# ELECTRON BEAMS FROM NEEDLE PHOTOCATHODES AND A NEW THEORY OF THE SMITH-PURCELL FREE-ELECTRON LASER

By

Charles Herbert Boulware, III

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR IN PHILOSOPHY

in

Physics

May, 2007

Nashville, Tennessee

Approved:

Professor Charles A. Brau

Professor Leonard C. Feldman

Professor Richard F. Haglund

Professor Marcus H. Mendenhall

Professor David W. Piston

To Sarynna,

whose love is the coal

for my engine.

“The more I dive into this matter of whaling, and push my researches up to the very spring-head of it so much the more am I impressed with its great honorableness and antiquity; and especially when I find so many great demi-gods and heroes, prophets of all sorts, who one way or other have shed distinction upon it, I am transported with the reflection that I myself belong, though but subordinately, to so emblazoned a fraternity.”

– Herman Melville

# ACKNOWLEDGMENT

Thanks to everyone who helped me in my several years at Vanderbilt. Special thanks go to the members of my research group. Charlie Brau has been an advisor in every sense of the word, and what physical intuition I have gained in my time here was learned by example from him. Jonathan Jarvis and Heather Andrews have helped me in more ways than I could count, and it will have to suffice for me to just to say thank you to them for all the hard work they did. We had fun, too.

Thanks also to my Ph.D. committee members who improved the quality of this thesis greatly by their comments and questions. Their toughest questions always seemed motivated by a sincere desire to help me solve problems.

Once upon a time, several of the physics graduate students would gather at the NanoCafe for afternoon coffee and I thank them for passing on to me the strange mixture of stress, freedom, depression, curiosity, cynicism and joy that is the grad student experience. Particularly I have to thank Carlos Hernandez Garcia, who preceded me in the needle photocathode project and who built things I used every day. Jonathan Newport also deserves a tip of the hat for his contributions, also having put together several widgets that got used long after he left Vanderbilt.

John Fellenstein and Bob Patchin in the shop deserve credit as co-designers on my experiment. I have the highest respect for their craftsmanship, and cannot thank them enough for talking me out of so many bad ideas.

I could not possibly list all the people who have impacted my life outside in physics in these years, but I know that everything I do is a product of all their influences. Nashville is quite a town, it turns out.

Sarynna, thank you for putting up with me when I was preoccupied with my wandering electrons and for sharing my happiness on the days they behaved. I love you with all my heart and I am so glad that I am taking you with me!

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER I

## INTRODUCTION

Electrical oscillations and radiated heat waves were understood to be parts of a single continuous electromagnetic spectrum long before it was technologically possible to produce and observe radiation at wavelengths spanning the range between the two sets of phenomena. Even today, there is a region of the spectrum that is not well served by practical sources. This region is called the terahertz (THz) gap because it lies between the fastest available transistors at around 0.5 THz and the lowest frequency conventional lasers around 30 THz. The lack of sources in this region is not due to a lack of applications for THz light, as there are many materials (including water) that have interesting dielectric properties at THz frequencies. This thesis concerns the theory of one compact source of THz radiation, the Smith-Purcell free-electron laser (SPFEL), and the development of high-brightness electron sources for driving such a device.

This thesis is organized into four principal sections, the chapters II through V which follow this introduction. Chapter II discusses the currently available THz sources and detectors, as well as a range of possible applications for THz light. This chapter also reviews the Smith-Purcell effect, which is the radiation output mechanism for the SPFEL. There have been several proposals for a coherent Smith-Purcell (SP) radiation source and these are discussed. The primary motivation for this thesis is the work done at Dartmouth College using a modified scanning electron microscope (SEM) to produce SP radiation whose power output varies nonlinearly with the beam current above a certain threshold. The end of Chapter II discusses these experiments.

Chapter III develops our new theory of the SPFEL. The bunching mechanism for this kind of laser is the interaction between the electron beam and an evanescent wave that propagates along the surface of the grating. Previous theories have not taken into account the dispersion of these evanescent waves, but we find that the group velocity of the evanescent wave is key to the device operation. The group velocity is also the velocity for energy flow in the wave, and we show that the wave energy can move in the same direction as the electrons or in the opposite direction. When the energy moves in the same direction as the beam, the wave grows as a convective instability in the manner of a traveling-wave tube (TWT) amplifier. When the energy flow is in the opposite direction as that of the electrons, the device more closely resembles a backward-wave oscillator (BWO). From our theory, we calculate the gain length for the TWT amplifier regime and the time growth rate for the BWO regime.

The development of new photocathodes to produce high quality electron beams is discussed in Chapter IV. We have developed several kinds of high-field metal photocathodes based on protruding wires of tungsten or yttrium with various tip shapes. We show that understanding the performance of these photocathodes requires some sort of model for the electric field on the tip and at least a simple model of the photoelectric emission efficiency. We discuss how the cathodes are made, smoothed, and cleaned, and present photoemission data at several drive laser wavelengths.

Chapter V discusses the electron beams that we obtain from needle photocathodes. The theoretical description of electron beam quality in terms of emittance and brightness is presented, as are the stringent requirements on the electron beam for an SPFEL. The beam quality of thermionic and photoelectric electron sources is calculated and contrasted. Then, using the concepts introduced to describe the electron beam, we calculate the spot size to which the beam can be focused and discuss contributions to that spot size arising from factors other than intrinsic beam quality at the source. The electron beam apparatus that has been built at Vanderbilt is described, including the focusing solenoid which serves as the main electron optical element in the system. We conclude the chapter by presenting examples of measured spot size data and interpretation of that data in light of the theoretical description of the electron source and focusing system.

A short conclusion and summary follows the principal sections as Chapter VI. Our contributions to the development of needle photocathodes are briefly discussed. Though the purpose of our research has been to make high-brightness electron beams for SPFELs, high-field metal photocathodes are also possibly relevant to X-ray FELs, which also require high-brightness beams. The discrepancies between the Dartmouth experiments and our theory of the SPFEL are then summarized. The final chapter concludes with thoughts on the continuing efforts to resolve experimentally the disagreement between the Dartmouth results and our theory.

# CHAPTER II

## TERAHERTZ RADIATION AND THE SMITH-PURCELL EFFECT

This chapter will discuss the terahertz (THz) region of the electromagnetic spectrum, which lies between electric waves and infrared light. The so-called THz gap has been recognized for over 75 years, and though there has been much progress in sources and detectors in the region, there are many proposed applications for light at these wavelengths that cannot be fully realized because the source technology does not exist.

Several current sources of THz radiation will be discussed including conventional lasers and a wide variety of electron-beam based technologies. The possible applications in biology and medicine, as well as in fundamental physics, will also be reviewed. The dielectric properties of water display many strong features in the THz range and these will be briefly discussed.

One very promising compact source of THz radiation has been demonstrated at Dartmouth College. The Dartmouth group uses a modified scanning electron microscope to produce Smith-Purcell (SP) radiation from an electron beam passing over a metal grating. Several proposals to produce coherent SP radiation were written in the wake of the successful operation of the first free-electron laser. Most of these involved adding external feedback elements to return the SP radiation back to the grating.

The work of the Dartmouth group shows that the SP radiation becomes nonlinear in the electron beam current above a certain threshold, without any external feedback.

They interpret their superradiant emission results as coherent amplification in the manner of a traveling-wave tube. Details of these experiments will be discussed in this chapter, as they are a primary motivation for the theoretical efforts in this thesis as well as the experimental program to produce high-brightness electron beams.

## Terahertz radiation

### *The THz gap*

The technological gap between electric waves and the infrared was recognized as early as the 1920s. Already, both phenomena were clearly understood as electromagnetic radiation, but experimentally, there remained an inaccessible region of the spectrum between the infrared and the electric. At that time, Nichols and Tear identified an "unexplored region" to exist between 1 to 30 THz, between filtered mercury arc lamps and Hertzian oscillators [1]. They recognized that the optical properties of many substances, with water as an important example, have significant variations across this region.

The THz gap today is between the fastest transistors up to several hundred GHz and the longest wavelength lasers, which extend down to almost 1 THz. Moving closer to the gap from either side, the options for sources are restricted, and therefore the range of desired qualities such as compactness, tunability, and spectral brightness is also limited. Fast, sensitive detection in the submillimeter and infrared range is dominated by bolometers, but this portion of the electromagnetic spectrum is also covered to some degree by pyroelectric detectors, diodes, and Golay cells. In the next sections, we will consider the sources and detectors available in the THz region, generation and detection of THz radiation, quality of available sources, and sensitivity of available detectors.

### *THz sources*

Though the quality and flexibility of sources available is not ideal, several technologies are currently used to produce electromagnetic radiation in the THz. These sources fit into several broad categories: solid- and gas-state lasers, nonlinear optical devices, and electron beam-based devices.

Molecular gas lasers exist in the far infrared (FIR) and THz frequencies. Carbon dioxide ($CO_2$) lasers were among the first to be discovered and can be operated continuously at high power on several emission lines around 200 THz [2]. Other molecular gas lasers also provide several intense line sources, and in the THz range. Methyl fluoride lasers pumped by $CO_2$ lasers provide MW pulsed power on several lines around 0.3 THz ($C^{13}H_3F$ ) and one at 0.6 THz ($C^{12}H_3F$). Heavy water ($D_2O$) lasers provide power in the tens of MWs at 0.78, 2.7, and 4.6 THz. The application of these sources to spectroscopy has been limited by the lack of line tuning [3].

Quantum cascade lasers (QCLs) are semiconductor heterostructure devices that sandwich micron-thick layers of doped and undoped semiconductor along with injector and collector structures to create a series of population-inverted units. The layers also serve as a waveguide to produce optical confinement and control attenuation. Thermal processes compete against the collection/injection process, and so these lasers must be operated at low temperatures. The lasing wavelength is set by the difference between the energy levels in each sandwich. A device of this kind was reported (in 2002) to produce

2 mW of 4.4-THz radiation when operated at temperatures of 8 K and hundreds of μW as high as 50 K [4].

Nonlinear optical techniques have also been exploited to produce THz light. Optical rectification of femtosecond laser pulses using slabs of optically active material has been demonstrated. As an example, Bonvalet et al. produced near-single-cycle pulses by hitting a 0.1 mm GaAs plate with a ~10 fs pulse from a Ti:sapphire laser. Optical rectification of this pulse by the second-order polarizability of the slab produced ~10 fs electromagnetic pulses with broadband spectra extending to the THz region. By this method, they measured 30 nW average power between 20-38 THz using incident pulses (10-15 fs at 100 MHz) of 175 mW average power at ~800 nm. This suggests a peak power of 20 mW [5]. More recently, peak powers near 200 W have been measured in a similar system by line focusing the pump beam with a cylindrical lens and taking advantage of self-focusing within the crystal. This line-source geometry has been demonstrated in optical rectification from $LiNbO_3$ to produce broadband fs pulses centered around 1 THz (the full-width half maximum (FWHM) of the spectrum is ~1.5 THz) [6].

Various THz sources are based on radiation from electron beams. For example, coherent synchrotron radiation (CSR) provides a strong source of broadband THz. Synchrotron radiation comes from accelerating electrons through bending magnets. When the electrons are bunched, this radiation can become coherent at wavelengths down to the bunch length. CSR from a storage ring at BESSY was recently reported to give broadband radiation with tens of watts average power over a range from 1 mm down to about 200 μm (0.3 to 1.5 THz) [7]. Using a superconducting linac to accelerate electrons to 40 MeV, researchers at Jefferson Lab are able to produce relativistic electron bunches with picosecond bunch lengths. These electrons radiate broadband THz with a spectral power of ~1 W per $cm^{-1}$ from 0.2-1 THz (average total power ~20 W) [8].

Conventional free-electron lasers (FELs) are very good sources for THz radiation. The FIR FEL at UC-Santa Barbara produces kW of output power in microsecond pulses tunable over a range from 1-5 THz. UCSB operates a 6-MV electrostatic accelerator to feed the FIR FEL and two other FEL devices [9]. A FEL facility in Novosibirsk, Russia was recently commissioned with the express intent to produce THz radiation. This new laser provides 200 W of average power tunable from 1.7-2.5 THz with measured linewidths <1%. The electron beam for the Novosibirsk FEL is provided by a 50-MeV accelerator in an energy-recovery configuration [10]

A compact source of radiation extending into the lower THz range is the backward wave oscillator (BWO). The BWO is an electron beam device in which a magnetically confined electron beam is injected into a corrugated slow-wave structure. The slow wave on the corrugated waveguide moves with a phase velocity synchronous with the electrons and with a group velocity in the direction opposite of the electron-beam motion. The slow wave bunches the electron beam and the bunching is carried toward the downstream end of the device by the electron beam velocity. In this way, the slow wave starts up from a bunched beam and the device goes into oscillation above a certain electron beam current [11]. BWOs are being used for THz and sub-THz spectroscopy from 0.03-1.5 THz [12], and have been used in some simple THz imaging applications as well [13].

Another compact alternative for broadband THz comes from coherent transition radiation. Electrons emit radiation as they cross a boundary between two regions of different dielectric constant. With short bunches of electrons, this radiation can be coherent down to wavelengths comparable to the bunch length. By focusing a laser pulse onto a supersonic helium jet, electrons can be accelerated out of the plasma stream into the vacuum. As they cross the vacuum-plasma boundary, they emit coherent transition radiation. Using a terawatt laser (0.4 J in 50-fs pulses) to accelerate the electrons, nanojoules of THz radiation have been measured in a 30-mrad collection angle, and the possibility has been suggested that this could be scaled up into a device that would produce 100 μJ/pulse (into a somewhat larger cone angle) [14].

Another possibility for a small THz source is one based on high electron mobility transistors. In these transistors, the transistor channel is filled with electron plasma that resonates at THz frequencies. The radiation actually comes from a plasma instability in the field effect transistor channel, driven by the current in the switch. Average radiation power in the nW range is observed from an InGaAs transistor with a 60-nm gate. Because the transistor is so small, this corresponds to several W/cm$^2$ [15].

One group at Dartmouth College succeeded in getting superradiant THz emission from an electron beam passing above a metal grating. This emission was interpreted as coherent Smith-Purcell radiation. This experiment will be discussed in detail below, as it presents an exciting, compact source for THz radiation. They observed broadly tunable, narrow-band emission near 2 THz. The Dartmouth experiment is a primary motivation for this thesis, and the requirements for electron beam quality presented by a SP radiation source are the connection between THz radiation production and the electron photocathode development that will be described.

### *THz detectors*

Because of the lack of good THz sources, and because the atmosphere screens astronomical sources of THz rays, detector technology is also less mature in the THz gap. One of the most important advances beyond the simple wire bolometer was the Golay cell, which became available after 1947. This cell is simply a gas-based calorimeter, where a gas expands against a flexible diaphragm after absorbing infrared radiation. The deflection of the membrane is measured by a change in capacitance between the membrane itself and an electrode, or by reflecting laser light from a surface mechanically connected to the membrane [16]. A modern Golay cell distributed by the Russian company Tydex has an optical noise equivalent power (NEP) at 15 Hz of $<100\ \mathrm{pW}\big/\sqrt{\mathrm{Hz}}$, 40 ms response speed, and operates at room temperature.

Bolometers have advanced considerably from the strips of iron used at the turn of the 20th century. Modern bolometers use detecting elements made of InSb, Ge, or Si, and operate at cryogenic temperatures. Our InSb bolometer built by QMC Instruments has an optical NEP of $3.1\ \mathrm{pW}\big/\sqrt{\mathrm{Hz}}$ and is much faster than a Golay cell with response time of 1 μs. This bolometer operates at liquid He temperature (4.2 K or below). The signal-to-noise issues involved in using this bolometer to detect spontaneous (incoherent) SP radiation are discussed below.

For imaging, several groups are working on micromachined antennas etched in silicon. Tiny aerials can be made on silicon wafers using techniques borrowed from the integrated circuit industry. These types of detectors were very expensive when first developed and used only in detectors designed for space flight. More recently, arrays of hundreds of micro-antennas functioning at 100 GHz have been fabricated. These arrays currently have a NEP of $25\ \mathrm{pW}/\sqrt{\mathrm{Hz}}$ (for the whole array, comparable in total area to the other detectors discussed) and a fast response time of 400 ns [17].

*THz applications*

The applications proposed for THz light sources are far ranging and include the fundamental research on water, THz imaging, atmospheric sensing and astronomy, plasma diagnostics, electron spin resonance, and various biomedical applications. As happened when other regions of the electromagnetic spectrum became accessible, it seems likely that new, compact, bright sources of THz light will create new applications and open up new avenues of research.

The dielectric constant of distilled water varies from ~80 at 1 GHz to 2 above 1 THz. Water also absorbs THz radiation strongly, with an absorption coefficient greater than $5\times10^{4}\ \mathrm{m}^{-1}$ at 3 THz. The broad absorption feature in this region varies with wavelength according to the Debye relaxation model [18]. The molecular mechanism for relaxation in a hydrogen-bonded solvent like water is more complicated than just single-molecule rotation, and may include reorganization of five-molecule tetrahedral clusters [19]. Detailed THz spectroscopy of water may illuminate these processes and lay the groundwork for more sophisticated spectroscopy and imaging in aqueous environments.

Direct THz imaging has been proposed for food testing, industrial quality control, and security applications. The lack of strong, bright sources has meant long acquisition times for images except in small, controlled experiments. But even with the broadband sources available, there have been impressive demonstrations. Because of the strong absorption bands in water, moisture-mapping images can be taken in the THz. Using optical rectification sources, pictures have been taken of water content inside a drying leaf, for example [20]. Recently, other groups have used THz from a parametric oscillator to examine large-scale integrated circuits for faults and to nondestructively identify drugs (distinguishing aspirin from MDMA and methamphetamine in this example) [21].

With the launch of the Stratospheric Observatory For Infrared Astronomy (SOFIA) program, molecular signatures and atmospheric absorption in the THz are becoming important subfields of space and atmospheric science. The SOFIA 2.5-m telescope is designed to look at light from 1-1,000 THz. It is estimated that 98% of the photons emitted since the Big Bang have been in the far-infrared [22]. Much is absorbed by the atmosphere, but airborne experiments (like SOFIA, which is carried aloft by a modified Boeing 747) to detect the radiation are coming online. Interpreting the spectra received from heavenly objects will certainly motivate better databases of THz spectra, and therefore brighter sources and detectors.

A strong THz source could also find application in fusion plasma diagnostics. The refractive index of a plasma is dependent on the electron density, so lasers can be used to make interferometric measurements of the varying density of electrons in a plasma tokamak. An upper limit on the wavelength of the laser is set by the cut-off frequency in the plasma, but there are arguments for working near this limit, specifically vibration and fringe-counting difficulties. The optimum range for probing fusion plasma electron density turns out to be 100-200 μm, or 1.5-3 THz. Early choices for diagnostic systems were DCN lasers at 119 μm or $H_20$ lasers at 118.6 μm [23]. A strong tunable THz source could make an impact in this field.

Electron spin resonance experiments take advantage of the Zeeman splitting between electrons whose spin is aligned and those that are anti-aligned to a strong magnetic field. The splitting and probe frequency both increase with the magnetic field. Standard ESR experiments are now performed with X-band sources at around 10 GHz and magnetic fields around 1 T. Submillimeter experiments are being done with Gunn diodes and BWOs up to 1 THz and using discrete lines from FIR lasers up to 7 THz [24, 25]. This field would benefit from a tunable narrow-linewidth THz source.

The strong absorption of THz radiation by water also leads to the possibility of great contrast in medical imaging with submillimeter waves. The low energy of THz photons (1-12 meV) seems to preclude DNA damage in cells. It is possible that the conformational states of proteins could be assessed with THz spectroscopy, or that different kinds of tissues could be distinguished with THz images. Already, THz waves have been used to identify some kinds of dental problems, to assess the severity of some burns, and to analyze the healing and scarring of wound tissues, sometimes even underneath the dressings [18]. Imaging of subcutaneous basal cell carcinoma has been demonstrated, and may prove to be an important advance in the treatment of these lesions [26]. Research on biomedical applications of this radiation is in its early stages, but again, a compact, bright source could be a real catalyst for further development.

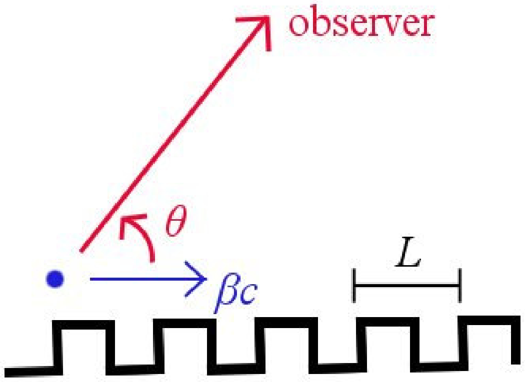

Figure 2-1. Diagram defining the parameters in the Smith-Purcell relation.

