Special Topics - Accelerators and Beams*, 6:094202, Sep 2003.

[57] I. M. Kapchinskij and V. V. Vladimirskij. Limitations Of Proton Beam Current In A Strong Focusing Linear Accelerator Associated With The Beam Space Charge. In *2nd International Conference on High-Energy Accelerators*, pages 274–287, 1959.

[58] Frank James Sacherer. *Transverse Space - Charge Effects in Circular Accelerators*. PhD thesis, UC, Berkeley, 1968.

[59] P. Piot, Y.-E Sun, T. J. Maxwell, J. Ruan, E. Secchi, and J. C. T. Thangaraj. Formation and acceleration of uniformly filled ellipsoidal electron bunches obtained via space-charge-driven expansion from a cesium-telluride photocathode. *Phys. Rev. ST Accel. Beams*, 16:010102, Jan 2013.

[60] R. K. Li, K. G. Roberts, C. M. Scoby, H. To, and P. Musumeci. Nanometer emittance ultralow charge beams from rf photoinjectors. *Phys. Rev. ST Accel. Beams*, 15:090702, Sep 2012.

[61] S. B. van der Geer, O. J. Luiten, and M. J. de Loos. Design of a 2-kA, 30-fs RF-photoinjector for waterbag compression. *Int. J. Mod. Phys. A*, 22:4000–4005, 2007.

[62] P. Musumeci, J. T. Moody, R. J. England, J. B. Rosenzweig, and T. Tran. Experimental generation and characterization of uniformly filled ellipsoidal electron-beam distributions. *Phys. Rev. Lett.*, 100:244801, Jun 2008.

[63] C. C. Lin, L. Mestel, and F. H. Shu. The gravitational collapse of a uniform spheroid. *Astrophysical Journal*, 142, November 1965.

[64] R. L. Gluckstern. Scalar potential for charge distributions with ellipsoidal symmetry. Technical Report FNAL-TM–1402, Fermi National Accelerator Laboratory (FNAL), United States, May 1986.

[65] B. W. Reed, T. LaGrange, R. M. Shuttlesworth, D. J. Gibson, G. H. Campbell, and N. D. Browning. Solving the accelerator-condenser coupling problem in a nanosecond dynamic transmission electron microscope. *Review of Scientific Instruments*, 81(5):053706, 05 2010.

[66] Armin Feist, Nora Bach, Nara Rubiano da Silva, Thomas Danz, Marcel Möller, Katharina E. Priebe, Till Domröse, J. Gregor Gatzmann, Stefan Rost, Jakob Schauss, Stefanie Strauch, Reiner Bormann, Murat Sivis, Sascha Schäfer, and Claus Ropers. Ultrafast transmission electron microscopy using a laser-driven field emitter: Femtosecond resolution with a high coherence electron beam. *Ultramicroscopy*, 176:63–73, 2017. 70th Birthday of Robert Sinclair and 65th Birthday of Nestor J. Zaluzec PICO 2017 – Fourth Conference on Frontiers of Aberration Corrected Electron Microscopy.

[67] J. D. Lawson. *The Physics of Charged-Particle Beams*, volume 75 of *International Series of Monographs on Physics*. Oxford University Press, Oxford, England, 2 edition, 1988.

[68] D. W. O. Heddle. Electron optics by p. grivet. *Journal of Applied Crystallography*, 7(4):450–450, 1974.

[69] P. A. Sturrock, L. Marton, and H. Mendlowitz. Static and Dynamic Electron Optics. *Physics Today*, 9(3):26–26, 03 1956.

[70] J. R. Pierce. Rectilinear Electron Flow in Beams. *Journal of Applied Physics*, 11(8):548–554, 08 1940.

[71] Ivan V. Bazarov et al. Comparison of dc and srf photoemission guns for high brightness high average current beam production. *arXiv:1103.5108 [physics.acc-ph]*, 2011.