## Smith-Purcell radiation

### *Discovery and early theories*

E. M. Purcell suggested in the 1950s that electrons passing near the metal surface of a diffraction grating induce surface currents on the grating, and that these surface currents in turn radiate electromagnetic radiation. From a simple wavefront construction, S. J. Smith and Purcell showed that the wavelength of this radiation should be

$$\lambda = \frac{L}{|n|}\left(\beta^{-1} - \cos\theta\right) \tag{2-1}$$

where $L$ is the period of the grating, $\beta$ is the velocity of the electron beam, $\theta$ is the observation angle as shown in Figure 2-1, and $n$ is the diffraction order. Using a Van de Graaff accelerator and a 5-μA beam, they observed this effect in 1953 [27]. Though the original physical insight that prompted the discovery of the effect was in terms of induced surface currents in the grating, one of the first detailed theories of the effect by di Francia used a different description. Di Francia expanded the field of the moving electrons in a set of evanescent waves and then allowed the waves to be scattered by the grating [28]. The grating scatters the evanescent waves into plane waves, which correspond to the SP radiation.

The intensity of the radiation was later analyzed in terms of the surface current induced by the electron beam [29, 30]. In this approach the surface current induced by the moving electrons is modulated by the periodicity of the grating surface. There exist spatial components of the current modulation with superluminal phase velocities and these are considered source terms for the radiation.

The evanescent wave approach of di Francia was used to obtain rigorous solutions for SP radiation from nonrelativistic electrons above an infinite grating by van den Berg [31, 32]. Van den Berg's method uses a Green's function formulation of the 2D problem to calculate the SP radiation for an arbitrary grating profile. SP radiation from a rectangular grating profile was also calculated using a separation of variables approach [33]. The validity of these various models for relativistic electrons has also been discussed [34].

### *Proposals for a source of coherent SP radiation*

After the free-electron laser was demonstrated [35], the possibility of a device utilizing coherent SP radiation was considered with various feedback mechanisms. Some formulations were based on periodic corrugations in closed waveguide structures [36], and others on an open grating waveguide together with feedback mirrors [37].

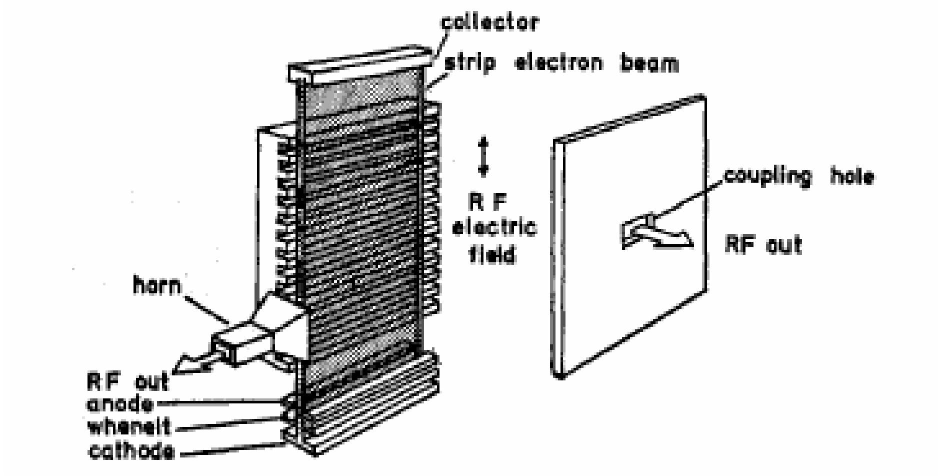

Figure 2-2. The ledatron (graphic from Mizuno et al., 1973), showing the Fabry-Perot output through the roof above the grating and the BWO output from the upstream end of the grating.

One device, called the ledatron, has been shown to function as either a BWO or a SP coherent device with a flat mirror over the grating constituting a Fabry-Perot cavity [38]. Adjustments of the mirror are shown to select between BWO operation, dominated by a surface mode, and the Fabry-Perot mode, dominated by the SP radiation at 90°. A sketch of the ledatron from Ref. [38] is shown in Figure 2-2. The authors discuss the two operation modes separately, and do not consider using the bunched beam from the surface wave interaction as a source of coherent SP radiation.

*The Dartmouth superradiant SP experiments*

A primary motivation for the work in the thesis comes from the nonlinear SP radiation experiments performed at Dartmouth College [39]. These experiments used a modified scanning electron microscope (SEM) and a rectangularly grooved grating to build a SP radiation source. The SEM produced an electron beam of up to 1 mA at 20-40

Table 2-1. Physical parameters for one of the gratings for which nonlinear SP radiation was observed in the Dartmouth experiments.

| Dartmouth grating parameters | |
|---|---|
| grating period | 173 μm |
| groove width | 62 μm |
| groove height | 100 μm |
| grating length | 12.7 mm (73 grating periods) |

keV beam energy from a thermionic tungsten hairpin cathode. They were able to focus the electron beam spot to a root-mean-square (rms) diameter of ≥20 μm. The emittance of the electron beam (the concept of beam emittance will be discussed in a later chapter) was measured using a 0.5 mm slit to be on the order of $10^{-7}$ m, varying by less than a factor of 2 as the current was changed.

The SEM setup allowed the Dartmouth group to look for SP radiation only at an observation angle of 90° (as defined in Figure 2-1). The SEM steering coils were used to sweep the electron beam in the direction perpendicular to the grating surface at a rate of 200 Hz. Sweeping the beam in this way imposes a modulation on the radiation signal, allowing the use of a lock-in amplifier. The theory of van den Berg [33] was used to design a rectangular grating optimized for first order SP radiation at that angle. The parameters of one of the gratings used are shown in Table 2-1.

The Dartmouth group observed a signal in the bolometer that increased nonlinearly with the electron beam current as shown in Figure 2-3. The experiments demonstrated linear emission up to a certain threshold current, above which the SP signal increased nonlinearly with current. The threshold current was seen to increase for larger beam diameters. In Figure 2-3, $\alpha$ is the exponent of the curve used to fit the data on the

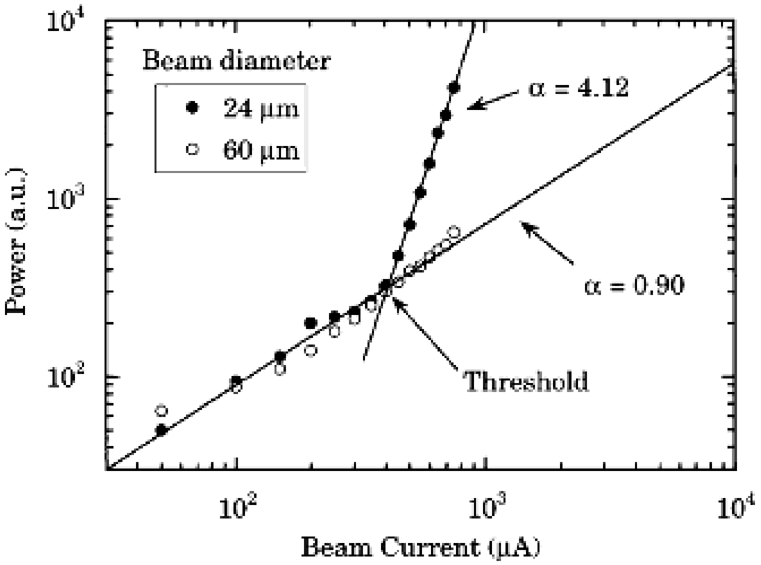

Figure 2-3. Bolometer signal data from the Dartmouth experiments using the 173 μm-period grating (graphic from Urata et al., 1998).

log-log plot. The data shown in the figure increase according to the fourth power of the beam current above threshold, but this exponent was seen to vary from 3 to 6 in the various experiments performed.

The SP radiation was measured to be strongly polarized above and below the current threshold, and the wavelength was observed to be the predicted first-order SP radiation wavelength according to Equation (2-1) both above and below the current threshold. A change in the linewidth of the radiation between the linear and nonlinear regimes was not reported. The nonlinear radiation was interpreted to be the result of interaction between the electron beam and a synchronously moving surface evanescent wave. The energy transferred from the electron beam to the surface wave is redistributed by the grating into the radiating orders. The frequency of the evanescent wave is not calculated or discussed.

The Dartmouth group considered their nonlinear SP radiation device to function in the manner of a traveling-wave tube. Detailed experiments showed that the power from the grating just below the current threshold displayed a sinusoidal dependence on the third root of the beam current [40]. This oscillation was interpreted as the beating of three waves (two space charge waves and one resonant structure wave) as described by Pierce for traveling-wave-tubes [41]. They used a modified scanning electron microscope with a hairpin tungsten thermionic cathode for their experiments, and suggested that a higher quality electron source could be used to build a useful source of spectrally bright, tunable THz radiation.

# References

1 E. F. Nichols and J. D. Tear, Phys. Rev. **21**, 587 (1923).

2 C. K. N. Patel, Physical Review **136**, A1187 LP (1964).

3 T. A. DeTemple, in *Infrared and Millimeter Waves*, edited by K. J. Button, 1979), Vol. 1, p. 129.

4 R. Kohler, A. Tredicucci, F. Beltram, et al., Nature **417**, 156 (2002).

5 A. Bonvalet, M. Joffre, J. L. Martin, et al., Appl. Phys. Lett. **67**, 2907 (1995).

6 A. G. Stepanov, J. Hebling, and J. Kuhl, in *Applied Physics B: Lasers and Optics*, 2005), Vol. 81, p. 23.

7 M. Abo-Bakr, J. Feikes, K. Holldack, et al., Phys. Rev. Lett. **90**, 094801 (2003).

8 G. L. Carr, M. C. Martin, W. R. McKinney, et al., Nature **420**, 153 (2002).

9 G. Ramian, Nucl. Inst. Meth. Phys. Res. A **318**, 225 (1992).

10 V. P. Bolotin, N. A. Vinokurov, D. A. Kayran, et al., in *FEL2004* (FEL2004 website, Trieste, Italy, 2004).

11 R. W. Gould, Electron Devices, IEEE Transactions on **2**, 37 (1955).

12 B. Gorshunov, M. Dressel, T. Kakeshita, et al., in *Infrared and Millimeter Waves, 2004 and 12th International Conference on Terahertz Electronics, 2004. Conference Digest of the 2004 Joint 29th International Conference on*, 2004), p. 91.

13 A. Dobroiu, M. Yamashita, Y. N. Ohshima, et al., in *Infrared and Millimeter Waves, 2004 and 12th International Conference on Terahertz Electronics, 2004. Conference Digest of the 2004 Joint 29th International Conference on*, 2004), p. 825.

14 W. P. Leemans, C. G. R. Geddes, J. Faure, et al., Phys. Rev. Lett. **91**, 074802 (2003).

15 W. Knap, J. Lusakowski, T. Parenty, et al., Appl. Phys. Lett. **84**, 2331 (2004).

16 M. J. E. Golay, Rev. Sci. Inst. **18**, 357 (1947).

17 A. J. Miller, A. Luukanen, and E. N. Grossman, in *Proceedings of the SPIE*, edited by R. J. a. D. L. W. Hwu, 2004), Vol. 5411, p. 18.

18 P. H. Siegel, IEEE Trans. on Microwave Theory and Techniques **52**, 2438 (2004).

19 N. Agmon, J. Phys. Chem. **100**, 1072 (1996).

20 B. B. Hu and M. C. Nuss, Optics Lett. **20**, 1716 (1995).

21 K. Kawase, Optics and Photonics News **15**, 34 (2004).

22 P. H. Siegel, IEEE Trans. on Microwave Theory and Techniques **50**, 910 (2002).

23 N. C. Luhmann and W. A. Peebles, Rev. Sci. Inst. **55**, 279 (1983).

24 J. Krzystek, S. A. Zvyagin, A. Ozarowski, et al., Journal of Magnetic Resonance **178**, 174 (2006).

25 H. Ohta, S. Okubo, K. Kawakami, et al., J. Phys. Soc. Jpan. **72**, 26 (2003).

26 R. M. Woodward, V. P. Wallace, R. J. Pye, et al., Journ. Invest. Derm. **120**, 72 (2003).

27 S. J. Smith and E. M. Purcell, Phys. Rev. **92**, 1069 (1953).

28 G. T. di Francia, Nuovo Cimento **16**, 61 (1960).

29 J. H. Brownell, J. Walsh, and G. Doucas, Phys. Rev. E **57**, 1075 (1998).

30 J. E. Walsh, Nucl. Inst. Meth. Phys. Res. A **445**, 214 (2000).

31 P. M. v. d. Berg, J. Opt. Soc. Am. **63**, 689 (1973).

32 P. M. v. d. Berg, J. Opt. Soc. Am. **63**, 1588 (1973).

33 P. M. v. d. Berg and T. H. Tan, J. Opt. Soc. Am. **64**, 325 (1974).

34 D. V. Karlovets and A. P. Potylitsyn, Physical Review Special Topics - Accelerators and Beams **9**, 080701 (2006).

35 J. M. J. Madey, Journal of Applied Physics **42**, 1906 (1971).

36 A. Gover and Z. Livni, Optics Communications **26**, 375 (1978).

37 J. M. Wachtel, Journal of Applied Physics **50**, 49 (1979).

38 K. Mizuno and S. Ono, in *Infrared and Millimeter Waves*, edited by K. J. Button, 1979), Vol. 1, p. 213.

39 J. Urata, M. Goldstein, M. F. Kimmitt, et al., Phys. Rev. Lett. **80**, 516 (1998).

40 A. Bakhtyari, J. E. Walsh, and J. H. Brownell, Phys. Rev. E **65**, 066503 (2002).

41 J. R. Pierce, *Traveling-Wave Tubes* (D. Van Nostrand, New York, 1950).

# CHAPTER III

## THEORY OF THE SMITH-PURCELL FREE-ELECTRON LASER

In this chapter, we present the two-dimensional theoretical model for a SPFEL. We describe the electron beam as a plasma dielectric, and use Floquet's theorem to establish a form for TM waves traveling along the grating surface. The surface waves are evanescent above the grating, therefore they have a small longitudinal component of the electric field which allows them to interact with the electron beam. This interaction happens for waves whose phase velocity equals the electron beam velocity, as we see by analyzing the susceptibility of our plasma dielectric model. By deriving the dispersion relation for surface waves on the grating without the electron beam, we calculate the operating point, where this condition on the phase velocity is satisfied. We then introduce the electron beam as a perturbation to get an expansion of the dispersion relation around the operating point.

Obviously, the phase velocity has the same sign as the electron velocity at the synchronous point, but we see from the empty-grating dispersion relation that the group velocity can be positive or negative there, depending on the electron beam energy. Because the sign of the group velocity indicates the direction of energy flow in the wave, it has significant consequences for the device operation which were ignored by previous treatments of SPFELs. When the group velocity at the operating point is positive, the wave grows as a convective instability. A negative group velocity indicates the disturbance grows as an absolute instability.

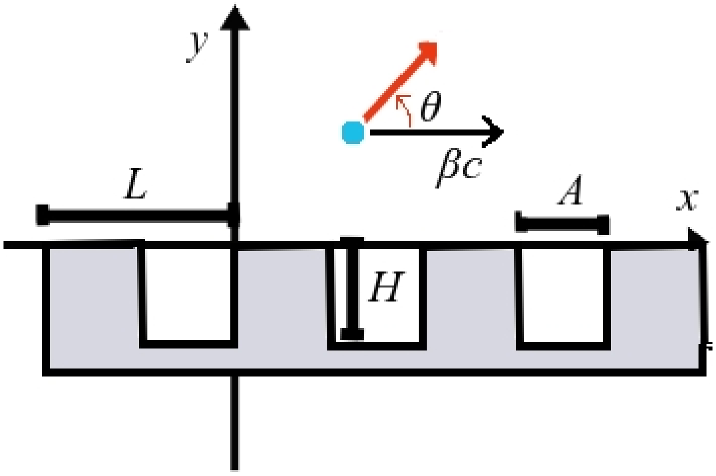

Figure 3-1. Coordinate system for Smith-Purcell radiation.

The figure of merit for a device operating on a convective instability is the gain length, as the wave only grows in space along the grating. For a device operating on an absolute instability, the wave grows everywhere along the grating, and we calculate the time growth rate. In both regimes, we analyze the effect of resistive losses in the grating, and we examine the effect of reflections at the grating ends on a wave with negative group velocity.

## Dispersion relation for evanescent surface waves

### *Model for the electron beam and fields above a grating*

The electron beam is represented as uniform plasma with a constant velocity, $v$, equal to the velocity of the electrons. The plasma fills the semi-infinite half-plane above the grating, defined by $y > 0$ in Figure 3-1. The motion of the plasma is reflected only in the frequency dependence of the dielectric susceptibility. The dielectric susceptibility of the plasma in its rest frame is

$$\chi'_e(\omega') = -\frac{\omega_p'^2}{\omega'^2} \tag{3-1}$$

where $\omega'_p$ is the plasma frequency, and the prime indicates the rest frame of the plasma. The plasma frequency is defined in terms of the rest frame electron density as

$$\omega'_p(n'_e) = \sqrt{\frac{n'_e q_e^2}{\varepsilon_0 m_e}} \tag{3-2}$$

where $n'_e$ is the electron number density in the plasma rest frame, $q_e$ is the charge on the electron, $\varepsilon_0$ is the vacuum permittivity, and $m_e$ is the mass of the electron.

In the lab frame, Lorentz contraction increases the electron number density and therefore the plasma frequency by a factor $\gamma = 1\big/\sqrt{1-\beta^2}$, where $\beta = \frac{v}{c}$ and $c$ is the speed of light. To transform the variable frequency on which the susceptibility depends, we consider waves that vary according to $e^{-ik^\alpha r_\alpha}$, where $k^\alpha$, the wave vector, is

$$k^\alpha = \left(\omega/c, k_x, k_y, k_z\right) \tag{3-3}$$

where $\omega$ is the angular frequency in a given frame and $k_x, k_y$ and $k_z$ are the wavenumbers in the three Cartesian directions. The covariant position vector is

$$r_\alpha = \left(ct, -x, -y, -z\right) \tag{3-4}$$

where $t$ is the time coordinate, and $x, y$, and $z$ represent the spatial coordinates in a given frame. The rule for transforming the frequency comes directly from the Lorentz transform of the wave vector. For a boost in the $x$-direction (which is the beam direction in Figure 3-1), that rule is

$$\omega' = \gamma(\omega - \beta c k_x) \tag{3-5}$$

The dielectric susceptibility of the moving plasma in the lab frame is therefore

$$\chi_e(\omega, k_x) = -\frac{\omega_p^{\ 2}}{\gamma^3 (\omega - \beta c k_x)^2} \tag{3-6}$$

The expression we have obtained for the dielectric susceptibility diverges for $\omega = \beta c k_x$. Physically, this means that the electrons interact with waves that have a phase velocity equal to the electron velocity. We now turn our attention to describing those waves. We consider transverse magnetic (TM) waves, which have a nonzero electric field component in the $x$- direction, allowing them to bunch the electrons. Floquet's theorem tells us that these waves can be expanded in the form

$$H_z = \sum_{p=-\infty}^{\infty} H_p e^{-\alpha_p y} e^{ipKx} e^{i(kx - \omega t)} \tag{3-7}$$

$$E_x = \sum_{p=-\infty}^{\infty} E_p e^{-\alpha_p y} e^{ipKx} e^{i(kx - \omega t)} \tag{3-8}$$

where $H_p$, $E_p$ and $\alpha_p$ are constants, $k = k_x$, $K = \frac{2\pi}{L}$ is the grating wave number, and $L$ is the grating period.

The wave equation for the electric field in the plasma rest frame is

$$\nabla^2 \mathbf{E} - \mu\varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} = \nabla^2 \mathbf{E} - \mu_0 \varepsilon_0 (1 + \chi_e) \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \tag{3-9}$$

where $\mu$ is the magnetic permeability of the medium (in the plasma rest frame, it is assumed to be equal to $\mu_0$, the magnetic permeability of free space), and $\varepsilon$ is the electric permittivity of the medium (related to the free space value by the dielectric susceptibility). For the form of the fields given in (3-7) and (3-8), we get

$$\alpha_p{}^2 = (pK+k)^2 - (1+\chi_e)\frac{\omega^2}{c^2} = (pK+k)^2 - \frac{\omega^2}{c^2} + \frac{\omega_p{}^2}{c^2} \tag{3-10}$$

where the $p$ on $\omega_p$ is a subscript that indicates the plasma frequency, not an index.

Computations show that $\alpha_p$ is always real and the sign of $\alpha_p$ is therefore chosen to satisfy the boundary condition that the fields vanish at infinity. For our convention stated in (3-7) and (3-8), this means that $|\alpha_p| > 0$. The TM wave is therefore evanescent in the $y$-direction, away from the grating. We will then refer to $\alpha_p{}^{-1}$ as the evanescent scale height for the $p^{th}$ space harmonic of the wave.

The Maxwell-Ampère law relates the electric and magnetic fields in the wave. From it, we find

$$\alpha_p H_p = i\varepsilon_0\omega(1+\chi_p)E_p \tag{3-11}$$

where the dielectric susceptibility of the $p^{th}$ component is

$$\chi_p(\omega,k) = -\frac{\omega_p{}^2}{\gamma^2\left[\omega - \beta c(k+pK)\right]^2} \tag{3-12}$$

When the wave is nearly synchronous with the electron beam, we ignore the susceptibility for all $p$ except $p=0$, where it is nearly divergent. In that case, we write (3-11) in the form

$$\alpha_p H_p = i\varepsilon_0\omega(1+\delta_{0p}\chi_0)E_p \tag{3-13}$$

where $\delta_{ij}$ is the Kronecker delta.

Inside the grooves of the grating, we expand the fields in the Fourier series

$$H_z = \sum_{n=0}^{\infty} \bar{H}_n \cos\left(\frac{n\pi x}{A}\right)\frac{\cosh\left[\kappa_n(y+H)\right]}{\sinh\left[\kappa_n H\right]}e^{-i\omega t} \tag{3-14}$$

$$E_x = \sum_{n=0}^{\infty} \bar{E}_n \cos\left(\frac{n\pi x}{A}\right) \frac{\sinh\left[\kappa_n \left(y+H\right)\right]}{\cosh\left[\kappa_n H\right]} e^{-i\omega t} \tag{3-15}$$

where $\bar{H}_n$, $\bar{E}_n$ and $\kappa_n$ are constants, $A$ is the width of one grating groove, and $H$ is the depth of the groove. These expressions are constructed to fulfill the boundary conditions at the conducting surface of the grating: that the electric field vanish at the bottom of the groove ( $y = -H$ ) and that the normal derivative of the magnetic field ( $\partial H_z / \partial x$ ) vanish on the sides of the groove ( $x = 0, A$ ). In the plasma-free grooves, the wave equation (3-9) becomes

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \tag{3-16}$$

and, inserting the forms of the fields, we get

$$\kappa_n{}^2 = \left(\frac{n\pi}{A}\right)^2 - \frac{\omega^2}{c^2} \tag{3-17}$$

The Maxwell-Ampère law gives

$$\bar{H}_n = -i\varepsilon_0 \frac{\omega}{\kappa_n} \tanh\left(\kappa_n H\right) \bar{E}_n \tag{3-18}$$

The electric field is continuous across the imaginary surface between the plasma region and the grooves region, and vanishes on the tops of the groove teeth. For $y = 0$, supressing the time dependence, this gives

$$\sum_{p=-\infty}^{\infty} E_p e^{i(pK+k)x} = \begin{cases} \sum_{n=0}^{\infty} \bar{E}_n \cos\left(\frac{n\pi x}{A}\right) \tanh\left[\kappa_n H\right] & \text{for } x = \left[0, A\right) \\ 0 & \text{for } x = \left[A, L\right] \end{cases} \tag{3-19}$$

If we multiply by $e^{-i(k+qK)x}$ and integrate along the $x$-axis from 0 to $L$, we get

$$E_q = \sum_{n=0}^{\infty} \bar{E}_n \tanh\left(\kappa_n H\right) \frac{K_{qn}}{L} \tag{3-20}$$

where

$$K_{qn} = iA \frac{\left(k+qK\right)A}{\left(k+qK\right)^2 A^2 - n^2\pi^2} \left[\left(-1\right)^n e^{-i\left(k+qK\right)A} - 1\right] \tag{3-21}$$

The magnetic field is also continuous between the two regions, though not restricted to vanish along the tops of the grating teeth. For $y = 0$, supressing the time dependence, this gives

$$\sum_{p=-\infty}^{\infty} H_p e^{i\left(k+pK\right)x)} = \sum_{n=0}^{\infty} \bar{H}_n \cos\left(\frac{n\pi x}{A}\right) \coth\left(\kappa_n H\right) \tag{3-22}$$

If we multiply by $\cos\left(m\pi x/A\right)$ and integrate along the $x$-axis from 0 to $A$, we get

$$\bar{H}_m \frac{1+\delta_{m0}}{2} \coth\left(\kappa_m H\right) = \sum_{p=-\infty}^{\infty} H_p \frac{K_{pm}^*}{A} \tag{3-23}$$

where the asterisk indicates the complex conjugate.

Substituting Equations (3-13) and (3-18) into (3-23), we get

$$-\frac{\bar{E}_m}{\kappa_m} \frac{1+\delta_{m0}}{2} = \sum_{p=-\infty}^{\infty} \frac{\left(1+\delta_{p0}\chi_0\right)}{\alpha_p} E_p \frac{K_{pm}^*}{A} \tag{3-24}$$

If we then substitute from Equation (3-20), we get

$$\bar{E}_m = \sum_{n=0}^{\infty} -\frac{\bar{E}_n}{LA} \frac{2\kappa_m \tanh\left(\kappa_n H\right)}{1+\delta_{m0}} \left[\frac{K_{0n}K_{0m}^*\chi_0}{\alpha_0} + \sum_{p=-\infty}^{\infty} \frac{K_{pn}K_{pm}^*}{\alpha_p}\right] \tag{3-25}$$

We can simplify the form of this matrix equation by making two definitions

$$R_{mn}=\frac{\tanh\left(\kappa_n H\right)}{1+\delta_{m0}}\sum_{p=-\infty}^{\infty}\frac{\kappa_m A}{\alpha_p L}\frac{4\left(k+pK\right)^2 A^2}{\left[\left(k+pK\right)^2 A^2-n^2\pi^2\right]\left[\left(k+pK\right)^2 A^2-m^2\pi^2\right]}\times \\ \begin{cases}(-1)^m 1-\cos\left[\left(k+pK\right)A\right]-1 & \text{for } m+n=\text{even}\\ i(-1)^m\sin\left[\left(k+pK\right)A\right] & \text{for } m+n=\text{odd}\end{cases} \tag{3-26}$$

and

$$S_{mn}=\frac{\tanh\left(\kappa_n H\right)}{1+\delta_{m0}}\frac{\kappa_m A}{\alpha_0 L}\frac{4k^2A^2}{\left(k^2A^2-m^2\pi^2\right)\left(k^2A^2-n^2\pi^2\right)}\times \\ \begin{cases}(-1)^m\cos\left(kA\right)-1 & \text{for } m+n=\text{even}\\ i(-1)^m\sin\left(kA\right) & \text{for } m+n=\text{odd}\end{cases} \tag{3-27}$$

With these definitions, the matrix Equation (3-25) becomes

$$\bar{E}_m=\sum_{n=0}^{\infty}\left(R_{mn}+\chi_0 S_{mn}\right)\bar{E}_n \tag{3-28}$$

For the solution of (3-28) to exist, the determinant of the coefficients in that equation must vanish, that is

$$\left|R_{mn}+\chi_0 S_{mn}-\delta_{mn}\right|=0 \tag{3-29}$$

Both $R_{mn}$ and $S_{mn}$ depend on $k$ and $\omega$ (through $\kappa_m$ and $\alpha_p$), and the Equation (3-29) is the dispersion relation for surface waves on the grating in the presence of the electron plasma. From the values of $k$ and $\omega$, the sets of coefficients $\bar{E}_m$, $\bar{H}_m$, $E_p$ and $H_p$ can be found.

*Dispersion relation without the electron beam*

Without the electron beam term, the dispersion relation is

$$\left|R_{mn}-\delta_{mn}\right|=0 \tag{3-30}$$

Table 3-1. Grating and beam parameters for the Dartmouth experiments.

| Parameters for the Dartmouth experiments | |
|---|---|
| grating period | 173 μm |
| grating length | 12.7 mm |
| groove width | 62 μm |
| groove depth | 100 μm |
| electron energy | 20-40 keV |
| electron-beam current | <1 mA |
| electron-beam diameter (FWHM) | >20 μm |

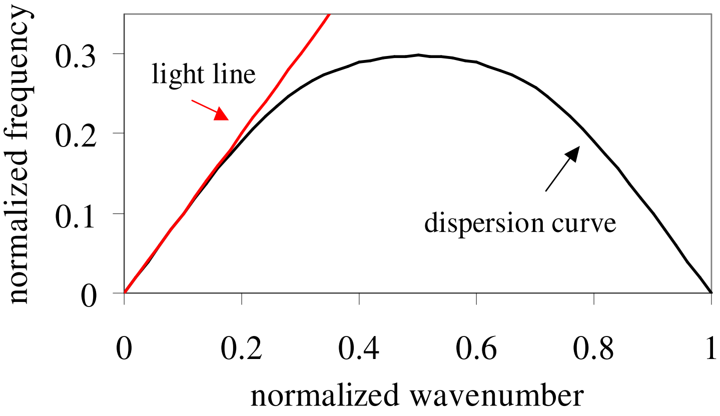

Figure 3-2. Dispersion relation for the Dartmouth grating.

Using Mathcad, we can evaluate the roots of this equation and plot the allowed values of $\omega$ and $k$. Using the parameters for the Dartmouth experiments, shown in Table 3-1[1], we calculate the empty-grating (no electron beam) dispersion curve shown in Figure 3-2. The curve is presented in terms of the normalized frequency

$$\omega_n = \frac{\omega L}{2\pi c} = \frac{\omega K}{c} \tag{3-31}$$

and normalized wavenumber

$$k_n = \frac{kL}{2\pi} = \frac{k}{K} \tag{3-32}$$

On this figure, straight lines through the origin connect points of equal phase velocity, where the phase velocity is

$$v_\phi = \frac{\omega}{k} = c\frac{\omega_n}{k_n} \tag{3-33}$$

and the derivative of the dispersion curve gives the group velocity

$$v_g = \frac{d\omega}{dk} = c\frac{d\omega_n}{dk_n} \tag{3-34}$$

The dispersion diagram shown in Figure 3-2 only contains points with positive phase velocities, but the diagram represents only the first Brillouin zone for the grating. The dispersion curve repeats with a period of one unit in normalized wavenumber in either direction along the abscissa. There exists, therefore, a Brillouin zone immediately to the left of the origin which contains points with only negative phase velocities, but the same range of group velocities displayed in the figure. This neighboring zone will be important in discussing reflections from the grating ends. An important point on the curve is the Bragg point, where the group velocity vanishes. This point represents the highest frequency (shortest corresponding free-space wavelength) for the set of grating solutions represented by the dispersion curve. The Bragg point occurs at a normalized wavenumber equal to one-half, and separates the set of points with positive group velocity on the left from points with negative group velocity on the right.

Points representing phase and group velocity equal to the speed of light lie along a straight line with slope equal to unity in the normalized coordinates. This light line is shown in red on Figure 3-2. The region above the light line in the Figure represents

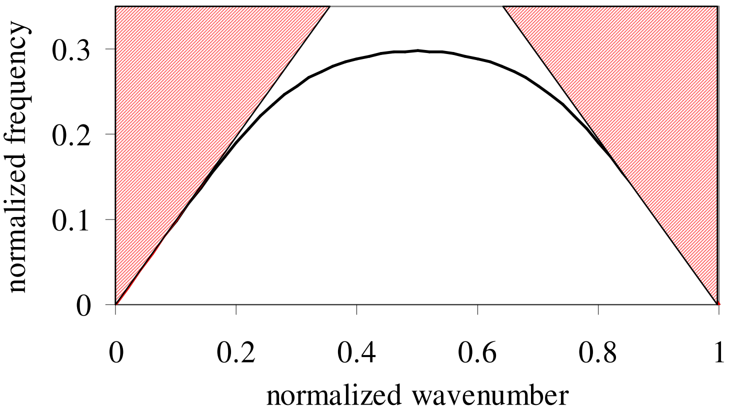

Figure 3-3. The forbidden region for evanescent waves is shown on the dispersion diagram as the red hatched region.

points for which the phase velocity exceeds that of light. For waves where $\alpha_p$ is real, we find from Equation (3-10) (in the absence of the electron beam) that

$$\omega_n < |k_n + p| \tag{3-35}$$

for all integers $p$. In the region where $k = (0,1)$, this becomes

$$\omega_n < k_n + 1 \quad \text{and} \quad \omega_n < k_n - 1 \tag{3-36}$$

Evanescent waves are therefore forbidden from the region above the light line and from the region above the negative slope light line originating from the origin of the next Brillouin zone, as shown in Figure 3-3. The presence of forbidden regions in the dispersion diagram is a characteristic of open structures[2].

We can find other forbidden regions, and other solutions, to the dispersion relation by examining its analytic form. From Equations (3-17) and (3-26), we see that

$$R_{00} \propto \omega_n KH \tan(\omega_n KH) \tag{3-37}$$

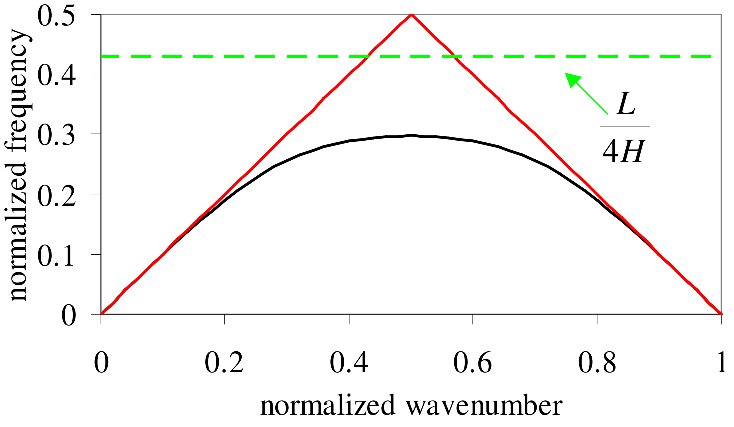

Figure 3-4. The allowed region for the dispersion curve is restricted by the top of the first passband.

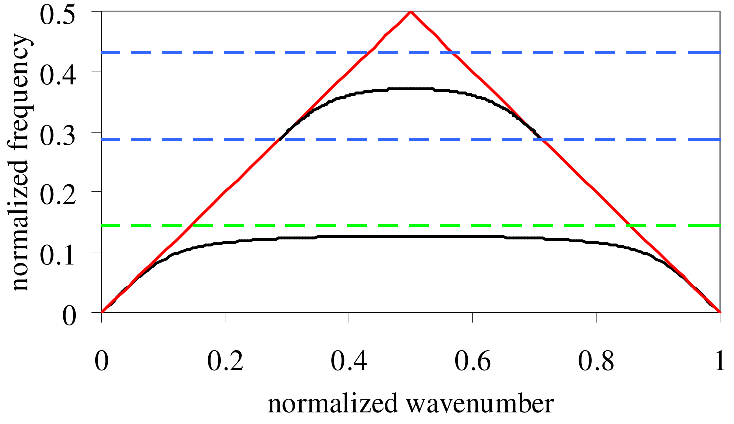

Figure 3-5. For a grating with deep grooves, more solutions to the dispersion relation appear as shown in black. The dotted green line represents the top of the first passband and the dotted blue lines the top and bottom of the second passband.

Solutions of the approximate dispersion relation $R_{00} = 1$ are therfore only possible for positive frequency when $\tan\left(\omega_n KH\right) > 0$. This is true when

$$2\pi n < \omega_n KH < \frac{(4n+1)\pi}{2} \tag{3-38}$$

for some integer $n$. The result is a series of passbands and stopbands in frequency. From this equation, we see that the first passband has a cutoff at

$$\omega_n = \frac{\pi}{2KH} = \frac{L}{4H} \tag{3-39}$$

For the Dartmouth parameters, $\frac{L}{4H} = 0.43$, placing only a small restriction on the allowed dispersion curve as shown in Figure 3-4. If, however, we triple the depth of the grating to 300 μm, the dispersion diagram appears as shown in Figure 3-5. A second band of solutions appears, corresponding to a frequency passband satisfying Equation (3-39) for $n = 2$. There are no solutions to the dispersion relation for frequencies that do not satisfy Equation (3-39). The results on SPFEL performance presented in this dissertation can be applied to any solutions of the dispersion relation, and in fact, the existence of higher passbands has been confirmed in simulations[3]. From this point forward, however, we discuss the solutions in the first passband.

### *Convergence of the numerical calculations*

The accuracy of the numerical calculations of the grating dispersion relation depends on the number of terms taken in the series expansions for the fields in the grooves and in the plasma above the grating. If we use the Dartmouth parameters with a 40 keV electron beam and take four terms in the expansion for the fields in the grooves

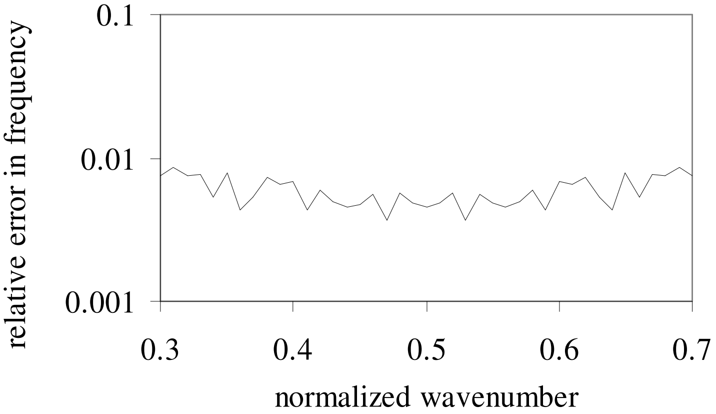

Figure 3-6. Relative error in frequency introduced by using one term in the field expansion inside the grooves instead of four terms.

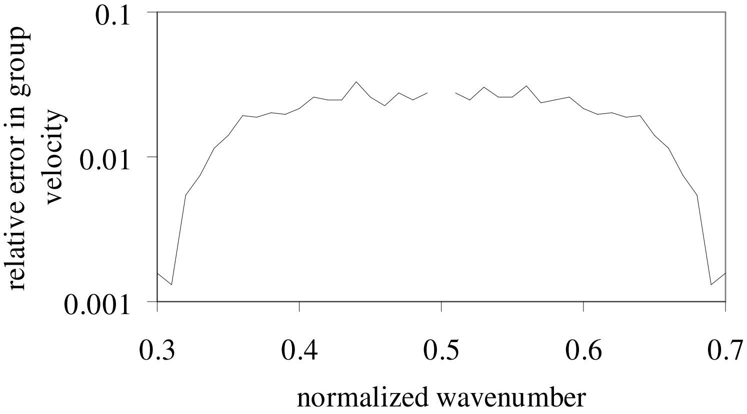

Figure 3-7. Relative error in group velocity introduced by using one term in the field expansion inside the grooves instead of four terms.

(and 23 terms for the fields above the grating), we find that the ratio between the electric fields in the different terms of the expansion in the grooves is

$$|E_0|^2 : |E_1|^2 : |E_2|^2 : |E_3|^2 = 1 : 0.193 : 0.211 : 0.039 \qquad (3\text{-}40)$$

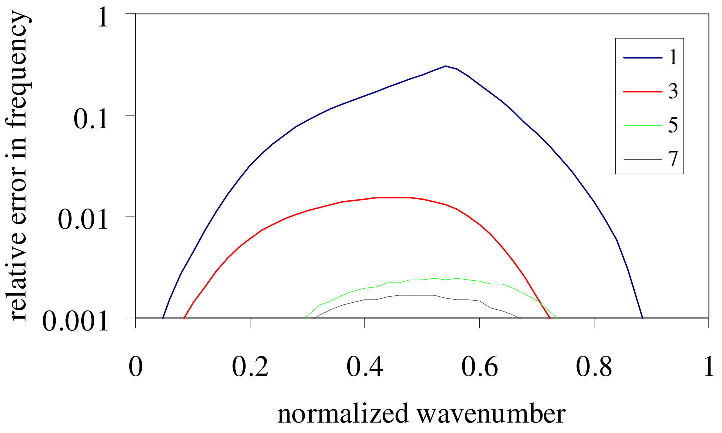

Figure 3-8. Relative error in frequency (compared to a 23-term expansion) with one, three, five, and seven terms in the expansion of the fields above the grating and one term in the grooves.

The dominance of the zeroth-order term suggests it is possible to do calculations with only one term in the grooves, that is, to approximate the dispersion relation by $R_{00} = 1$. Comparing this approximation to the calculation with four terms in the grooves, we find that the error in the frequency is <3% and the relative error in the group velocity is <1%. The relative errors in these two quantities as a function of wavenumber are shown in Figure 3-6 and Figure 3-7.

Using one term in the field expansions in the grooves, we then analyze the required number of terms in the series expansion for the fields above the grating. The relative error in frequency (compared to a 23-term expansion) is significant, >10%, but can be reduced by adding just a few more terms, as shown in Figure 3-8. The relative error in the group velocity is greater, but a five-term expansion is enough to reduce the error to less than 1% for all wavenumbers, as shown in Figure 3-9. The computations to be presented on performance of the SPFEL repeatedly make use of the frequency and

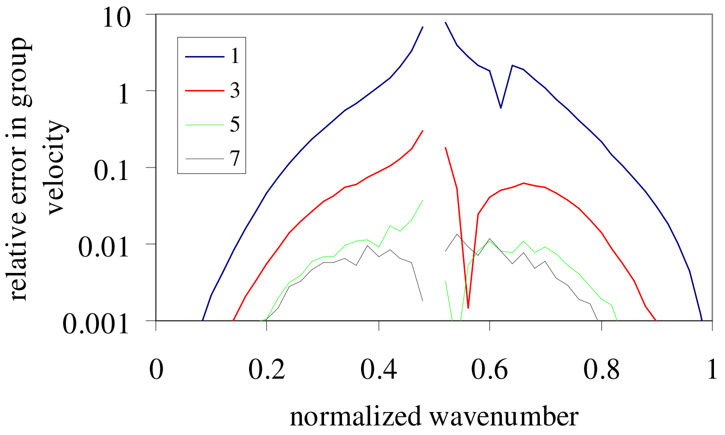

Figure 3-9. Relative error in the group velocity (compared to a 23-term expansion) with one, three, five, and seven terms in the expansion of the fields above the grating and one term in the grooves.

group velocity calculations. To control the computing time required to produce results, we limit the field expansion to one term in the grooves and five terms above the grating.