[72] Cameron JR Duncan, David A Muller, and Jared M Maxson. Lossless monochromation for electron microscopy with pulsed photoemission sources and radio-frequency cavities. *Physical Review Applied*, 14(1):014060, 2020.

[73] Germán Sciaini and RJ Dwayne Miller. Femtosecond electron diffraction: heralding the era of atomically resolved dynamics. *Reports on Progress in Physics*, 74(9):096101, 2011.

[74] Bradley J Siwick, Jason R Dwyer, Robert E Jordan, and RJ Dwayne Miller. An atomic-level view of melting using femtosecond electron diffraction. *Science*, 302(5649):1382–1385, 2003.

[75] S. P. Weathersby, G. Brown, M. Centurion, T. F. Chase, R. Coffee, J. Corbett, J. P. Eichner, J. C. Frisch, A. R. Fry, M. Gühr, N. Hartmann, C. Hast, R. Hettel, R. K. Jobe, E. N. Jongewaard, J. R. Lewandowski, R. K. Li, A. M. Lindenberg, I. Makasyuk, J. E. May, D. McCormick, M. N. Nguyen, A. H. Reid, X. Shen, K. Sokolowski-Tinten, T. Vecchione, S. L. Vetter, J. Wu, J. Yang, H. A. Dürr,

and X. J. Wang. Mega-electron-volt ultrafast electron diffraction at slac national accelerator laboratory. *Review of Scientific Instruments*, 86(7):073702, Jul 2015.

[76] Pengfei Zhu, Y Zhu, Y Hidaka, L Wu, J Cao, H Berger, J Geck, R Kraus, S Pjerov, Y Shen, et al. Femtosecond time-resolved mev electron diffraction. *New Journal of Physics*, 17(6):063004, 2015.

[77] Stephanie Manz, Albert Casandruc, Dongfang Zhang, Yinpeng Zhong, Rolf A Loch, Alexander Marx, Taisuke Hasegawa, Lai Chung Liu, Shima Bayesteh, Hossein Delsim-Hashemi, et al. Mapping atomic motions with ultrabright electrons: towards fundamental limits in space-time resolution. *Faraday discussions*, 177:467–491, 2015.

[78] D. Filippetto and H. Qian. Design of a high-flux instrument for ultrafast electron diffraction and microscopy. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 49(10):104003, 2016.

[79] T Van Oudheusden, EF De Jong, SB Van der Geer, WPEM Op 't Root, OJ Luiten, and BJ Siwick. Electron source concept for single-shot sub-100 fs electron diffraction in the 100 kev range. *Journal of Applied Physics*, 102(9):093501, 2007.

[80] T Van Oudheusden, PLEM Pasmans, SB Van Der Geer, MJ De Loos, MJ Van Der Wiel, and OJ Luiten. Compression of subrelativistic space-charge-dominated electron bunches for single-shot femtosecond electron diffraction. *Physical review letters*, 105(26):264801, 2010.

[81] Jared Maxson, David Cesar, Giacomo Calmasini, Alexander Ody, Pietro Musumeci, and David Alesini. Direct measurement of sub-10 fs relativistic electron beams with ultralow emittance. *Physical review letters*, 118(15):154802, 2017.

[82] Lingrong Zhao, Zhe Wang, Chao Lu, Rui Wang, Cheng Hu, Peng Wang, Jia Qi, Tao Jiang, Shengguang Liu, Zhuoran Ma, et al. Terahertz streaking of few-femtosecond relativistic electron beams. *Physical Review X*, 8(2):021061, 2018.

[83] EC Snively, MAK Othman, M Kozina, BK Ofori-Okai, SP Weathersby, S Park, X Shen, XJ Wang, MC Hoffmann, RK Li, et al. Femtosecond compression dynamics and timing jitter suppression in a thz-driven electron bunch compressor. *Physical review letters*, 124(5):054801, 2020.