*Approximate dispersion relation in the presence of the electron beam*

We take advantage of the simplification discussed in the previous section and look at the dispersion relation in the presence of the electron beam with only one term in the field expansion in the grooves

$$R_{00} - 1 + \chi_0 S_{00} = 0 \tag{3-41}$$

When the effect of the electron beam is small, we expand $R_{00}$ around the solution to the approximate empty-grating dispersion relation given by

$$R_{00} = 1 \tag{3-42}$$

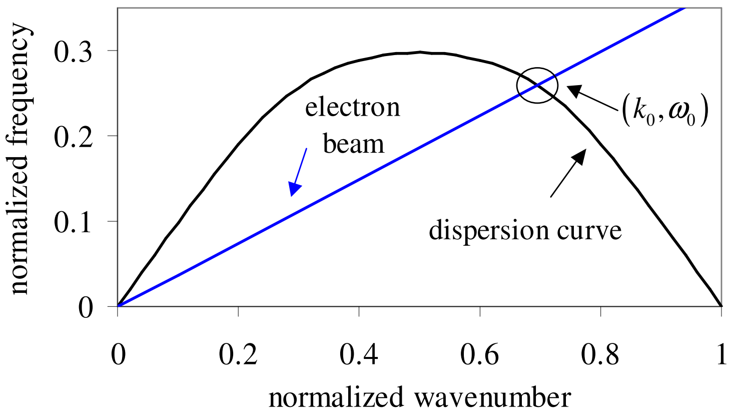

Figure 3-10. Intersection between the line of constant phase velocity representing the electron beam and the dispersion curve selects the expansion point for calculating the gain in the interaction.

We choose the expansion point $(k_0,\omega_0)$ by finding the solution on the empty-grating dispersion curve corresponding to a phase velocity that equals the electron velocity. Graphically, this solution for the operating point is shown in Figure 3-10.

The first-order expansion of the approximate dispersion relation around the operating point is

$$R_{00}(k,\omega) \approx 1 + R_\omega \delta\omega + R_k \delta k \tag{3-43}$$

where $\delta\omega = (\omega - \omega_0)$, $\delta k = (k - k_0)$, and

$$R_\omega = \left. \frac{\partial R_{00}}{\partial \omega} \right|_{k_0,\omega_0} \tag{3-44}$$

$$R_k = \left. \frac{\partial R_{00}}{\partial k} \right|_{k_0,\omega_0} \tag{3-45}$$

Differentiating Equation (3-42) with respect to $k$ we get

$$R_\omega \frac{d\omega}{dk} + R_k = \beta_g c R_\omega + R_k = 0 \tag{3-46}$$

where $\beta_g$ is the group velocity of the wave. We use this result to eliminate $R_k$ from Equation (3-43), yielding

$$R_{00}\left(k,\omega\right) \approx 1 + R_\omega\left(\delta\omega - \beta_g c\delta k\right) \tag{3-47}$$

Substituting into Equation (3-41), we produce an approximate dispersion relation including the electron beam, given by

$$R_\omega\left(\delta\omega - \beta_g c\delta k\right) + \chi_0 S = 0 \tag{3-48}$$

where $S = S_{00}\left(k_0, \omega_0\right)$.

As shown in Equation (3-12), the susceptibility of the electron beam plasma diverges at the point we have now defined at the operating point, where $\omega_0 = \beta c k_0$. In terms of the variables $\delta k$ and $\delta\omega$, the susceptibility is

$$\chi_0 = -\frac{{\omega_p}^2}{\gamma^3\left(\delta\omega - \beta c\delta k\right)^2} \tag{3-49}$$

Substituting the expression for the susceptibility into our dispersion relation gives

$$\left(\delta\omega - \beta c\delta k\right)^2\left(\delta\omega - \beta_g c\delta k\right) = \Delta \tag{3-50}$$

where

$$\Delta = \frac{{\omega_p}^2 S}{\gamma^3 R_\omega} \tag{3-51}$$

is real and positive (shown by calculations).

*Effect of losses in the grating*

So far we have ignored the attenuation of TM waves that results from the finite conductivity of the grating. This resistive loss can be important, especially when the group velocity of the wave is low. Losses introduce both a real and an imaginary shift in the frequency, and therefore should be included in the dispersion relation. To analyze the losses, we follow the succesive approximation approach taken by Jackson [4] for computing the fields near the metal surface.

First, Jackson assumes that the only fields outside the conductor are the normal electric field and the tangential magnetic field given by the solution to the problem for the perfect conductor. The boundary conditions on the fields are

$$\mathbf{n}\cdot\left(\mathbf{B}-\mathbf{B_c}\right)=0 \tag{3-52}$$

$$\mathbf{n}\cdot\left(\mathbf{E}-\mathbf{E_c}\right)=0 \tag{3-53}$$

where $\mathbf{n}$ is the unit vector normal to the conductor surface pointing away from the conductor and the subscript $\mathbf{c}$ indicates fields inside the conductor. For a perfect conductor, $\mathbf{E_c}=\mathbf{B_c}=0$ everywhere, but this requires that the surface charges create a current given by

$$\mathbf{K}=\mathbf{n}\times\mathbf{B} \tag{3-54}$$

and such a current is not possible for finite conductivity. Ignoring displacement current, the Maxwell equations give

$$\mathbf{E_c}\cong\frac{1}{\sigma}\nabla\times\mathbf{H_c} \tag{3-55}$$

$$\mathbf{H_c}=-\frac{i}{\mu_0\omega}\nabla\times\mathbf{E_c} \tag{3-56}$$

Considering only the normal direction, the gradient operator is

$$\nabla = -\mathbf{n}\frac{\partial}{\partial \xi} \tag{3-57}$$

Equations (3-55) and (3-56) are then

$$\mathbf{E_c} \cong -\frac{1}{\sigma}\mathbf{n}\times\frac{\partial \mathbf{H_c}}{\partial \xi} \tag{3-58}$$

$$\mathbf{H_c} = \frac{i}{\mu_0 \omega}\mathbf{n}\times\frac{\partial \mathbf{E_c}}{\partial \xi} \tag{3-59}$$

Combining the two equations yields

$$\left(\frac{\partial^2}{\partial \xi^2} + \mu_0 \omega \sigma i\right)\left(\mathbf{n}\times\mathbf{H_c}\right) \cong 0 \tag{3-60}$$

$$\mathbf{n}\cdot\mathbf{H_c} \cong 0 \tag{3-61}$$

where $\sigma$ is the conductivity of the conductor. The solution for the magnetic field inside the conductor in terms of the field outside is

$$\mathbf{H_c} = \mathbf{H}e^{(-1+i)\frac{\xi}{\delta}} \tag{3-62}$$

where the skin depth is

$$\delta = \sqrt{\frac{2}{\mu_0 \omega \sigma}} \tag{3-63}$$

The electric field inside the conductor is

$$\mathbf{E_c} \cong \sqrt{\frac{\mu_0 \omega}{2\sigma}}\left(1-i\right)\left(\mathbf{n}\times\mathbf{H}\right)e^{(-1+i)\frac{\xi}{\delta}} \tag{3-64}$$

The boundary condition expressed by Equation (3-53) then requires a small tangential electric field component outside the conductor given by

$$\mathbf{E}_{\perp} \cong \sqrt{\frac{\mu_0 \omega}{2\sigma}}\left(1-i\right)\left(\mathbf{n}\times\mathbf{H}\right) \tag{3-65}$$

With tangential components to both the electric and magnetic fields at the surface of the conductor, calculation of the Poynting vector results in an energy flow into the surface. Averaging over one cycle of the wave, we see that the loss into the grating surface is

$$Q = \langle S \rangle = -\langle \mathbf{n} \cdot (\mathbf{E} \times \mathbf{H}) \rangle = -\frac{1}{2} \mathrm{Re}\left[ \mathbf{n} \cdot (\mathbf{E} \times \mathbf{H}) \right] \tag{3-66}$$

For the fields in Equations (3-64) and (3-62), this gives

$$Q = -\frac{1}{2} \mathrm{Re} \sqrt{\frac{\mu_0 \omega}{2\sigma}} |\mathbf{H}|^2 \tag{3-67}$$

The conductivity of the aluminum metal in the grating is given by

$$\sigma = \frac{\sigma_0}{1 - i\omega\tau} = \sigma_0 \frac{1 - i\omega\tau}{1 + \omega^2\tau^2} \tag{3-68}$$

where $\sigma_0$ is the DC conductivity and $\tau$ is the mean collision time for conduction electrons[5 (p. 345)]. For aluminum, $\sigma_0 \approx 4 \times 10^7 \ \Omega^{-1}\mathrm{m}^{-1}$ and $\tau \approx 1 \times 10^{-14}$ s. For low THz frequencies, $\omega^2\tau^2 \ll 1$, and we can use the DC value for the conductivity, so the grating loss per unit area is

$$Q = -\frac{1}{2} \sqrt{\frac{\mu_0 \omega}{2\sigma_0}} H_z^{\,2} \tag{3-69}$$

This expression can be calculated for any point on the grating surface by using the empty-grating solution for the magnetic field, valid in this case for low losses. Details of this calculation are provided in a separate section.

As mentioned above, the resistive losses introduce a real and an imaginary shift in the frequency. To compute this shift, we consider a section of the grating as if it were a cavity with perfectly reflecting end mirrors. By analogy to perturbation theory in waveguides, it can be shown that the losses in the cavity walls induce equal shifts in the real and imaginary parts of the frequency, given by [4 (p. 374)]

$$\delta\omega = -\frac{\omega_0}{2Q_c}(1+i) \tag{3-70}$$

where $Q_c$, the Q-factor for the cavity, is

$$Q_c = \omega_0 \frac{\langle U \rangle}{\langle Q \rangle} \tag{3-71}$$

$U$ is the power density in the wave, $Q$ is the power loss into the grating (calculated in Equation (3-67) for a reasonable conductor), and the angled brackets here represent averages over one cycle of the wave and over one period of the grating. As noted by Jackson, the equality of the real and imaginary frequency shifts follows from the boundary condition given by Equation (3-65).

With the inclusion of the frequency shift due to resistive losses, the expansion of the approximate dispersion relation around the operating point is

$$(\delta\omega - \beta c \delta k)^2 \left[ \delta\omega - \beta_g c \delta k + \frac{\omega_0}{2Q_c}(1+i) \right] = \Delta \tag{3-72}$$

## SPFEL operation: amplifier and oscillator

### *Operation regimes and the group velocity*

To calculate the growth of the evanescent wave for the dispersion relation given in Equation (3-50), we give primary consideration to the sign of the group velocity $\beta_g$. The group velocity represents the energy flow in the evanescent wave. A positive sign for the group velocity indicates energy flow in the same direction as the electron motion.

A negative sign for $\beta_g$ indicates energy flowing in a direction opposite to the electron velocity. As the growth of the evanescent wave becomes nonlinear, the behavior for positive and negative group velocity is very different. Energy propagating with the electron beam leads to a convective instability, meaning we see growth of the evanescent wave with time only if we move along with the wave. Energy propagating against the electron beam flow leads to an absolute instability, meaning the evanescent wave grows with time at all spatial points [6].

In terms of device operation, positive group velocity (convective instability) leads to single-pass growth limited by the length of the grating. External feedback of some kind would be required to achieve oscillation. However, when the group velocity is negative (absolute instability), feedback is not necessary. The wave grows toward the upstream end of the grating, but bunching in the electron beam is carried along with the electron motion. The wave starts up from an increasingly bunched electron beam, and the device can go into oscillation without any external feedback. We refer to the positive group-velocity case as the amplifier regime, as the output evanescent wave is simply a function of the input, whether that input be noise or some kind of seeding. The negative group-velocity will be referred to as the oscillator regime, as the evanescent wave grows to saturation when the electron beam current exceeds some threshold value. The figure of merit for the amplifier is the gain length, and the appropriate quantities to describe the oscillator are the start current and the time rate of growth.

*Gain length in the amplifier regime*

In general the wavenumber shift $\delta k$ and the frequency shift $\delta\omega$ are complex. For the case of positive group velocity, however, we can consider a steady-state amplifier where $\delta\omega$ is real (an imaginary component to the frequency gives a time rate of change to the evanescent wave). We find the largest gain in this steady-state case by differentiating Equation (3-72) with respect to the frequency shift. If we set $\delta\omega = 0$, the derivative

$$\frac{d\delta k}{d\delta\omega} = \frac{\beta + 2\beta_g}{3\beta\beta_g c} \tag{3-73}$$

is real, so the derivative of the imaginary part of $\delta k$ is zero. The imaginary part of the wavenumber then has an extremum for $\delta\omega = 0$. By the convention established in Equations (3-7) and (3-8), the negative of the imaginary part of the wavenumber represents the gain length. For this steady-state case, Equation (3-72) becomes

$$-\left(\beta c\delta k\right)^2 \left[\beta_g c\delta k - \frac{\omega_0}{2Q_c}\left(1+i\right)\right] = \Delta \tag{3-74}$$

and, ignoring losses for the moment $\left(Q_c \to \infty\right)$, we have

$$\delta k^3 = -\frac{\Delta}{\beta^2 \beta_g c^3} \tag{3-75}$$

Of the three solutions for $\delta k$, the one with the most negative $\operatorname{Im}\delta k$ has the highest gain. The amplitude gain parameter (inverse gain length) for that solution is

$$\mu_\infty = -\operatorname{Im}\delta k = \frac{\sqrt{3}}{2}\left|\frac{\Delta}{\beta^2 \beta_g c^3}\right|^{1/3} \tag{3-76}$$

where the subscript infinity represents the no-loss case. The parameter $\Delta$ varies slowly with the group velocity, so the gain varies approximately as $\beta_g^{-1/3}$, particularly near the Bragg point. Such a dependence agrees with that shown for traveling-wave-tubes by Pierce[7]. The gain is therefore greatest when the group velocity is small.

In the absence of the electron beam $(\Delta \to 0)$, Equation (3-74) becomes simply

$$\delta k = \frac{\omega_0}{2\beta_g c Q_c}(1+i) \tag{3-77}$$

In this case, the imaginary part of the wavenumber is positive, which describes an attenuation parameter

$$\nu_0 = \operatorname{Im}\delta k = \left|\frac{\omega_0}{2\beta_g c Q_c}\right| = \left|\frac{\langle Q\rangle}{2\beta_g c\langle U\rangle}\right| \tag{3-78}$$

where the subscript zero indicates the no-beam case. Both $\langle Q\rangle$ and $\langle U\rangle$ are slowly varying functions of the wavenumber, therefore the attenuation parameter varies as $\beta_g^{-1}$ near the Bragg point. The ratio of ideal gain to no-beam loss, then, varies as $\beta_g^{2/3}$. Though the ideal gain parameter is largest where the group velocity is small, the attenuation parameter grows faster with decreasing group velocity.

When we have both gain and loss, we can cast the dispersion relation given by Equation (3-74) in the dimensionless form

$$(\Delta k)^2\left[\Delta k - (1+i)\right] + J^3 = 0 \tag{3-79}$$

where

$$\Delta k = \frac{2\beta_g c Q_c}{\omega_0}\delta k = \frac{\delta k}{\nu_0} \tag{3-80}$$

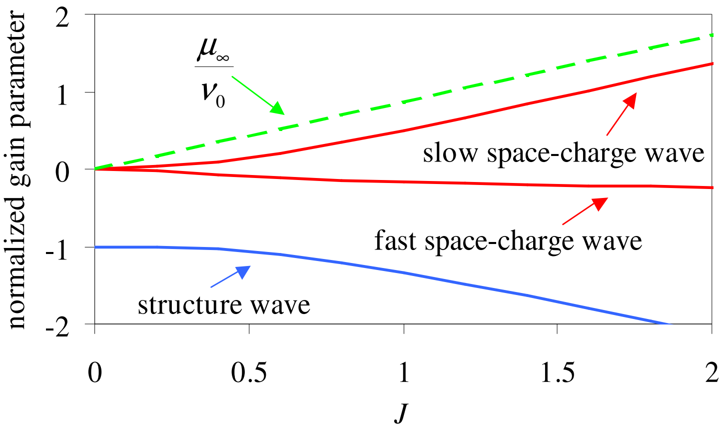

Figure 3-11. Negative imaginary part, or normalized gain parameter, for the three solutions to Equation (3-79) (red and blue lines), and their dependence on $J$. For reference, the ratio of ideal gain to no-beam loss is plotted (green dotted line).

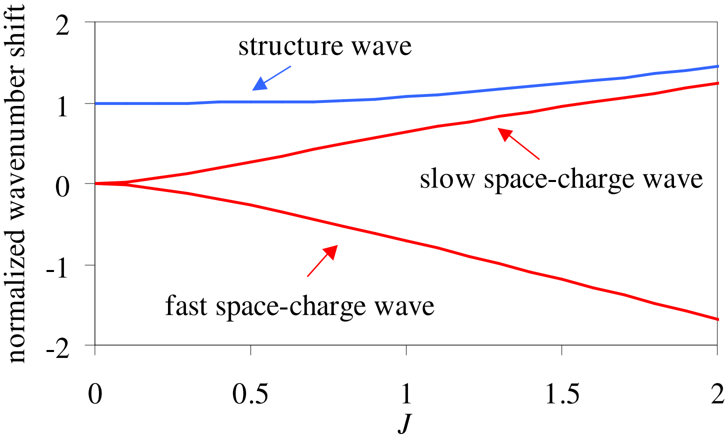

Figure 3-12. Real part, or normalized wavenumber shift, for the three solutions to Equation (3-79).

$$J = \frac{2Q_c}{\omega_0}\left(\frac{\beta_g{}^2\Delta}{\beta^2}\right)^{1/3} = \frac{2\mu_\infty}{\sqrt{3}\nu_0} \tag{3-81}$$

This equation allows us to solve for the gain parameter away from the limiting cases given by Equations (3-75) and (3-78). For a given value of $Q_c$, the three solutions vary with $J$ as shown in Figure 3-11. As $\Delta \to 0$, we get the two trivial solutions $(\delta k = 0)$ and the solution given by Equation (3-77). As we add the electron beam, the two trivial solutions separate. We identify these two solutions as density fluctuations or space-charge waves in the electron beam. As shown in Figure 3-12, one has a positive real part and the other a negative real part. All three solutions share the same frequency, so a positive shift in $\Delta k$ corresponds to a reduction in the phase velocity, and vice versa. Based on the phase velocity difference, we call one of the solutions the slow space-charge wave and the other the fast space-charge wave. The third solution we call the structure wave, as it represents a wave traveling along the grating structure even when the beam is absent.

The imaginary part of $\Delta k$ represents the gain or attenuation parameter for the wave. Again, as per the convention in Equations (3-7) and (3-8), a negative imaginary part represents gain and a positive imaginary part represents attenuation. The negative imaginary part for the three solutions is shown in Figure 3-11. As shown, the structure wave is always attenuated, increasingly so as the electron beam current grows. The fast space-charge wave is always slightly attenuated, and the slow space-charge wave displays increasing gain as the current is increased. The gain on the slow space-charge wave approaches the ratio of ideal gain to no-beam loss as the current is increased.

Interestingly, the gain is always positive on the slow space-charge wave, even when the ideal gain is less than the no-beam attenuation coefficient.

*Growth rate in the oscillator regime*

The analysis of SPFEL operation is very different when the group velocity is negative. For points on the right hand side of the dispersion curve $(k > 1/2)$, the device operates on an absolute instability, and will oscillate above a threshold current called the start current. For the oscillator case, there is a nonzero shift in both frequency and wavenumber. The three solutions to the dispersion relation have the same frequency, meaning they grow together, but different wavenumbers, allowing them to interfere to fulfill the boundary conditions at the grating ends.

We consider the evanescent electric field above the grating as a sum of the contributions from the three solutions to the dispersion relation

$$E_x = \sum_{j=1}^{3} E_j = \sum_{j=1}^{3} \sum_{p=-\infty}^{\infty} E_p{}^{(j)} e^{-\alpha_p{}^{(j)} y} e^{ipKx} e^{i(k_0 x - \omega_0 t)} e^{i(\delta k_j x - \delta\omega_j t)} \tag{3-82}$$

where the index $j = 1, 2, 3$ represents the three different solutions. We then make the approximation that the decay constants $\alpha_p{}^{(j)}$ for each of the three solutions are equal to the empty-grating values. Also, we assume that the ratios between the different spatial harmonic coefficients $E_p{}^{(j)}$ are the same as for the empty grating. Then, apart from an overall constant representing the amplitude of the mode, the field can be written

$$E_x = \sum_{j=1}^{3} E_j = E_0 \sum_{j=1}^{3} A_j e^{i(\delta k_j x - \delta\omega_j t)} \tag{3-83}$$

where

$$E_0 = \sum_{p=-\infty}^{\infty} E_p e^{-\alpha_p y} e^{ipKx} e^{i(k_0 x - \omega_0 t)} \tag{3-84}$$

is the empty-grating form of the longitudinal field, and the coefficients $A_j$ are constants. The coeffficients $A_j$ reflect the weighted contribution of the three solutions that will satisfy the boundary conditions.

In the amplifier case, we forced a steady-state solution and looked for the frequency shift with the greatest gain parameter, finding the maximum when $\delta\omega = 0$. For the oscillator, we only require that the three solutions have the same frequency shift, that is, we require $\delta\omega_j = \delta\omega$. The waves are allowed to grow with time, but the time rate of growth of the three waves is locked together so that they continue to satisfy the boundary conditions as the mode grows.

At the upstream end of the grating, the electron beam arrives without spatial fluctuations in electron density or velocity. Zero spatial variation in electron density translates to zero polarization for the dielectric plasma used to represent the electron beam. No spatial fluctuations in velocity equates to a vanishing convective derivative for the polarization. Since the plasma is a linear dielectric, the polarization is the sum of the polarizations due to the three components. The three components have the same frequency in the lab frame, but different wavenumbers. Therefore, they have different frequencies and different susceptibilities in the plasma rest frame. The total polarization is [5 (p. 292)]

$$P_x = -\frac{\varepsilon_0 \omega_p{}^2 E_0}{\gamma^3} \sum_{j=1}^{3} \frac{A_j e^{i(\delta k_j x - \delta\omega t)}}{\left(\delta\omega - \beta c \delta k_j\right)^2} \tag{3-85}$$

where we have used the result in Equation (3-49) for the dielectric susceptibility and the expansion from Equation (3-83) for the longitudinal electric field. The corresponding boundary condition of vanishing polarization at $x = 0$ is

$$\sum_{j=1}^{3} \frac{A_j}{\left(\delta\omega - \beta c \delta k_j\right)^2} = 0 \tag{3-86}$$

The convective derivative of the polarization is

$$\frac{dP_x}{dt} = \left(\frac{\partial}{\partial t} + \beta c \frac{\partial}{\partial x}\right) P_x = \frac{i\varepsilon_0 \omega_p{}^2 E_0}{\gamma^3} \sum_{j=1}^{3} \frac{A_j e^{i\left(\delta k_j x - \delta\omega t\right)}}{\delta\omega - \beta c \delta k_j} \tag{3-87}$$

and the corresponding boundary condition (also at $x = 0$) is

$$\sum_{j=1}^{3} \frac{A_j}{\delta\omega - \beta c \delta k_j} = 0 \tag{3-88}$$

At the downstream end, there are no boundary conditions on the electron beam, but we require the fields to be zero because the downstream end is the origin for the backward moving wave. That boundary condition at $x = Z$ is

$$\sum_{j=1}^{3} A_j e^{i\delta k_j Z} = 0 \tag{3-89}$$

where $Z$ is the length of the grating.

We simplify the form of the boundary conditions by defining

$$\delta_j = \left|\frac{\beta^2 \beta_g c^3}{\Delta}\right|^{1/3} \left(\frac{\delta\omega}{\beta c} - \delta k_j\right) \tag{3-90}$$

In these normalized variables, the boundary conditions are

$$\sum_{j=1}^{3} \frac{A_j}{\delta_j{}^2} = 0 \tag{3-91}$$

$$\sum_{j=1}^{3} \frac{A_j}{\delta_j} = 0 \tag{3-92}$$

$$\sum_{j=1}^{3} A_j e^{-i\xi\delta_j} = 0 \tag{3-93}$$

where we make the further definition

$$\xi = \left| \frac{\Delta}{\beta^2 \beta_g c^3} \right|^{1/3} Z = \frac{2}{\sqrt{3}} \mu_\infty Z \tag{3-94}$$

We solve the boundary conditions subject to the constraint imposed by the dispersion relation. The dispersion relation Equation (3-72) is

$$\delta^2 \left(\delta - \kappa\right) + 1 = 0 \tag{3-95}$$

where $\delta$ refers to any one of the three $\delta_j$, and

$$\kappa = \left| \frac{\beta^2 \beta_g}{\Delta} \right|^{1/3} \left[ \left( \frac{1}{\beta} - \frac{1}{\beta_g} \right) \delta\omega - \frac{\omega_0}{2\beta_g Q_c} \left(1 + i\right) \right] = \frac{\sqrt{3}}{2\mu_\infty} \left[ \frac{\beta_g - \beta}{\beta \beta_g c} \delta\omega + \nu_0 \left(1 + i\right) \right] \tag{3-96}$$

where we recall that the ideal gain and no-beam attenuation were defined using absolute value so that the same values obtain when the group velocity $\beta_g$ is negative.

Solving the boundary conditions simultaneously is equivalent to finding the roots in the determinant of the matrix of coefficients

$$\begin{vmatrix} \dfrac{1}{\delta_1^2} & \dfrac{1}{\delta_2^2} & \dfrac{1}{\delta_3^2} \\ \dfrac{1}{\delta_1} & \dfrac{1}{\delta_2} & \dfrac{1}{\delta_3} \\ e^{-i\xi\delta_1} & e^{-i\xi\delta_2} & e^{-i\xi\delta_3} \end{vmatrix} = 0 \tag{3-97}$$

The solution of Equations (3-95) and (3-97) appears in the literature for backward-wave oscillators (BWOs) and comes in the form of a solution for $\kappa$ in terms of $\xi$ [8]. We get

Table 3-2. Grating and beam parameters for the MAGIC simulation by Donohue and Gardelle[9].

| Parameters for the simulation by Donohue and Gardelle | |
|---|---|
| grating period | 20 mm |
| grating length | 700 mm |
| groove width | 10 mm |
| groove depth | 10 mm |
| electron energy | 100 keV |
| electron-beam current density | 25-500 A/m |
| electron-beam thickness | 5 mm |
| center of beam height above grating | 4.5 mm |

the same solution, derived without considering loss, and apply to the SPFEL, incorporating the loss in our interpretation of $\kappa$. The time growth rate of the mode is

$$\operatorname{Im}\delta\omega = \frac{\beta\beta_g c}{\beta_g - \beta}\left(\frac{2\mu_\infty}{\sqrt{3}}\operatorname{Im}\kappa - \nu_0\right) \tag{3-98}$$

The starting condition that this growth rate be positive is

$$\operatorname{Im}\kappa(\xi) > \frac{\sqrt{3}\nu_0}{2\mu_\infty} \tag{3-99}$$

We compare this theory to published simulations of evanescent wave growth in SPFELs[9]. The parameters for the simulation using the particle-in-cell (PIC) code MAGIC are given in

Table 3-2. For the simulation parameters, the second harmonic of the evanescent wave frequency appears in the Smith-Purcell band with an observation angle of 78°. By examining the magnetic field at a position corresponding to that angle in the simulation, Donohue and Gardelle find the growth rate at different values for the

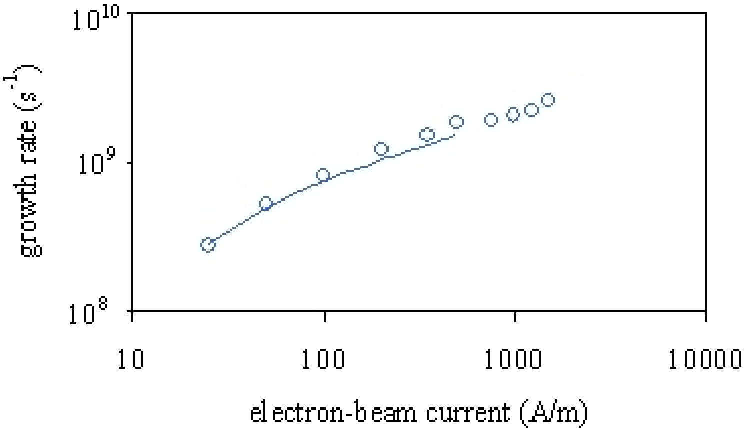

Figure 3-13. Growth rate of coherent Smith-Purcell radiation on the 2nd harmonic of the evanescent wave as computed by PIC simulations(blue circles) and the corresponding growth rate of the evanescent wave itself computed by our theory (blue line).

electron-beam current. Figure 3-13 shows the growth rate in the simulation and the growth rate calculated from Equation (3-98) (with a zero attenuation coefficient, as the simulation does not include losses). The agreement is remarkable, and there are no adjustable parameters used to fit the curve.

The theory disagrees, however, with the results from the Dartmouth experiments. For a 24-μm diameter electron beam and grating parameters given in Table 3-1, Dartmouth saw a threshold current for nonlinear growth at 800 μA. Our two-dimensional theory predicts a start current of 2 mA. The agreement is actually even worse when we attempt to include three-dimensional effects. The width of the optical mode is on the order of [10]

$$w \approx \sqrt{\frac{Z_g}{k_0}} \tag{3-100}$$

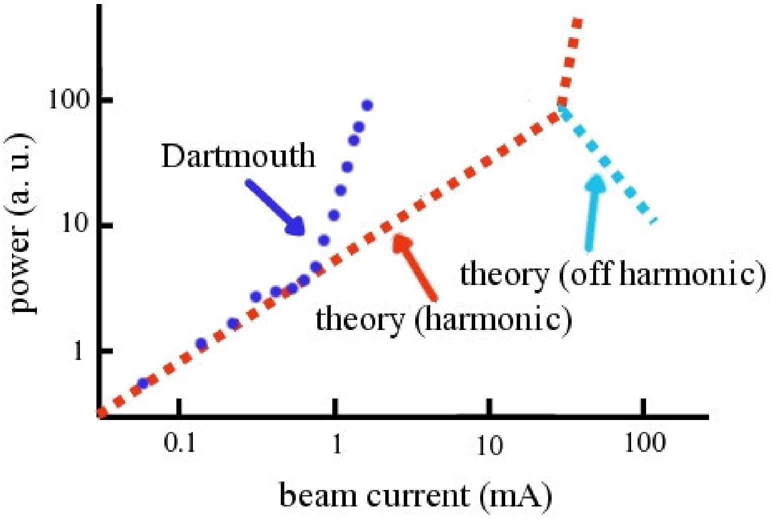

Figure 3-14. Predicted power for given beam current for the Dartmouth experimental parameters. Theoretical curves are shown for conditions where the SP radiation is a harmonic of the fundamental and when the SP radiation at 90˚ is off-harmonic.

where $Z_g$ is the gain length and $k_0$ is the wavenumber at the operating point. At the oscillation threshold, the gain length is

$$Z_g = \frac{2Z}{\sqrt{3}\xi} \tag{3-101}$$

where $Z$ is the grating length and $\xi = 1.97$ in the absence of losses [8].

For the Dartmouth parameters, the optical mode with is ~500 μm. The start current should be increased by the ratio between the mode width and the electron beam width. For this case, that ratio is ~20, and we then expect the start current to be closer to 40 mA. The predicted power vs. current curve is contrasted with the Dartmouth data in Figure 3-14. Since the Dartmouth experiment was observing radiation at 90˚ from the electron beam where the SP radiation is not on a harmonic of the evanescent wave, the theory predicts that the SP radiation is actually suppressed above threshold. The

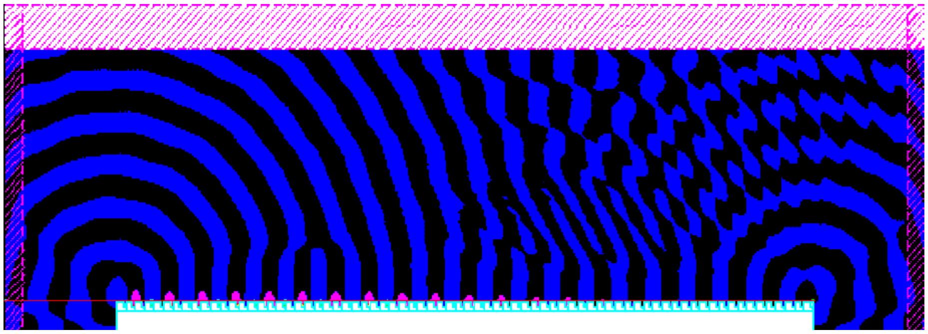

Figure 3-15. Contour map of magnetic field in the $z$-direction from simulations by Donohue and Gardelle for the Dartmouth experiments (oscillator regime). The purple crosshatched regions represent absorbers.

radiation is enhanced on the harmonics by constructive interference and simultaneously suppressed elsewhere by destructive interference.

*Oscillator growth rate with reflections*

The results of PIC code simulations of SPFEL devices make it clear that the reflection and radiation of the evanescent wave at the grating ends plays an important role in the operation of the device. Figure 3-15 shows a contour map of the magnetic field in the $z$-direction produced using the PIC code MAGIC by Donohue and Gardelle[11]. In contrast to earlier simulations, the grating was raised above the absorbers placed at the ends. We note the device is clearly operating in the backward-wave regime, as seen by the growing fields above the grating teeth near the upstream end. We also see that, unlike the previous simulations with absorbers at the ends, radiation is clearly scattering away from both grating ends.

Figure 3-16 shows a contour map of magnetic field produced by Dazhi Li, also using the MAGIC code[12] This figure shows results for an SPFEL operating in the

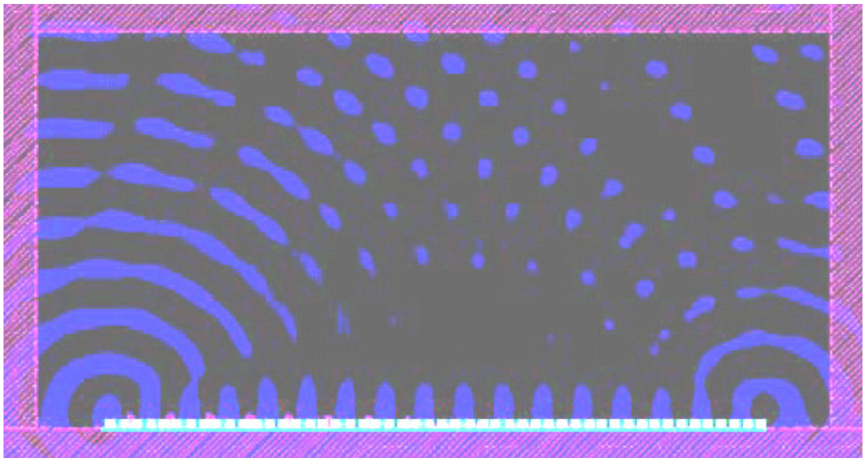

Figure 3-16. Contour map of magnetic field in the $z$-direction from simulations by Dazhi Li for a SPFEL in the oscillator regime (purple crosshatched regions are absorbing).

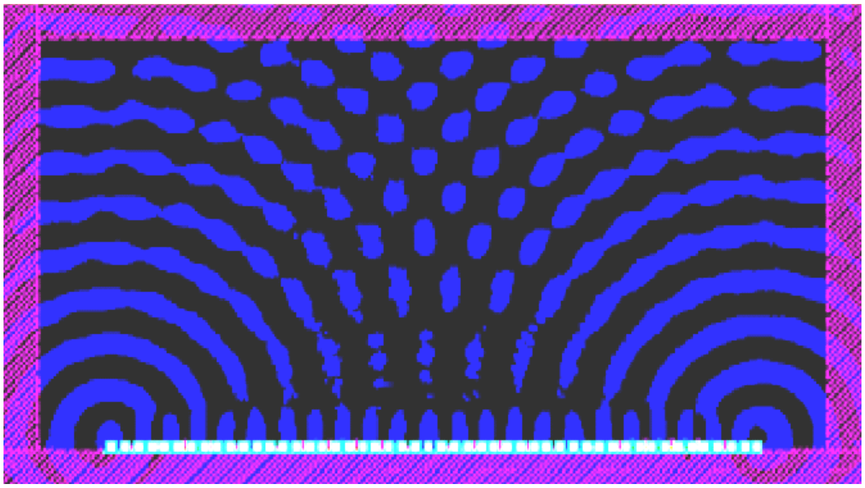

Figure 3-17. Contour map of magnetic field in the $z$-direction from simulations by Dazhi Li for a SPFEL in the amplifier regime (purple crosshatched regions are absorbing).

oscillator regime. Like Donohue and Gardelle's results for the Darmouth grating, we look very closely and see the surface wave stronger at the upstream end than the downstream end. Again, spherical wavefronts emanating from both ends of the grating make it clear that the radiation of the evanescent wave from the ends is significant. Figure 3-17 shows a similar calculation, for the same grating, with a beam voltage corresponding to the amplifier regime. Here, we cannot discern growth of the surface wave in either direction. The radiation of the evanescent wave from the ends of the grating is pronounced, however, and the interference between emission from the upstream and downstream end is clearly observed.

We note that in all of these figures, the evanescent wave radiation appears to come from both ends of the grating. According to the theory presented so far, we would expect to see scattered radiation only from the upstream end of the grating in the oscillator regime and from the downstream end in the amplifier case. Obviously, reflection of the evanescent wave from the ends of the grating is important in these simulation results. The symmetry of the dispersion diagram in the Brillouin zone immediately to the left of the origin insures that for any solution to the dispersion curve, there exists at least one solution having the same frequency and a negative wavenumber. One of the solutions will have the same wavenumber magnitude, as shown in Figure 3-18. The red arrow in the figure shows the equivalent solution in the first Brillouin zone.

To consider a nonzero reflection coefficient for the surface wave at the ends of the grating, we modify our expression for the fields above the grating from Equation (3-82) to

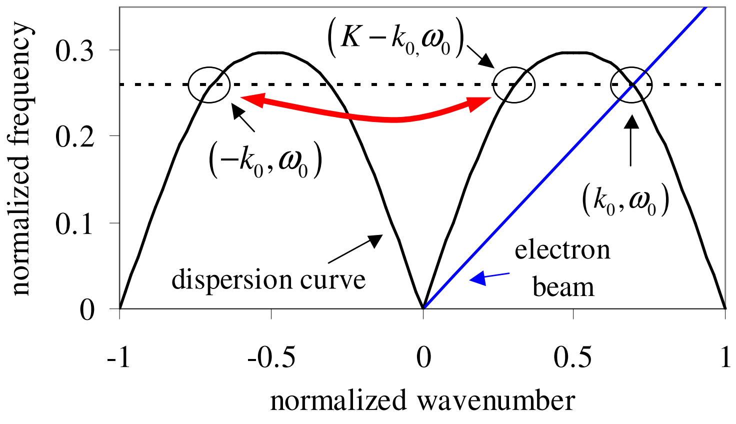

Figure 3-18. Dispersion diagram showing the dispersion curve (black line) and the electron beam phase velocity line (blue line) in the first Brillouin zone and in another complete zone immediately to the left of the origin. The red arrow connects equivalent points on the dispersion relation in different zones.

$$\begin{aligned} E_x = \sum_{j=1}^{3}\left( \sum_{p=-\infty}^{\infty} E_p{}^{(j)} e^{-\alpha_p{}^{(j)} y} e^{ipKx} e^{i\left(k_0 x-\omega_0 t\right)} e^{i\left(\delta k_j x-\delta\omega_j t\right)} \right) \\ + \sum_{p=-\infty}^{\infty} E_p{}^{(F)} e^{-\alpha_p{}^{(F)} y} e^{ipKx} e^{i\left(k_F x-\omega_0 t\right)} e^{i\left(\delta k_F x-\delta\omega_F t\right)} \end{aligned} \tag{3-102}$$

As before, we assume that the shifts in wavenumber among the three backward waves are small enough that the coefficients $E_p$ are the same for the three waves, and the decay constants $\alpha_p$ are equal to the empty-grating values, so Equation (3-102) becomes

$$E_x = E_B e^{i\left(k_0 x-\omega_0 t\right)} \sum_{j=1}^{3} A_j e^{i\left(\delta k_j x-\delta\omega_j t\right)} + E_F e^{i\left(k_F x-\omega_0 t\right)} A_F e^{i\left(\delta k_F x-\delta\omega_F t\right)} \tag{3-103}$$

where

$$E_B = \sum_{p=-\infty}^{\infty} E_p e^{-\alpha_p y} e^{ipKx} \tag{3-104}$$

$$E_F = \sum_{p=-\infty}^{\infty} E_p^{(F)} e^{-\alpha_p^{(F)} y} e^{ipKx} \tag{3-105}$$

and $A_{1,2,3,F}$ are the four constants that weight the contribution of the four waves. We now impose reflection boundary conditions. At the upstream end of the grating the three backward waves are reflected to give the forward wave, so

$$E_F(0, y) A_F = R_0 E_B(0, y) \sum_{j=1}^{3} A_j \tag{3-106}$$

where $R_0$ is the complex reflection coefficient at the upstream end. At the downstream end, the forward wave is reflected into the three backward waves, giving the boundary condition

$$E_B(Z, y) e^{ik_0 Z} \sum_{j=1}^{3} A_j e^{i\delta k_j Z} = R_Z E_F(Z, y) e^{ik_F Z} A_F e^{i\delta k_F Z} \tag{3-107}$$

where $R_Z$ is the complex reflection coefficient at the downstream end of the grating. We can eliminate $A_F$ from these two boundary conditions to get the single condition

$$\sum_{j=1}^{3} \left( e^{i\delta k_j Z} - R_{RT} e^{i\delta k_F Z} \right) A_j = 0 \tag{3-108}$$

where the round-trip reflection coefficient is

$$R_{RT} = R_0 \frac{E_B(0, y)}{E_F(0, y)} R_Z \frac{E_F(Z, y)}{E_B(Z, y)} e^{i(k_F - k_0)Z} \tag{3-109}$$