[84] Catherine Kealhofer, Waldemar Schneider, Dominik Ehberger, Andrey Ryabov, Ferenc Krausz, and Peter Baum. All-optical control and metrology of electron pulses. *Science*, 352(6284):429–433, 2016.

[85] General particle tracer. *http://www.pulsar.nl/gpt/*.

[86] Klaus Floettmann. Generation of sub-fs electron beams at few-mev energies. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 740:34–38, 2014.

[87] Benno Zeitler, Klaus Floettmann, and Florian Grüner. Linearization of the longitudinal phase space without higher harmonic field. *Physical Review Special Topics-Accelerators and Beams*, 18(12):120102, 2015.

[88] ARRA Bacci, AR Rossi, et al. Ultra-short electron bunches by velocity bunching as required for plasma wave accelerations. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 740:42–47, 2014.

[89] Thomas P. Wangler. *RF Linear Accelerators*. Wiley-VCH Verlag GmbH & Co. KGaA, 2008.

[90] Lingrong Zhao, Heng Tang, Chao Lu, Tao Jiang, Pengfei Zhu, Long Hu, Wei Song, Huida Wang, Jiaqi Qiu, Chunguang Jing, et al. Femtosecond relativistic electron beam with reduced timing jitter from thz driven beam compression. *Physical review letters*, 124(5):054802, 2020.

[91] P. Denham, A. Ody, and P. Musumeci. X-band harmonic longitudinal phase space linearization at the pegasus photoinjector. In *Proc. NAPAC'22*, pages 508–511, New Mexico, USA, 2022.

[92] J. Maxson et al. Direct measurement of sub-10 fs relativistic electron beams with ultralow emittance. *Phys. Rev. Lett.*, 118:154802, 2017.

[93] Emilio A Nanni, Wenqian R Huang, Kyung-Han Hong, Koustuban Ravi, Arya Fallahi, Gustavo Moriena, RJ Dwayne Miller, and Franz X Kärtner. Terahertz-driven linear electron acceleration. *Nature communications*, 6(1):1–8, 2015.

[94] D Cesar, J Maxson, X Shen, KP Wootton, S Tan, RJ England, and P Musumeci. Enhanced energy gain in a dielectric laser accelerator using a tilted pulse front laser. *Optics express*, 26(22):29216–29224, 2018.

[95] Kwang-Je Kim. Rf and space-charge effects in laser-driven rf electron guns. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 275(2):201–218, 1989.

[96] Z Huang. Intrabeam scattering in an x-ray fel driver.

[97] F. Stulle and J. Bergoz. Turbo-ict pico-coulomb calibration to percent-level accuracy. In *Proceedings of FEL2015*, page MOP041. Bergoz Instrumentation, Saint-Genis-Pouilly, France, 2015.

[98] N. Barov, R. H. Miller, and D. J. Newsham. Development of the dual-slot resonance linac. In *Conf.Proc.*, volume C110328, pages 1897–1899, 2011.

[99] X. Wang, D. Xiang, T. Kim, and H. Ihee. Potential of femtosecond electron diffraction using near-relativistic electrons from a photocathode rf electron gun. *Journal of the Korean Physical Society*, 48(3):390–396, 2006.

[100] C. Lu, T. Jiang, S. Liu, R. Wang, L. Zhao, P. Zhu, Y. Liu, J. Xu, D. Yu, W. Wan, Y. Zhu, D. Xiang, and J. Zhang. Imaging nanoscale spatial modulation of a relativistic electron beam with a mev ultrafast electron microscope. *Applied Physics Letters*, 112(11):113102, Mar 2018.

[101] M. Hachmann and K. Flöttmann. Measurement of ultra low transverse emittance at regae. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 829:318–320, 2016. 2nd European Advanced Accelerator Concepts Workshop - EAAC 2015.

[102] D.J. Newsham, N. Barov, and R.H. Miller. Development of the dual-slot resonance linac. In *Proceedings of the 2011 Particle Accelerator Conference (PAC11)*, San Diego, CA, 2011.