The round-trip reflection does not depend on $y$ so we can evaluate it at a large distance from the grating where only the $p = -1$ terms in the sums survive. At large $y$ we have

$$R_{RT} = R_0 R_Z e^{i(k_F - k_0)Z} = R_0 R_Z e^{i(K - 2k_0)Z} \tag{3-110}$$

where we eliminate $k_F$, seeing from Figure 3-18 that $k_F = K - k_0$. We also see in that figure that the group velocity at $\left(k_F, \omega_0\right)$ is of the same magnitude and opposite sign as the group velocity at the operating point. Therefore,

$$\frac{\delta\omega}{\delta k_F} = -\beta_g \tag{3-111}$$

and the boundary condition from Equation (3-108) is

$$\sum_{j=1}^{3}\left(e^{-i\xi\delta_j} - R_L e^{-ib\kappa\xi}\right)A_j = 0 \tag{3-112}$$

where

$$R_L = R_{RT} e^{b\nu_0(i-1)Z} \tag{3-113}$$

$$b = \frac{\beta_g + \beta}{\beta_g - \beta} \tag{3-114}$$

The boundary conditions on the backward waves given in Equations (3-91) and (3-92) remain valid, and we replace Equation (3-93) with the condition given in (3-114). The determinant of the coefficients matrix is

$$\begin{vmatrix} 1/\delta_1^2 & 1/\delta_2^2 & 1/\delta_3^2 \\ 1/\delta_1 & 1/\delta_2 & 1/\delta_3 \\ e^{-i\xi\delta_1} - R_L e^{-ib\kappa\xi} & e^{-i\xi\delta_2} - R_L e^{-ib\kappa\xi} & e^{-i\xi\delta_3} - R_L e^{-ib\kappa\xi} \end{vmatrix} = 0 \tag{3-115}$$

The result of including reflections in the calculation is that the start current varies periodically with wavenumber (or voltage) as the reflected forward wave interferes, constructively or deconstructively, with the growing backward wave. Similar behavior has been described for BWOs[13]. The present treatment includes losses as part of the round-trip reflection coefficient. The theory requires as inputs the amplitude reflection coefficients $R_0$ and $R_Z$, which are difficult to estimate analytically. Data from

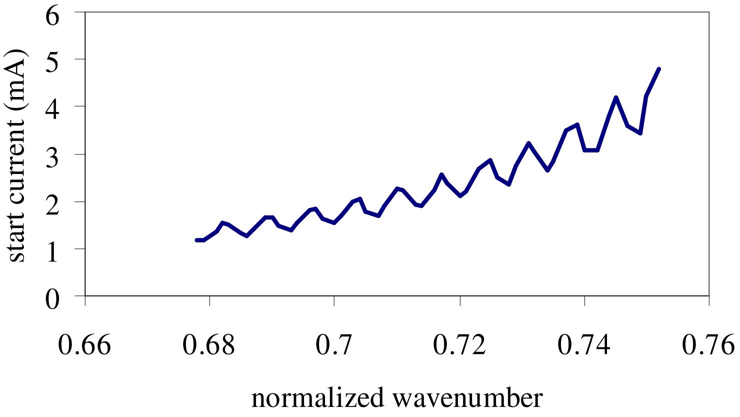

Figure 3-19. Variation of start current with wavenumber considering reflection coefficients estimated from computer simulations.

simulations using the PIC code MAGIC have been utilized to estimate these coefficients[14]. It is found for the Dartmouth experiment parameters that a grating ending in a groove has an amplitude reflection coefficient $R_0 \approx 0.30 - 0.17i$ and one that ends in a tooth has a coefficient $R_0 \approx 0.35 - 0.16i$. The variation in the start current resulting from these reflection coefficients for a relevant range of wavenumbers is shown in Figure 3-19.

## Summary

This chapter has detailed a two-dimensional theoretical framework for analyzing a SPFEL with a rectangularly-grooved grating. The electron beam is represented as a plasma dielectric and Floquet's theorem is used to describe the electromagnetic waves traveling along the grating. The dispersion relation for surface waves on a grating without the influence of an electron beam is derived, and an approximation for the dispersion relation of surface waves in the presence of an electron beam is obtained. Some results for the dispersion relation are presented with a short description of the convergence of numerical calculations of the frequency and group velocity for a given wavenumber. The operation point for the device is calculated by finding the solution along the dispersion curve with phase velocity equal to the electron beam velocity.

Examination of the dispersion relation reveals that the group velocity of the evanescent waves may be positive or negative. The sign of the group velocity at the operating point determines the operation regime of the SPFEL device. A positive group velocity at the operating point indicates the wave grows as a convective instability in the manner of a TWT amplifier. A negative group velocity indicates the disturbance grows as an absolute instability in the manner of a BWO. The theory allows the calculation of the gain parameter (inverse gain length) in the amplifier regime and the growth rate in the oscillator regime. The effect of resistive losses in the grating surface is calculated for both regimes, and the effect of reflections at the grating ends is considered for the oscillator regime.

# References

1 J. Urata, M. Goldstein, M. F. Kimmitt, et al., Phys. Rev. Lett. **80**, 516 (1998).

2 D. A. Watkins, *Topics in Electromagnetic Theory* (John Wiley & Sons, Inc., New York, 1958).

3 J. Donohue, private communication, 2005

4 J. D. Jackson, *Classical Electrodynamics* (John Wiley & Sons, Inc., New York, 1999).

5 C. A. Brau, *Modern Problems in Classical Electrodynamics* (Oxford University Press, New York, 2004).

6 J. A. Swegle, Phys. Fluids **28**, 3696 (1985).

7 J. R. Pierce, *Traveling-Wave Tubes* (D. Van Nostrand, New York, 1950).

8 J. A. Swegle, Phys. Fluids **30**, 1201 (1987).

9 J. T. Donohue and J. Gardelle, Phys. Rev. ST-AB **8**, 060702 (2005).

10 C. A. Brau, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **407**, 1 (1998).

11 J. T. Donohue, and J. Gardelle, in *FEL 2006* (www.jacow.org, Berlin, Germany, 2006).

12 D. Li, in *FEL 2006* (www.jacow.org, Berlin, Germany, 2006).

13 B. Levush, T. M. Antonsen, A. Bromborksy, et al., IEEE Trans. on Plasma Sci. **20**, 263 (1992).

14 H. L. Andrews, C. H. Boulware, C. A. Brau, et al., submitted to New Journal of Physics (2006).

# CHAPTER IV

## ELECTRON EMISSION FROM METAL PHOTOCATHODES

In this chapter we use spheroidal coordinates to model the electric field at the surface of a metal needle and examine the consequences of the model for two kinds of electron emission from sharp and blunt metal needle cathodes. We discuss the theory of field emission based on the work of Fowler and Nordheim [1] and a very simplified version of the three-step model for photoelectric emission by Spicer [2]. We then apply the results via the electric field model to field emission from sharp needles and to photoemission from both sharp and blunt needles. From our model for the emission processes, we predict the emitting area and initial beam divergence for both field-emitted and photoemitted electron beams.

Experimental results are presented for three different photon energies. By changing the surface electric field, we explore the dependence on the emission probability on the height of the potential barrier at the vacuum-metal. Because the tuning of the barrier height is done without changing the wavelength of the incident photons, we can predict the dependence of the current on the electric field without knowing the mean free path for excited electrons within the metal. Production of both tungsten and yttrium needles is discussed along with the electron bombardment heating process used to smooth the needle surfaces.

## Theoretical description of emission from needle photocathodes

We analyze two types of electron emission from our metal photocathodes: field emission, and high-field photoemission. Field emission is a very useful diagnostic for sharp needle cathodes, and allows us to test our electron optical system with a low-current CW beam. Field emission is unsuitable for an SPFEL, however, as it becomes unstable at high current. Photoemission presents several attractive properties for an electron beam source for SPFELs. The emission is in short pulses with high peak currents, and is more stable than field emission. In the following paragraphs, we discuss the geometric enhancement of the electric field at a sharp tip, and briefly review the theory of both field emission and the photoelectric effect at high electric field.

### *Electric fields around a needle tip*

Electromagnetic fields are enhanced around any kind of conducting protrusion or sharp edge. For a needle tip, we use the geometry of the needle surface to calculate the field at the surface when the needle is immersed in an external electric field. The fields around the needle tip have a complicated form in Cartesian or even simple polar coordinates. However, we can model the needle surface as an elongated spheroid, and use prolate spheroidal coordinates to calculate the fields[3 (p. 149), 4]. A prolate spheroid is the surface of revolution generated by rotating an elongated ellipse around its long axis. We start in Cartesian coordinates $(x, y, z)$, and take an ellipse centered on the origin with its major axis along the $z$-axis. Rotating the ellipse around that axis, we get a surface defined by

$$\frac{z^2}{a^2}+\frac{r^2}{b^2}=1 \tag{4-1}$$

where $a$ is the semimajor axis, $b$ is the semiminor axis, and $r=\sqrt{x^2+y^2}$ is the perpendicular distance from the $z$-axis. The foci of the ellipsoid (or prolate spheroid) lie on the $z$-axis at the points

$$z=\pm\sqrt{a^2-b^2}=\pm c \tag{4-2}$$

We then define the spheroidal coordinates $(\rho,\zeta,\varphi)$ according to the equations

$$z=c\rho\zeta \tag{4-3}$$

$$x=r\cos\varphi=c\sqrt{\left(\rho^2-1\right)\left(1-\zeta^2\right)}\cos\varphi \tag{4-4}$$

$$y=r\sin\varphi=c\sqrt{\left(\rho^2-1\right)\left(1-\zeta^2\right)}\sin\varphi \tag{4-5}$$

In these coordinates, the family prolate spheroids with foci defined by Equation (4-2) are surfaces of constant $\rho$, and we can easily set them as constant potential surfaces and find the accompanying fields. The solution for the electric field at the surface of a conducting spheroid in an external electric field is

$$E=E_0\zeta\sqrt{\frac{a^2-c^2}{a^2-c^2\zeta^2}\left[1+\frac{c^2}{\frac{a}{c}\left(a^2-c^2\right)\left(\operatorname{arctanh}\frac{c}{a}-\frac{c}{a}\right)}\right]} \tag{4-6}$$

where $E_0$ represents the external field far from the conductor. For our elongated spheroid, $a \gg b$, and we see large field enhancement at the tip of the spheroid (where $\zeta=1$). The radius of curvature of the surface at the tip is

$$R=\frac{b^2}{a} \tag{4-7}$$

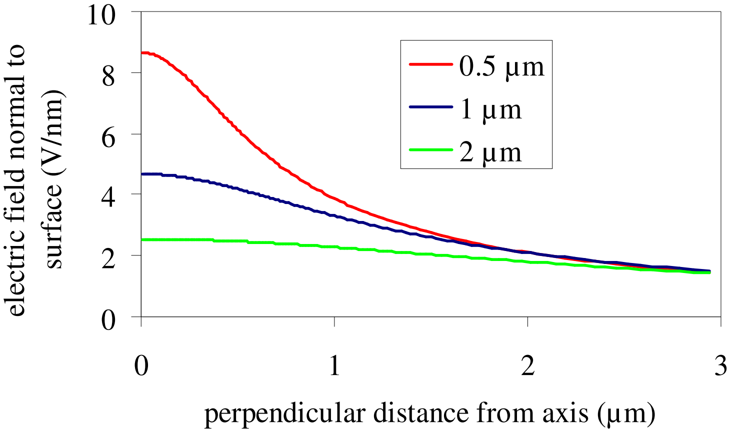

Figure 4-1. Surface electric field in units of V/nm near the tip of a sharp needle. The field is calculated for three different tip radii: 0.5 μm (red line), 1 μm (blue line), and 2 μm (green line).

and the field enhancement factor at the tip is

$$\beta \approx \frac{2L/R}{\ln\left(4L/R\right)} \tag{4-8}$$

where $\beta = E_{tip}/E_0$ and $L = a$ is the length of the needle. In our needle cathode source, we generally use 10-mm long needles. The geometric enhancement factor for such a needle with a tip radius of 1 μm is $\beta \approx 2000$. We can easily produce an external DC electric field of a few MV/m in the accelerator, which becomes several V/nm at the needle tip.

We also use the spheroidal model to calculate the electric field along the surface of the needle away from the tip. The electric field near the tip of a sharp tip is shown in Figure 4-1 for three different tip radii near 1 μm. The figure shows how the field falls off with perpendicular distance from the symmetry axis. This plot is useful to estimate the transverse size of the emitter, and we will use the perpendicular distance from the axis as

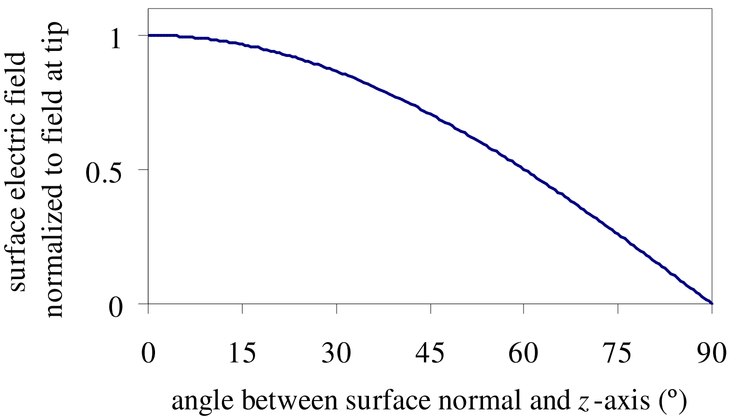

Figure 4-2. Surface electric field on a needle tip of any radius as a function of the angle made by the surface normal and the symmetry axis.

a parameter to evaluate the field emission and photoelectric emission current densities. We can also examine how the electric field changes in terms of the angle made by the surface normal with respect to the symmetry axis. When we plot the field against the surface normal, we see that the field drops off in the same way for any tip radius. Figure 4-2 shows the electric field as a function of the surface normal angle, normalized to the electric field at the tip. A rule of thumb to compare Figure 4-1 and Figure 4-2 is that the angle of the surface normal away from the axis equals 45° at a perpendicular distance from the axis equal to the radius of curvature at the tip. These figures must be considered together to find the effective divergence of an emitted electron beam. Even though the field drops off slowly with emission angle, the differential transverse area element (cross-sectional area presented to an incident laser beam) becomes very small as the emission angle increases.

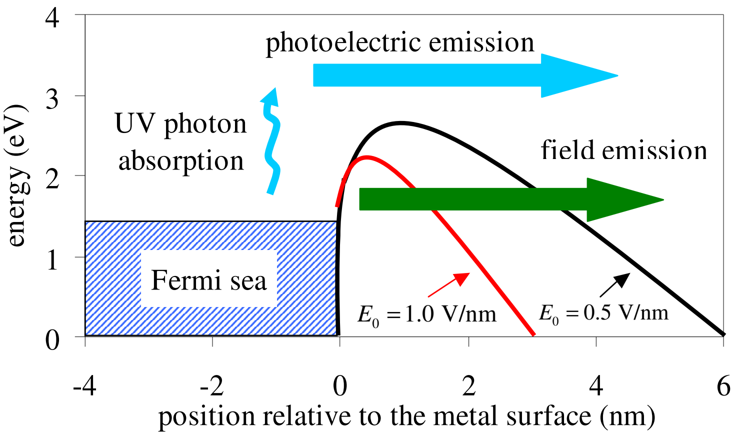

Figure 4-3. Schematic outlining the emission processes for field emission and the photoelectric effect. The potential energy barrier is shown for two different external electric fields. $E_0 = 0.5$ V/nm $E_0 = 1.0$ V/nm

*Field emission*

The theory of field emission by Fowler and Nordheim was an early success of the quantum theory in the 1920s. In a strong electric field, the potential barrier becomes thin between the Fermi sea of electrons inside the metal and the vacuum. Electrons from near the Fermi level are able to tunnel through the barrier and be emitted. A schematic of the emission process for field emission and photoelectric emission is shown in Figure 4-3. As depicted in the figure, the barrier thins with increasing external electric field. Field emission is therefore characterized by a steep voltage dependence of the total current, as the tunneling probability is exponential in the barrier thickness.

We plot the field emission current for sharp tips according to Fowler-Nordheim theory. Using prolate-ellipsoidal coordinates, we can calculate the surface electric field at any point on a needle for some given external field (set by the voltage on the needle and the geometry of the accelerator). The current density is calculated directly from the surface field according to the Fowler-Nordheim equation[1]

$$J_{FE}\left(E\right)=\frac{q_e^{\,3}E^2}{16\pi^2\hbar\phi}\exp\left[-\frac{4}{3}\left(\frac{2m_e}{\hbar^2}\right)^{1/2}\frac{\phi^{3/2}}{q_eE}f\left(y\right)\right] \tag{4-9}$$

where $q_e$ is the charge on the electron, $E$ is the surface electric field, $\hbar$ is the reduced Planck constant, $m_e$ is the mass of the electron, $\phi$ is the work function of the metal, and

$$y=\frac{\sqrt{q_e^{\,3}E/4\pi\varepsilon_0}}{\phi} \tag{4-10}$$

where $\varepsilon_0$ is the vacuum permittivity. The function $f\left(y\right)$ is an elliptic function introduced to take into account the perturbation of the barrier shape by image charge forces. The function has a value of order unity over a wide range of fields, and we approximate it by $\cos\left(y\right)$.

As the current density from any kind of emitter is increased, the field of the electrons in the beam increases. The field of the electrons counteracts the extraction field at the surface of the emitter, and there is a limiting current density for which electrons are actually reflected by the repulsive field created by the beam. The limiting current for a planar cathode-anode was derived by Child and Langmuir[5, 6], and later extended to cylindrical and spherical geometries[7]. The Langmuir-Blodgett current limit for a spherical emitter is given by

$$I_{LB}=\frac{16\pi\varepsilon_0}{9}\sqrt{\frac{2q_e}{m_e}}\frac{V^{3/2}}{g\left[\ln\left(r_c/r_e\right)\right]^2} \tag{4-11}$$

where $V$ is the voltage difference between the emitter and the spherical collector and $g\left[\ln\left(r_c/r_e\right)\right]$ is a function of the log ratio between $r_e$, the radius of curvature at the emitter, and $r_c$, the radius of curvature at the collector[7]. From the values tabulated by Langmuir and Blodgett, for sharp needles $g\left[\ln\left(r_c/r_e\right)\right]^2 \approx 7.2$.

*High-field photoelectric emission*

In photoelectric emission, electrons are promoted to energy levels above the metal-vacuum barrier and are directly emitted into the vacuum. The emission mechanism is contrasted with field emission in Figure 4-3. The conceptual framework for understanding the photoemission process is Spicer's three step model[2, 8]. The three steps are the excitation of an electron within the metal by photon absorption, the transport of the excited electron to the surface, and the escape of the electron over the potential barrier at the vacuum-metal interface. We can greatly simplify the third step by assuming that all excited electrons with energy above the surface potential barrier are transmitted with unit probability, and ignore any tunneling by excited electrons below the threshold. With these conditions, we can also make a simple estimate of the photoexcitation probability based on the number of electrons available in the Fermi sea that can be excited above the barrier by a photon of a given energy. The middle step, transport of excited electrons to the surface of the metal, is more problematic and will also be discussed.

We combine the processes of photoexcitation and escape across the metal-metal interface into a single parameter called the quantum efficiency (QE). The QE is varies with the energy difference between the excited electron and the barrier. As shown in the figure, the height of the barrier (and therefore the resulting energy above the barrier for the excited electron) is reduced at high field by the Schottky effect. The Schottky effect is simply the result of the image charge potential of an emitted electron, and the lowering of the barrier is given by

$$\Delta E_{Schottky} = \sqrt{\frac{q_e^{\,3} E}{4\pi\varepsilon_0}} \tag{4-12}$$

For sharp needles, the Schottky effect can be significant. An electric field of $10^9$ V/m corresponds to an effective reduction in the work function of 1.2 eV.

We ignore tunneling by excited electrons, setting the QE to zero for excited electron energies that are less than the energy at the top of the barrier. Above the barrier, the QE depends on matrix element corresponding to photon absorption and electron promotion from a set of initial states near the Fermi level to a set of final states well above the potential barrier. But we can make a simplified analysis of the dependence of the QE on the energy above the barrier, following Cardona and Ley[9]. As they describe, for a perfect metal surface, the component of the momentum parallel to the surface, call it $p_{\|}$, is conserved by symmetry. To be emitted, an electron must therefore already have a momentum perpendicular to the surface, $p_{\perp}$, corresponding to an energy above the barrier. Cardona and Ley suggest a condition on the "perpendicular energy" of the emitted electron given by (if we add the Schottky effect)

$$\frac{p_{\perp}^{\,2}}{2m} > \Phi + E_F - h\nu - \Delta E_{Schottky} \tag{4-13}$$

where $\Phi$ is the work function of the material, $E_F$ is the Fermi energy, and $h\nu$ is the photon energy. But all the occupied states have energy less than the Fermi energy, so

$$\frac{1}{2m}\left(p_{\|}^{\,2} + p_{\perp}^{\,2}\right) \le E_F \tag{4-14}$$

The QE should be proportional to the number of filled electron states that satisfy these two equations. We integrate in momentum space to get

$$QE \propto \int_{\sqrt{2m(E_F-\Delta E)}}^{\sqrt{2mE_F}} dp_\perp \int_0^{\sqrt{2mE_F-p_\perp{}^2}} p_{\|}dp_{\|} \tag{4-15}$$

where $\Delta E = h\nu + \Delta E_{Schottky} - \Phi$. Performing the integration, we find that the QE is

$$QE \propto 1-\left(1-\frac{\Delta E}{E_F}\right)^{\frac{1}{2}} - \frac{1}{3} + \frac{1}{3}\left(1-\frac{\Delta E}{E_F}\right)^{\frac{3}{2}} \tag{4-16}$$

and we use the binomial expansion (keeping terms up to second order in $\Delta E/E_F$ ) to show that

$$QE \propto \Delta E^2 \tag{4-17}$$

Corrections to this threshold law are possible using information from the electron density of states, but we find this simple analytic dependence law very useful.