[103] J. Rosenzweig and L. Serafini. Transverse particle motion in radio-frequency linear accelerators. *Phys. Rev. E*, 49:1599–1602, Feb 1994.

[104] Colwyn Gulliford and Ivan Bazarov. New method for generating linear transfer matrices through combined rf and solenoid fields. *Phys. Rev. ST Accel. Beams*, 15:024002, Feb 2012.

[105] P. M. Lapostolle, F. Meot, S. Valero, and E. Tanke. Proceedings of the 1990 linear accelerator conference. In *Proc. 1990 Linear Accel. Conf.*, page 315, Albuquerque, N. Mex., 1990. Los Alamos Report LA-112004-C.

[106] J.G. Power, C. Jing, and I. Jovanovic. High frequency bunch train generation from an rf photoinjector at the awa. In *Proceedings of PAC09, Vancouver, BC, Canada*, page MO6RFP047. IEEE, May 2009.

[107] A. Bartnik, C. Gulliford, G. Hoffstaetter, and J. Maxson. Ultimate bunch length and emittance performance of an mev ultrafast electron diffraction apparatus with a dc gun and a multi-cavity srf linac. *Phys. Rev. Accel. Beams*, 25:093401, 2022.

[108] I. Nozawa et al. Measurement of $<$ 20fs bunch length using coherent transition radiation. *Phys. Rev. ST Accel. Beams*, 2014.

[109] Z. Wang et al. Femtosecond diagnostics of ultrashort electron beam using a pinhole-based thz-driven deflector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2014.

[110] G. Andonian et al. Transverse beam size effects on longitudinal profile reconstruction. *arXiv*, 2010.

[111] M. V. Tsarev and P. Baum. Characterization of non-relativistic attosecond electron pulses by transition radiation from tilted surfaces. *New J. Phys.*, 20:033002, 2018.

[112] B. H. Schaap. *Laser-induced generation of coherent soft X-rays by relativistic electron bunches*. PhD thesis, Eindhoven University of Technology, 2023. Research TU/e / Graduation TU/e, Applied Physics and Science Education.

[113] L. Zhang et al. Terahertz streaking of few-femtosecond relativistic electron beams. *Phys. Rev. X*, 8:021061, 2018.

[114] H Dömer and O Bostanjoglo. High-speed transmission electron microscope. *Review of Scientific Instruments*, 74(10):4369–4372, 2003.

[115] G Berruto, I Madan, Y Murooka, GM Vanacore, E Pomarico, J Rajeswari, R Lamb, P Huang, AJ Kruchkov, Y Togawa, et al. Laser-induced skyrmion writing and erasing in an ultrafast cryo-lorentz transmission electron microscope. *Physical review letters*, 120(11):117201, 2018.

[116] John CH Spence. Outrunning damage: Electrons vs x-rays—timescales and mechanisms. *Structural dynamics*, 4(4):044027, 2017.

[117] Chao Lu, Tao Jiang, Shengguang Liu, Rui Wang, Lingrong Zhao, Pengfei Zhu, Yaqi Liu, Jun Xu, Dapeng Yu, Weishi Wan, et al. Imaging nanoscale spatial modulation of a relativistic electron beam with a mev ultrafast electron microscope. *Applied Physics Letters*, 112(11):113102, 2018.

[118] Gerrit H Jansen. Coulomb interactions in particle beams. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 298(1-3):496–504, 1990.

[119] BW Reed, MR Armstrong, ND Browning, GH Campbell, JE Evans, T LaGrange, and DJ Masiel. The evolution of ultrafast electron microscope instrumentation. *Microscopy and microanalysis*, 15(4):272, 2009.

[120] MJ De Loos and SB Van Der Geer. General particle tracer: A new 3d code for accelerator and beamline design. In *5th European Particle Accelerator Conference*, page 1241, 1996.