Our threshold law for the electron photoyield is then

$$QE = K_{PE}\left(h\nu + \Delta E_{Schottky} - \Phi\right)^2 = K_{PE}\left(\Delta E\right)^2 \tag{4-18}$$

where $h\nu$ is the photon energy, and $K_{PE}$ is a constant representing all the other relevant variables: transition probabilities (as mentioned above), densities of states, etc[10]. We measure the photoelectric current and laser energy directly to calculate the average QE over the surface, but need this threshold law to analyze emission from surfaces of nonuniform electric field. An example, which we presently discuss, is a sharp needle where the tip radius is smaller than the radius of the drive laser. The physical quantity that describes the process of photoemission at all electric fields is the coefficient $K_{PE}$, not the average QE we measure in the lab. $K_{PE}$ is a property of the cathode material, and does not depend on the laser energy or the surface electric field.

Once the electrons are excited with sufficient perpendicular momentum to cross the potential barrier, they must travel from where they are excited to the surface without losing significant momentum in electron-electron collisions. Electrons also experience elastic collisions with the lattice on their way to the surface, increasing the effective distance they travel from the point where they are excited. From the free electron theory of metals, we can calculate the mean time between elastic collisions directly from the DC value of the conductivity, which gives

$$\tau = \frac{m_e \sigma_0}{n_e q_e^{2}} \tag{4-19}$$

where $\tau$ is the mean collision time, $m_e$ and $q_e$ are the mass and charge of the electron, $n_e$ is the number density of free electrons, and $\sigma_0$ is the DC conductivity [3 (p. 345)]. The collisions between electrons do not contribute significantly to the conductivity, and cannot be estimated in a similar way. In the simplest approximation, the mean electron-electron collision time for "hot" electrons varies as

$$\tau(E) \propto \frac{1}{\left(E - E_f\right)^2} \tag{4-20}$$

where $E$ is the energy of the excited electron, and $E_f$ is the Fermi energy. This simple approximation is probably not valid for metals like tungsten with significant d-shell electron density [11]. Where we can estimate or calculate the lifetime of the electrons with respect to momentum-altering collision, we can estimate the mean free path and calculate the probability of the electron traveling to the metal surface with the necessary

Table 4-1. Values of the complex index of refraction and resulting penetration depth for UV photons at several relevant laser wavelengths incident on tungsten.

| Index of refraction and penetration depth for UV photons in tungsten | | | |
|---|---|---|---|
| λ(nm) | Re(n) | Im(n) | α(nm) |
| 337 | 3.1 | 2.6 | 21 |
| 266 | 3.2 | 2.4 | 18 |
| 213 | 2.0 | 3.6 | 9.4 |

momentum to cross the vacuum-metal barrier. Experimental measurements of the hot electron lifetime are available from time-resolved two-photon photoemission with fs laser pulses [12], but not in tungsten.

From the optical properties of the photocathode material, we can calculate the penetration depth for UV photons. That penetration depth is given by

$$\alpha = \frac{4\pi \operatorname{Im}(n)}{\lambda} \tag{4-21}$$

where $n$ is the complex index of refraction and $\lambda$ is the free-space wavelength of the incident photons[13 (p. 101)]. Data for tungsten (from the SOPRA database) are shown in Table 4-1. If we assume that a hot electron loses enough energy in a single electron-electron collision to no longer have enough energy to cross the vacuum-metal barrier, then the average probability for a photoexcited electron to be emitted is

$$P = e^{-f(\alpha)/\mu_f} \tag{4-22}$$

where $f(\alpha)$ is the distance traveled from the point of excitation to the surface and $\mu_f$ is the electron-electron collisional mean free path. The distance given by the function $f(\alpha)$ is the length of a random walk taken by the excited electron as a result of momentum-preserving lattice collisions. That distance is greater than the penetration depth.

We can infer the product of $P$ and $K_{PE}$ using data on emitted photocurrent in low-field experiments. From reported experiments on tungsten with a ps laser pulse at 213 nm, we estimate $K_{PE} \times P$ for tungsten $\approx 4\times10^{-6}$ eV$^{-2}$ [14]. Srinivasan-Rao et al. measured the photoyield at low surface field for a variety of metals with photons at 266 nm[10], and from their low-field QE for yttrium, we get $K_{PE} \times P \approx 2\times10^{-4}$ eV$^{-2}$.

## Consequences for photoemission from sharp and blunt needle cathodes

*Field emission from sharp needle cathodes*

We define sharp needle cathodes as those for which we can measure field emission current on the nA level in our apparatus using accelerating voltages up to 60 kV. Manufacture and smoothing of these tips is discussed below. With smooth, sharp needles, we can predict the field emission current based on the theory presented above. Integrating the Fowler-Nordheim current density over the surface of the needle, we calculate the total current from the needle tip. The total needle current is measured using a floating ammeter connected in series with the needle. We produce a so-called Fowler-Nordheim plot that shows the variation of the current with needle voltage. The only adjustable parameter in our theory is the needle tip radius, and we vary the radius to produce the best fit shown in Figure 4-4. The line in the figure represents the theoretical calculation with a tip radius of 0.72 μm, and the circles represent data taken with the

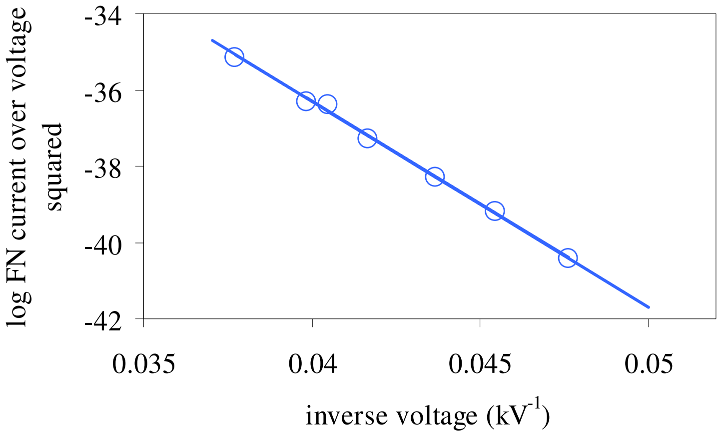

Figure 4-4. Fowler-Nordheim plot of current from a sharp tungsten tip (blue dots) and predicted current for a needle of tip radius 0.72 μm (blue line).

ammeter. Tip radius estimates are confirmed using the scanning electron microscope (SEM) and correlate to the Fowler-Nordheim estimate to within about 10%[15].

As discussed above, there is a space-charge limit for current from any cathode. The Langmuir-Blodgett current density for a sharp needle is

$$J_{LB} = \frac{4\varepsilon_0}{9{r_e}^2}\sqrt{\frac{2q_e}{m_e}}\frac{V^{3/2}}{7.2} \tag{4-23}$$

where we have considered the emitting area to be that of a sphere with radius of curvature equal to the radius of curvature at the needle tip. The ratio between the field emitted current density and the space-charge limit at the tip of the needle has a voltage dependence given by

$$\frac{J_{FE}}{J_{LB}} \propto V^{1/2}\exp\left(-\frac{K_{FE}}{V}\right) \tag{4-24}$$

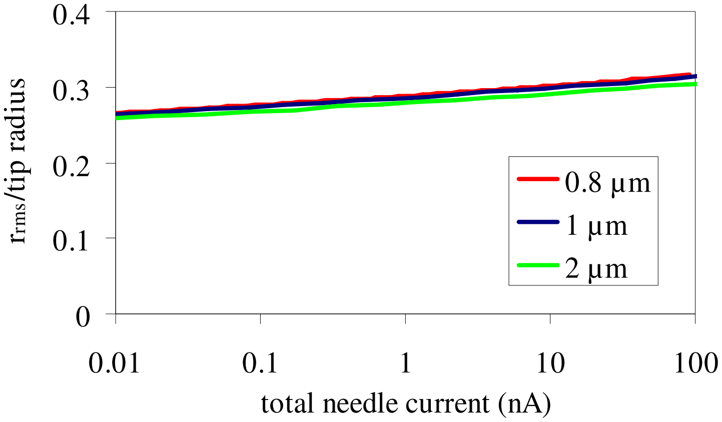

Figure 4-5. Root-mean-square radius for field-emitted electrons normalized by needle tip radius as a function of total needle current for three different needle tip radii.

where $K_{FE}$ varies with the workfunction of the metal. We see that the predicted field emission current exceeds the space-charge limit at high voltage, and at some voltage the Fowler-Nordheim current density will not be realized because of space-charge effects. For sharp needles, however, the Langmuir-Blodgett current density is

$$J_{LB} \approx 10^{12}\ \text{A/m}^2 \tag{4-25}$$

which is four or more orders of magnitude higher than the feasible field-emitted current densities from a single sharp tip. Space-charge can still play some role in limiting the field-emitted current from sharp needles, as field emission is very sensitive to reductions in the surface field [16].

Because field emission varies strongly with voltage (surface electric field), the emitting area is restricted to about one tip radius around the tip. We calculate the root-mean-square perpendicular distance from the axis for the current distribution using the Fowler-Nordheim current density

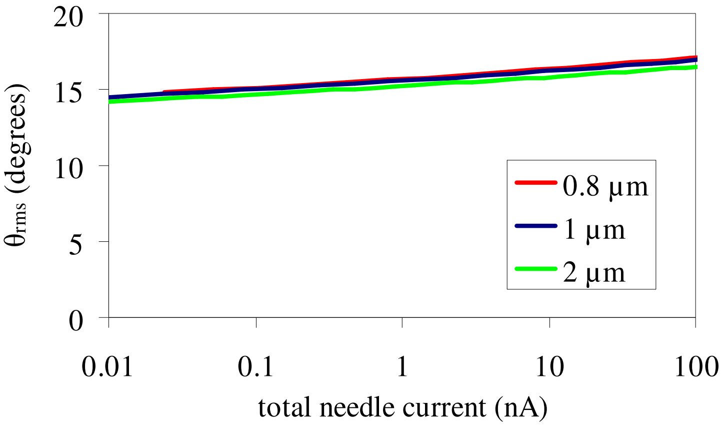

Figure 4-6. Root-mean-square emission angle for field-emitted electrons as a function of total needle current for three different needle tip radii.

$$r_{rms} = \sqrt{\frac{\int J_{FE}\left[E(\mathbf{r})\right]\mathbf{r}^2 dA}{\int J_{FE}\left[E(\mathbf{r})\right]dA}} \tag{4-26}$$

where $E(\mathbf{r})$ is the electric field at any given point $\mathbf{r}$ on the needle surface, and $dA$ is the differential area element on the surface. Figure 4-5 shows the root-mean-square radial extent of the emitted beam at the needle surface normalized to the tip radius for three different tip radii. We see that the emitting area is practically constant relative to the tip radius squared over many orders of magnitude in total needle current. The three tip radii chosen represent sharp tips from which hundreds of nA can be extracted by the field emission process using our spheroidal geometry and up to 50 kV accelerating voltage.

The divergence of the field-emitted beam at the source also changes very little over the range of interest. We calculate the root-mean-square of the emission angle according to

$$\theta_{rms} = \sqrt{\frac{\int J_{FE}\left[E(\mathbf{r})\right]\theta^2 dA}{\int J_{FE}\left[E(\mathbf{r})\right]dA}} \tag{4-27}$$

where $\theta$ is the angle of the electric field at the surface with respect to the $z$-axis. The electric field is normal to the surface everywhere, so Equation (4-27) gives an average of the surface normal angle over the emission area, weighted by the Fowler-Nordheim current density. The root-mean-square emission angles as a function of total needle current are shown in Figure 4-6. The rms emission angle at the surface varies only a few degrees over several orders of magnitude in total current.

The initial divergence angle of the beam is reduced in the accelerator region, as the longitudinal momentum of the beam is increased relative to the transverse components. However, we can estimate the impact of the fields at the emitter by considering that the initial divergence angle is appropriate to describe the electron trajectories within one tip radius of the emitter, and approximate the magnitude of the electric field in that region by the field at the tip. If we make these approximations, the fraction of the electron momentum imparted in the tip-controlled region is

$$\varphi = \sqrt{\frac{E_0 R\beta}{E_0 d}} = \sqrt{\frac{L}{d\ln(2L/R)}} \tag{4-28}$$

where $\beta$ is the electric field enhancement at the tip given by Equation (4-8), $E_0$ is the external electric field, $R$ is the tip radius of the needle, $L$ is the exposed needle length, and $d$ is the distance between the cathode cover and the anode. For a 1-cm long needle with a tip radius of 1 µm and an anode-cathode distance of 2.5 cm, $\varphi$=20%. From this crude analysis (ignoring space charge and the defocusing effect of the anode aperture), we would expect the field-emitted beam exiting the accelerator to have a divergence angle that is 20% of the initial angle, or a resulting divergence around 3° based on the emitted divergence shown in Figure 4-6.

Measurement of field emission current is a powerful diagnostic and allows detection of changes in emitter surface roughness and contamination by adsorbed molecules, both of which are difficult to monitor in other ways. The electron optical system (focusing and steering magnets) can be tested using field-emitted electron beams. However, field emission currents can be modulated only by changing the surface field, which for our accelerator means changing the beam voltage. Further, though a 60-W high voltage supply could source 1 mA at 60 keV, field emission is extremely unstable for currents above several hundred nA. A single sharp tip is damaged by field-emitted currents above tens of μA.

### *Photoemission from sharp needles*

Photoemission from sharp needle tips is a qualitatively different process. Below the space-charge limited current density given in Equation (4-23), the current density is proportional to the laser power. The proportionality constant, the probability of an electron being emitted for each photon absorbed, is the quantum efficiency (QE). Through the Schottky effect, the QE depends on the electric field, and the strong fields created at the tip of a sharp needle significantly enhance the QE. However, the dependence of the photoemitted current density does not vary with voltage as strongly as does the current density for field emission. According to our threshold law, Equation (4-18), the QE varies with the square of the difference between the photon energy and the

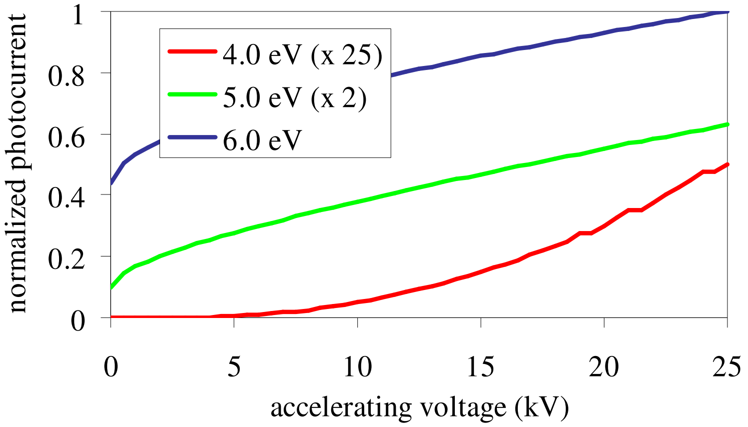

Figure 4-7. Photocurrent emitted from a 1-µm tungsten needle uniformly illuminated by a 10-µm radius laser spot for three different values of the laser photon energy as a function of accelerating voltage. Current is normalized to the current for 6.0 eV photons on tungsten at 25 kV accelerating voltage with additional multipliers identified in the legend.

Schottky-lowered workfunction. The QE then varies only linearly with the voltage, as opposed to the $V^2 \exp(-1/V)$ dependence we see in the Fowler-Nordheim equation.

The total photoemitted current is proportional to the total number of absorbed photons and to the integral of the QE across the surface of the needle

$$I_{PE} = \int_S \Phi(\mathbf{r})\left[1 - R(\theta)\right] QE\left[E(\mathbf{r})\right] \cos\theta dA \tag{4-29}$$

where $\theta$ is the angle made by the surface normal and the $z$-axis, $\Phi(\mathbf{r})$ is the photon flux (number of photons per cross-sectional beam area per second), $R(\theta)$ is the reflectance of the material for a given incident angle, $QE\left[E(\mathbf{r})\right]$ is the quantum efficiency which depends on the electric field at point $\mathbf{r}$ within the laser spot area $S$, and $dA$ is the differential surface area element. The dependence of the QE on the electric field through the Schottky effect is given by Equations (4-12) and (4-18). The $\cos\theta$ simply reflects that the laser intensity is distributed onto the needle cross sectional area, not the actual surface area.

Calculated photocurrent is shown for three different photon energies on a uniformly illuminated, 1-µm tip radius tungsten needle in Figure 4-7. We see from the figure that the calculated photocurrent simply reflects our threshold law for photons with energy equal to the tungsten workfunction. When the photon energy is equal to or above the workfunction, the photocurrent grows with the square root of the voltage for low voltage and approaches a linear growth at high voltage. For lower energy photons, however, there are three regions of current growth. At low voltage, the current rises rapidly with voltage as the Schottky effect increases the emitting area (we examine this in detail below). At higher voltages, the increasing current again reflects the threshold law, displaying growth with the square root of the voltage transitioning to linear growth with field.

In sharp needles, the barrier lowering by the Schottky effect is only pronounced near the tip. Accordingly, the emitting area on the needle surface varies considerably with voltage when the photon energy is below or just above the zero-field workfunction. For a 1-µm tungsten needle at 50 kV, the field at the tip is ~5 V/nm. With 337-µm wavelength laser photons, the QE drops to zero at the edge of a 20-µm diameter laser spot. If we consider instead laser illumination at 266 nm, the energy above the barrier changes from 2.8 eV at the tip of the emitter to 1.0 eV at the edge of a 20-µm diameter laser spot (corresponding to a 40% decrease in QE at the edge). At 213 nm, the energy

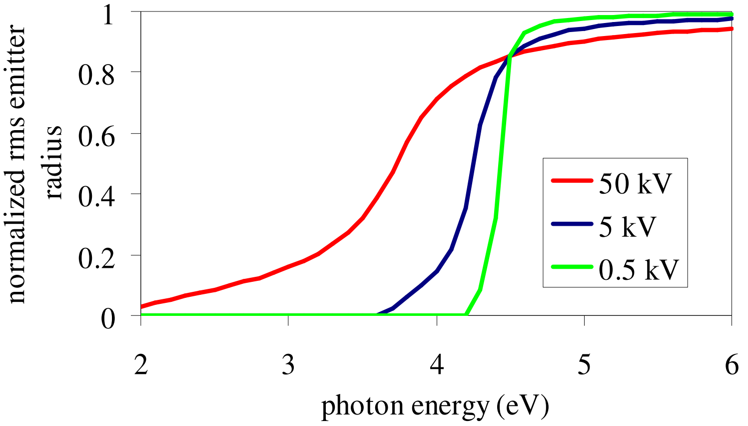

Figure 4-8. Root-mean-square emitting radius as a function of photon energy for a tungsten needle with a 1-µm tip radius illuminated by a uniform laser spot with a 10 µm radius. Rms radii are normalized to the rms emitter radius that would be expected for uniform emission from a circular area.

above the barrier changes from 3.9 eV at the tip to 2.1 eV at the edge of the laser spot (only a 25% decrease in QE).

These examples are helpful to consider the practical impact of the Schottky effect for various cathodes, but we can actually calculate an effective emitter size by integrating our threshold law for the quantum efficiency over the needle surface. The root-mean-square radius of a uniformly illuminated emitter is

$$r_{rms} = \sqrt{\frac{\int_S r^2 QE\left[E(\mathbf{r})\right]\cos\theta dA}{\int_S QE\left[E(\mathbf{r})\right]\cos\theta dA}} \tag{4-30}$$

where $\theta$ is the angle made by the surface normal and the $z$-axis, and $QE\left[E(\mathbf{r})\right]$ is the quantum efficiency which depends on the electric field via the Schottky effect as given in Equation (4-12) and (4-18). If we assume a uniform laser spot with a 10 µm radius on a

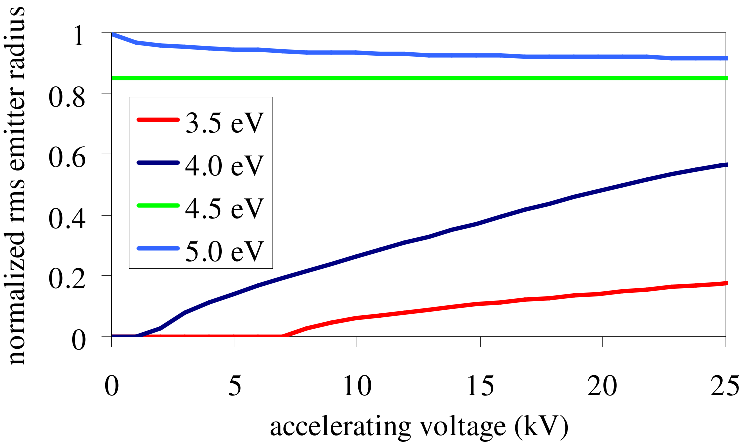

Figure 4-9. Root-mean-square emitting radius as a function of voltage for a tungsten needle with a 1-µm tip radius illuminated by a uniform laser spot with a 10 µm radius and four different values for the photon energy. Rms radii are normalized to the rms emitter radius that would be expected for uniform emission from a circular area.