[121] Peter W Hawkes and Erwin Kasper. *Principles of electron optics*, volume 3. Academic press, 1996.

[122] Jon Orloff. Charged particle optics. *Encyclopedia of Imaging Science and Technology*, 2002.

[123] John David Jackson. *Classical Electrodynamics*. John Wiley & Sons, New York, 3 edition, 1999.

[124] Helmut Wiedemann. *Particle accelerator physics; 3rd ed.* Springer, Berlin, 2007.

[125] G. Pöplau, Ursula van Rienen, Bas van der Geer, and Marieke De Loos. Multigrid algorithms for the fast calculation of space-charge effects in accelerator design. *Magnetics, IEEE Transactions on*, 40:714 – 717, 04 2004.

[126] O Scherzer. The theoretical resolution limit of the electron microscope. *Journal of Applied Physics*, 20(1):20–29, 1949.

[127] Albert Rose. The sensitivity performance of the human eye on an absolute scale∗. *J. Opt. Soc. Am.*, 38(2):196–208, Feb 1948.

[128] E Lessner, X Wang, and P Musumeci. Report of the basic energy sciences workshop on the future of electron sources. *SLAC National Accelerator Laboratory*, 2016.

[129] P Musumeci, J Giner Navarro, JB Rosenzweig, L Cultrera, I Bazarov, J Maxson, S Karkare, and H Padmore. Advances in bright electron sources. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 907:209–220, 2018.

[130] JB Rosenzweig, A Cahill, V Dolgashev, C Emma, A Fukasawa, R Li, C Limborg, J Maxson, P Musumeci, A Nause, et al. Next generation high brightness electron beams from ultrahigh field cryogenic rf photocathode sources. *Physical Review Accelerators and Beams*, 22(2):023403, 2019.

[131] R Legg, JJ Bisognano, M Bissen, R Bosch, D Eisert, M Fisher, M Green, K Kleman, J Kulpin, J Lawler, et al. Status of the wisconsin srf gun. In *Proceeding of this Conference, New Orleans, LA*, 2012.

[132] Jinfeng Yang, Yoichi Yoshida, and Hiromi Shibata. Femtosecond time-resolved electron microscopy. *Electronics and Communications in Japan*, 98(11):50–57, 2015.

[133] Theodore Vecchione, P Denes, RK Jobe, IJ Johnson, JM Joseph, RK Li, A Perazzo, X Shen, XJ Wang, SP Weathersby, et al. A direct electron detector for time-resolved mev electron microscopy. *Review of Scientific Instruments*, 88(3):033702, 2017.

[134] Ahmed H Zewail. 4d ultrafast electron diffraction, crystallography, and microscopy. *Annu. Rev. Phys. Chem.*, 57:65–103, 2006.

[135] Germán Sciaini and RJ Dwayne Miller. Femtosecond electron diffraction: heralding the era of atomically resolved dynamics. *Reports on Progress in Physics*, 74(9):096101, 2011.

[136] D. Filippetto, P. Musumeci, R. K. Li, B. J. Siwick, M. R. Otto, M. Centurion, and J. P. F. Nunes. Ultrafast electron diffraction: Visualizing dynamic states of matter. *Rev. Mod. Phys.*, 94:045004, Dec 2022.

[137] Jie Yang, Xiaolei Zhu, Thomas JA Wolf, Zheng Li, J Pedro F Nunes, Ryan Coffee, James P Cryan, Markus Gühr, Kareem Hegazy, Tony F Heinz, et al. Imaging cf3i conical intersection and photodissociation dynamics with ultrafast electron diffraction. *Science*, 361(6397):64–67, 2018.

[138] Fabrizio Carbone, Peter Baum, Petra Rudolf, and Ahmed H Zewail. Structural preablation dynamics of graphite observed by ultrafast electron crystallography. *Physical review letters*, 100(3):035501, 2008.

[139] Ralph Ernstorfer, Maher Harb, Christoph T Hebeisen, Germán Sciaini, Thibault Dartigalongue, and RJ Dwayne Miller. The formation of warm dense matter: Experimental evidence for electronic bond hardening in gold. *Science*, 323(5917):1033–1037, 2009.

[140] Anshul Kogar, Alfred Zong, Pavel E Dolgirev, Xiaozhe Shen, Joshua Straquadine, Ya-Qing Bie, Xirui Wang, Timm Rohwer, I-Cheng Tung, Yafang Yang, et al. Light-induced charge density wave in late3. *Nature Physics*, 16(2):159–163, 2020.

[141] Gerard Mourou and Steve Williamson. Picosecond electron diffraction. *Applied Physics Letters*, 41(1):44–45, 1982.

[142] Hyotcherl Ihee, Vladimir A Lobastov, Udo M Gomez, Boyd M Goodson, Ramesh Srinivasan, Chong-Yu Ruan, and Ahmed H Zewail. Direct imaging of transient molecular structures with ultrafast diffraction. *Science*, 291(5503):458–462, 2001.

[143] Pengfei Zhu, Y Zhu, Y Hidaka, L Wu, J Cao, H Berger, J Geck, R Kraus, S Pjerov, Y Shen, et al. Femtosecond time-resolved mev electron diffraction. *New Journal of Physics*, 17(6):063004, 2015.

[144] P Musumeci, JT Moody, CM Scoby, MS Gutierrez, and M Westfall. Laser-induced melting of a single crystal gold sample by time-resolved ultrafast relativistic electron diffraction. *Applied Physics Letters*, 97(6), 2010.

[145] T Van Oudheusden, PLEM Pasmans, SB Van Der Geer, MJ De Loos, MJ Van Der Wiel, and OJ Luiten. Compression of subrelativistic space-charge-dominated electron bunches for single-shot femtosecond electron diffraction. *Physical review letters*, 105(26):264801, 2010.

[146] WH Li, CJR Duncan, MB Andorf, AC Bartnik, E Bianco, L Cultrera, A Galdi, M Gordon, M Kaemingk, CA Pennington, et al. A kiloelectron-volt ultrafast electron micro-diffraction apparatus using low emittance semiconductor photocathodes. *Structural Dynamics*, 9(2), 2022.

[147] Hyun Woo Kim, Nikolay A Vinokurov, In Hyung Baek, Key Young Oang, Mi Hye Kim, Young Chan Kim, Kyu-Ha Jang, Kitae Lee, Seong Hee Park, Sunjeong Park, et al. Towards jitter-free ultrafast electron diffraction technology. *Nature photonics*, 14(4):245–249, 2020.

[148] Fengfeng Qi, Zhuoran Ma, Lingrong Zhao, Yun Cheng, Wenxiang Jiang, Chao Lu, Tao Jiang, Dong Qian, Zhe Wang, Wentao Zhang, et al. Breaking 50 femtosecond resolution barrier in mev ultrafast electron diffraction with a double bend achromat compressor. *Physical review letters*, 124(13):134803, 2020.

[149] Mark W Tate, Prafull Purohit, Darol Chamberlain, Kayla X Nguyen, Robert Hovden, Celesta S Chang, Pratiti Deb, Emrah Turgut, John T Heron, Darrell G Schlom, et al. High dynamic range pixel array detector for scanning transmission electron microscopy. *Microscopy and Microanalysis*, 22(1):237–249, 2016.

[150] Theodore Vecchione, P Denes, RK Jobe, IJ Johnson, JM Joseph, RK Li, A Perazzo, X Shen, XJ Wang, SP Weathersby, et al. A direct electron detector for time-resolved mev electron microscopy. *Review of Scientific Instruments*, 88(3), 2017.

[151] Fuhao Ji, Daniel B Durham, Andrew M Minor, Pietro Musumeci, Jorge G Navarro, and Daniele Filippetto. Ultrafast relativistic electron nanoprobes. *Communications Physics*, 2(1):54, 2019.

[152] M Gordon, WH Li, MB Andorf, AC Bartnik, CJR Duncan, M Kaemingk, CA Pennington, IV Bazarov, Y-K Kim, and JM Maxson. Four-dimensional emittance measurements of ultrafast electron diffraction optics corrected up to sextupole order. *Physical Review Accelerators and Beams*, 25(8):084001, 2022.

[153] Y. Murooka, N. Naruse, S. Sakakihara, M. Ishimaru, J. Yang, and K. Tanimura. Transmission-electron diffraction by MeV electron pulses. *Applied Physics Letters*, 98(25):251903, 06 2011.

[154] Lucio Rossi and Luca Bottura. Superconducting magnets for particle accelerators. *Reviews of accelerator science and technology*, 5:51–89, 2012.

[155] Martin Reiser and Patrick O'Shea. *Theory and design of charged particle beams*, volume 312. Wiley Online Library, 1994.

[156] D Cesar, J Maxson, P Musumeci, Y Sun, J Harrison, P Frigola, FH O'Shea, H To, D Alesini, and RK Li. Demonstration of single-shot picosecond time-resolved mev electron imaging using a compact permanent magnet quadrupole based lens. *Physical review letters*, 117(2):024801, 2016.

[157] J. K. Lim, P. Frigola, G. Travish, J. B. Rosenzweig, S. G. Anderson, W. J. Brown, J. S. Jacob, C. L. Robbins, and A. M. Tremaine. Adjustable, short focal length permanent-magnet quadrupole based electron beam final focus system. *Phys. Rev. ST Accel. Beams*, 8:072401, Jul 2005.

[158] European Synchrotron Radiation Facility (ESRF). *RADIA Technical Reference Manual*. ESRF, Grenoble, France, Year of publication or last update. `URLorotherrelevantdetails`.

[159] Harald A. Enge. Effect of Extended Fringing Fields on Ion-Focusing Properties of Deflecting Magnets. *Review of Scientific Instruments*, 35(3):278–287, 03 1964.

[160] R. Baartman. Quadrupole shapes. *Phys. Rev. ST Accel. Beams*, 15:074002, Jul 2012.

[161] Alex D'Audney. *Ultra-high resolution pulsed-wire magnet measurement system, an.* PhD thesis, Colorado State University, 2016.

[162] D. Arbelaez, T. Wilks, A. Madur, S. Prestemon, S. Marks, and R. Schlueter. A dispersion and pulse width correction algorithm for the pulsed wire method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 716:62–70, 2013.

[163] David Alesini, Antonio Battisti, Massimo Ferrario, Luca Foggetta, Valerio Lollo, Luca Ficcadenti, Valerio Pettinacci, Sean Custodio, Eylene Pirez, Pietro Musumeci, et al. New technology based on clamping for high gradient radio frequency photogun. *Physical Review Special Topics-Accelerators and Beams*, 18(9):092001, 2015.

[164] M.J. De Loos and Van der Geer. General particle tracer: A new 3d code for accelerator and beamline design. *5th European Particle Accelerator Conference*, page 1241, 1996.

[165] F. Salvat, J. D. Mart˘nez, R. Mayol, and J. Parellada. Analytical dirac-hartree-fock-slater screening function for atoms (z=1–92). *Phys. Rev. A*, 36:467–474, Jul 1987.

[166] Francesc Salvat and Ricardo Mayol. Elastic scattering of electrons and positrons by atoms. schrödinger and dirac partial wave analysis. *Computer Physics Communications*, 74(3):358–374, 1993.

[167] P. Denham and P. Musumeci. Space-charge aberrations in single-shot time-resolved transmission electron microscopy. *Phys. Rev. Appl.*, 15:024050, Feb 2021.