1 µm tip-radius tungsten needle, the rms emitter radius increases with photon energy as shown in Figure 4-8. The calculation of rms emitter radius shown in the figure is normalized to the rms radius for a uniformly emitting circular area, given by

$$r_{rms} = \sqrt{\frac{\int_0^{R_{laser}} r^3 dr}{\int_0^{R_{laser}} r dr}} = \frac{\sqrt{2}}{2} R_{laser} \tag{4-31}$$

where $R_{laser}$ is the radius of the laser spot. Figure 4-8 shows that for low voltage, the emitter radius dependence on photon energy approaches a step function. There is zero emission below the workfunction (4.5 eV for tungsten) and uniform emission across the laser spot above the work function. At higher voltages, Schottky barrier lowering near the tip allows an emitting area to develop even for photon energies below the workfunction. This emitting area expands to fill the laser spot as the photon energy

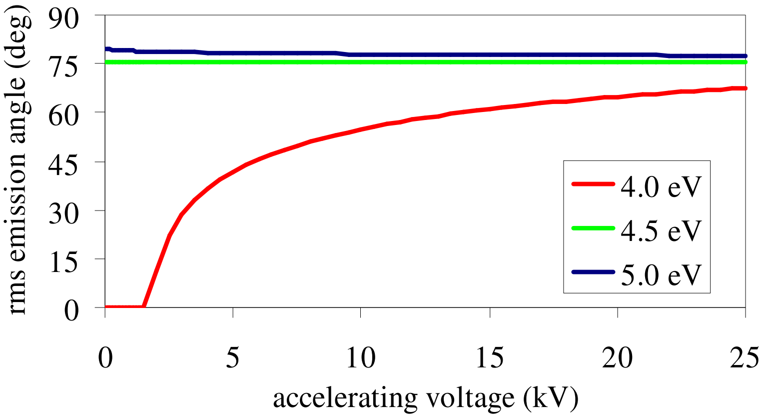

Figure 4-10. Root-mean-square emission angle from a 1-µm tip radius tungsten needle uniformly illuminated by a 10-µm radius laser spot for three different values of the laser photon energy as a function of accelerating voltage.

increases. In the lab, the photon energy is constant and we change the voltage on the needle. The growth of the emitting area with voltage is shown in Figure 4-9. For photon energies at or above the work function, the emitting area is essentially equal to the laser spot area and does not change appreciably with voltage.

Photoemission from a sharp needle tip produces a divergent beam at the source, more so than with field emission, because the emission is not restricted to a small area around the tip. The emission falls off more slowly with surface field than field emission. We can calculate the root-mean-square emission angle of a uniformly illuminated emitter from

$$\theta_{rms} = \sqrt{\frac{\int_S \theta^2 QE\left[E(\mathbf{r})\right]\cos\theta dA}{\int_S QE\left[E(\mathbf{r})\right]\cos\theta dA}} \tag{4-32}$$

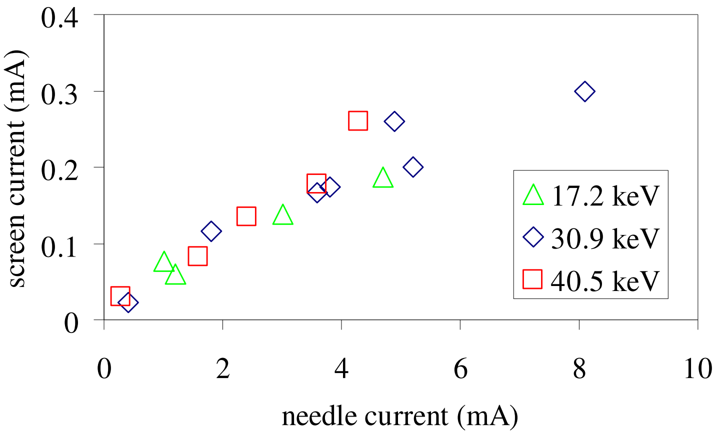

Figure 4-11. Detected current at the focus of the electron beam (phosphor screen) as a function of total needle current. Negligible current is lost after the anode and this screen current therefore represents the current passing the anode.

where $\theta$ is the angle made by the surface normal at the point $\mathbf{r}$ on the surface of the needle with the $z$-axis. Again, the $\cos\theta$ term reflects the cross-sectional area presented to the laser beam. As shown in Figure 4-10, photoemission is much more divergent than field emission, and the root-mean-square angle approaches 80º at high voltage.

In highly divergent beams, the fraction of beam current that passes through the hole in the anode is small. Current detected through the anode from a beam generated by 213 nm laser light on a 2-µm tungsten needle is shown in Figure 4-11 as a function of total needle current. We see that the fraction of the total beam current detected at the focus approaches 10% at low current, dropping to ~4% for total needle currents near 10 mA. In the low-current limit, the detected fraction is a function of the accelerator geometry (most importantly, the tip radius and anode aperture). We interpret the sublinear rise in screen current as the total needle current increases to transverse expansion of the beam due to space charge at the source.

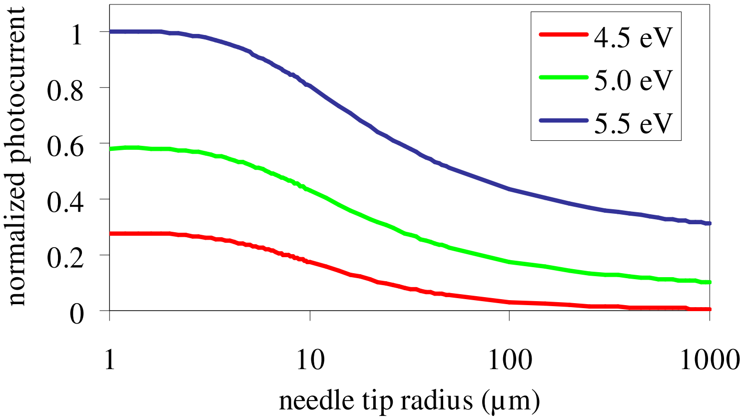

Figure 4-12. Total needle photocurrent as a function of needle tip radius for tungsten needles held at 50 kV and illuminated with a 10-μm radius laser spot at three different photon energies. Needle current is normalized to that for a 1-μm needle tip illuminated at 5.5 eV.

*Photoemission from blunt needle tips*

Any electron focusing system will have a certain acceptance angle, due to the geometry of the anode and the degree to which aberrations in the optics make large numerical apertures impractical. For these reasons, photoemission from sharp needle tips does not produce a useable beam for experiments requiring high current density such as the SPFEL. We therefore investigate photoemission from blunter needles, whose geometry will produce less divergent beams. However, the electric fields produced at the tip of the needle are greatly decreased when the tip radius is increased, so the QE is reduced, as is the space-charge current limit at the source.

We calculate the total photoemitted current from a needle by the integral presented in Equation (4-29). Blunt tips produce less total photocurrent for the same laser spot size than sharp tips, because the surface electric field of blunt tips is reduced

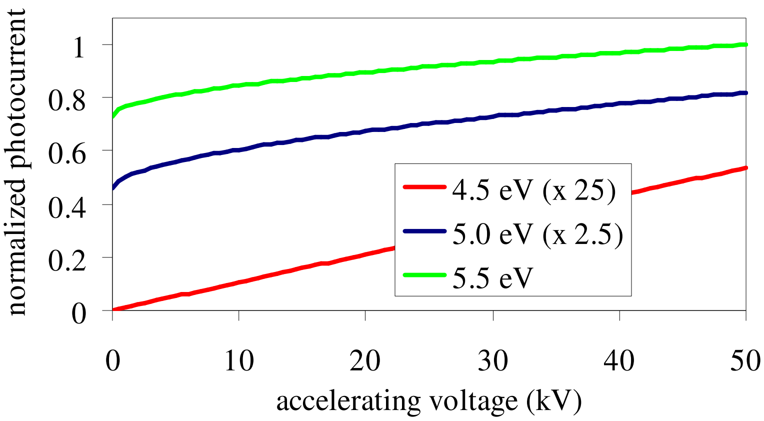

Figure 4-13. Total needle photocurrent as a function of accelerating voltage for 1-mm tip radius tungsten needles illuminated with a 10-µm radius laser spot at three different photon energies. Needle current is normalized to that for a tip at 50 kV illuminated by 5.0 eV photons.

compared to sharp ones. The QE of the photoemission process is therefore also reduced, as it depends on the electric field through the Schottky effect. The total current produced by a 10-µm laser spot on a tungsten needle is presented as a function of the needle tip radius in Figure 4-12. In contrast to the case for field emission, increasing the radius of curvature at the needle tip by several orders of magnitude only reduces the photoemitted current by a factor of two or three. As a function of voltage, photoemitted current increases slowly for photon energies above the workfunction, and more rapidly for energies at or below the workfunction, as shown in Figure 4-13.

In blunt needles, the root-mean-squared emission radius for photoemission is the entire laser spot. Similarly, the root-mean-square emission angle is constant over relevant voltages and photon energies used for photoemission. However, the value of the rms angle of emission for photocurrent is much smaller for blunt needles than sharp ones,

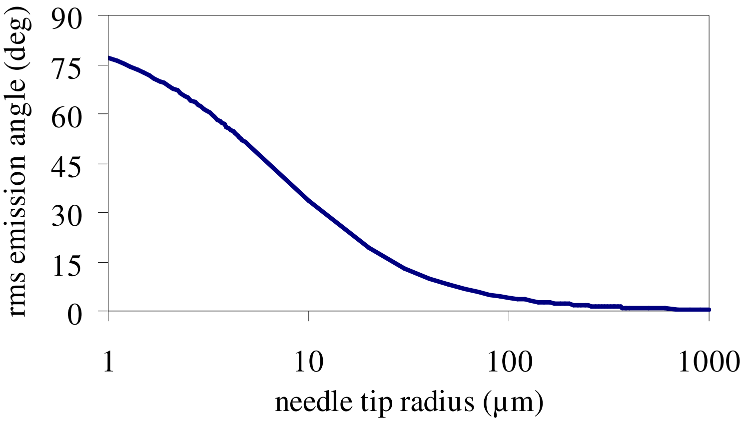

Figure 4-14. Root-mean-square emission angle for photocurrent from tungsten needles illuminated by a 10-µm radius laser spot as a function of radius of curvature at the needle tip. These data are calculated for 5.5 eV photons, but the result is essentially independent of energy for photons above the workfunction.

varying with tip radius as shown in Figure 4-14. In fact, by comparison with Figure 4-6, we see that the photoemitted beam has a lower divergence than field-emitted beams from sharp needles. The emitted beam from a blunt needle is therefore much less divergent than either photocurrent or field emission current from sharp needles and a much greater fraction of the emitted beam can be focused by an electron optical system. As we will show, the reduction in divergence at the source more than compensates the reduced total needle current at lower surface electric field.

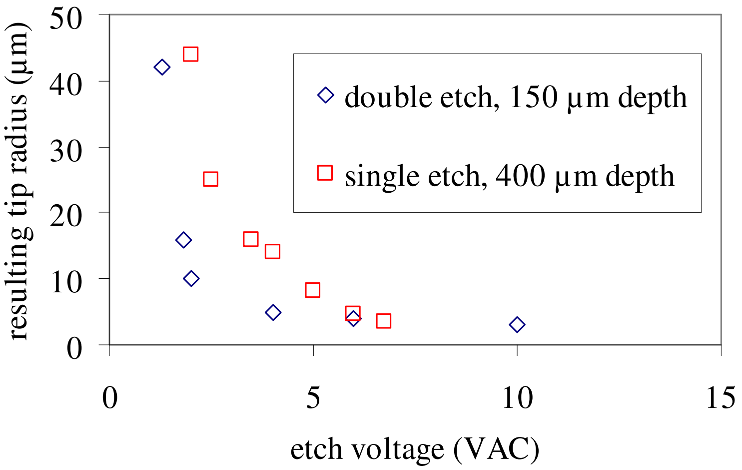

Figure 4-15. Tip radii for tungsten sharp needles made from 150-µm round wire stock by Newport (blue diamonds) and 500 µm round wire stock by Jarvis (red squares).

## Experimental results for photoemission from tungsten and yttrium tips

### *Production and smoothing of sharp needles*

Sharp tungsten needles are made by an etching process from 150-µm diameter tungsten wire stock[17]. The etching is done in a 1-M aqueous solution of sodium hydroxide (NaOH). The needle tip is dipped into the solution by a stage controlled with a finely threaded screw and the needle position is measured with a digital micrometer. Typical depths are .500 mm below the surface of the liquid. An AC voltage at 60 Hz is placed on the needle and a nickel wire, which serves as the other electrode and is also dipped into the solution. By varying the voltage in the electrochemical etch, a variety of tip radii can be fashioned, with the resulting tip radius inversely proportional to the etch voltage. Electrochemically etched tips have been produced in our lab with radii from less than 1 μm up to 50 μm by varying the etch voltage. Multiple etches and differing immersion depths also impact the final tip size, but etch voltage seems to be the primary influence. Sample data taken in our lab by J. D. Jarvis and J. Newport are shown in Figure 4-15 [18].

These sharp tips are mechanically fragile, so they are not polished to get a smooth emitting surface. Instead, they are smoothed by electron bombardment heating after installation in the chamber. The needle is held at several kV positive voltage and an ion gauge filament is rotated into place within several cm of the needle tip. Approximately 2 A of current through the ion gauge filament produces enough thermionic emission that several mA can be collected on the needle. Surface smoothing of sharp tungsten needles takes place when the needle surface is heated above ~2300 K. This is well below the melting point of tungsten metal, which is around 3700 K. (The laser ablation threshold for ultraviolet light is observed around $10^{12}$ W/m$^2$. In most cases, laser ablation damage can be removed by electron bombardment heating above 2400 K.) Evidence of smoothing on a larger-radius needle tip is shown in Figure 4-16.

It has not been possible to produce sharp tips made of yttrium metal. Various attempts at chemical etching left a rough surface covered in deposited salt. Also, the smoothing process by electron bombardment heating is unstable. This instability will be discussed below with respect to smoothing of blunt needles.

### *Previous experiments using Nd:YAG 4$^{th}$ harmonic on sharp tungsten needles*

Photoemission experiments using the 4$^{th}$ harmonic of a Nd:YAG laser at 266 nm (4.7 eV) on tungsten needles have been performed previously in our lab[15, 19]. Data for

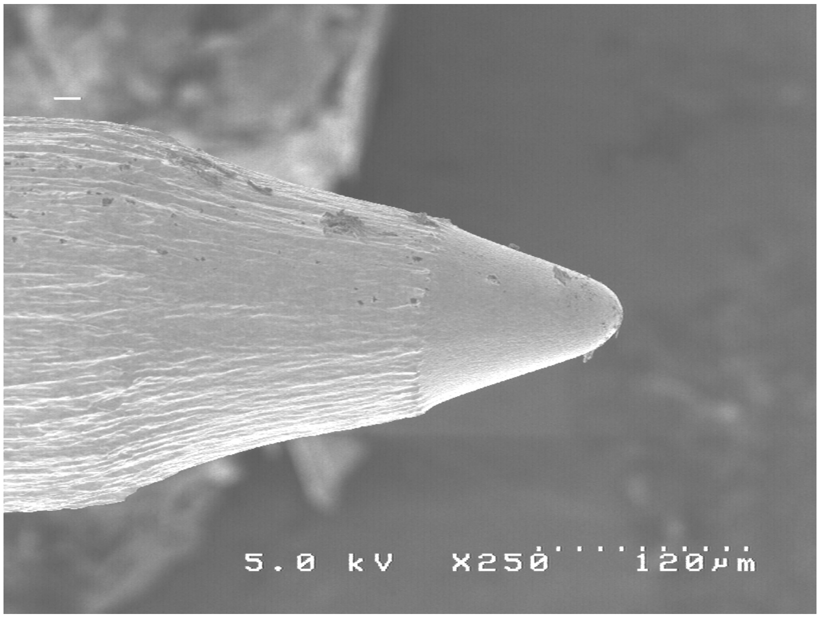

Figure 4-16. Tungsten needle showing evidence of smoothing from electron bombardment heating.

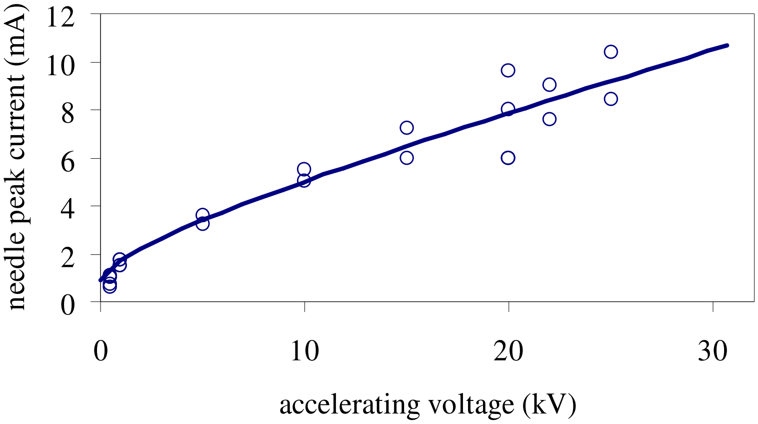

Figure 4-17. Photocurrent from a sharp tungsten needle (tip radius 0.63 μm) illuminated with laser light at 4.66 eV as a function of needle voltage.

a 0.63-μm needle are shown in Figure 4-17 as a function of voltage along with calculated photocurrents using Equation (4-29). The calculated data fit the shape of the curve well except at low fields, which is interpreted as the onset of space charge limiting the total current.

Photoyield (electrons emitted per incident photon) approaching unity has been reported[15], but this result was based on an emitting area corresponding to that for field emission from a sharp needle. The emitting area for photons with energy above the workfunction is much greater than the field emitting area, as we have discussed above, and the maximum photoyield for tungsten needles illuminated with 266-nm photons is closer to $10^{-4}$.

The change in current with voltage corresponds to the change in QE, not to a change in emitting area. From Figure 4-9, we estimate that the emitting area is equal to the laser spot area for all relevant voltages in this case. From Figure 4-10, we see that the emission is very divergent at the source, with a root-mean-square emission angle of ~75 degrees.

### *Photoemission from sharp tungsten needles using a nitrogen laser*

Photocurrent was observed from a sharp tungsten needle under illumination by a nitrogen laser (Laser Science Industries model VSL-337ND-S) at 337 μm (3.67 eV). This laser provided 4-ns long pulses from 1-10 Hz with a pulse energy of ~300 μJ. Beam profiles were taken using a knife-edge method (razor blade on a micrometer-controlled stage), measuring a beam waist of ~50 μm. Transmission of the optical system was measured at ~0.3, so the maximum laser power at the needle surface is $3\times10^{12}\ \mathrm{W/m^2}$.

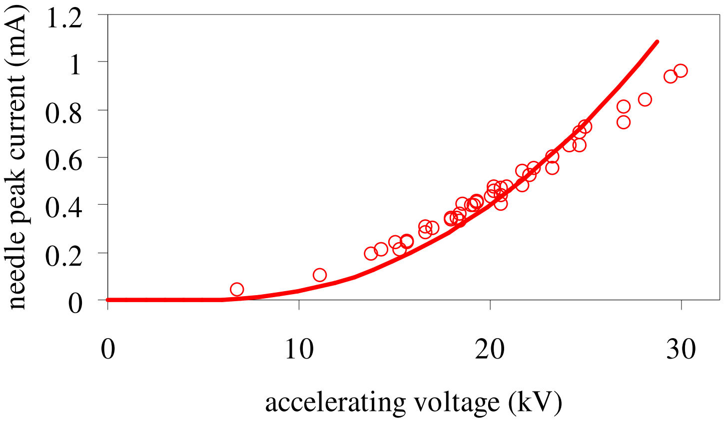

Figure 4-18. Photocurrent from a sharp tungsten needle (tip radius 0.75 μm) illuminated with laser light at 3.67 eV as a function of needle voltage.

The last lens in the optical system was mounted on a rail to allow adjustment of the size of the focal spot at the needle, and used to keep the ultraviolet (UV) laser power at the needle tip between $5\times10^{9}\ \mathrm{W/m^2}$ and $1\times10^{11}\ \mathrm{W/m^2}$. At these power levels and pulse lengths, laser ablation of the cathode was avoided. Spot radii for the UV laser at the needle tip were therefore 100-500 μm (always much greater than the needle tip radius). The laser light was directed onto the needle tip at normal incidence. Figure 4-18 shows measured photocurrent for an incident laser power of $1.2\times10^{10}\ \mathrm{W/m^2}$ as a function of needle voltage. Fowler-Nordheim plots of the tungsten needle used to produce these data indicated a tip radius of 0.75 μm, and this tip radius is used with the photon energy to produce the calculated solid line in the figure according to Equation (4-29).

In this case, the rapid rise of the current as a function of voltage is a result of both increasing QE and increasing emitting area. The situation for 337-nm photons on tungsten is described by the bottom two curves in Figure 4-9, where the root-mean-square emitting radius increases with voltage.

### *Production and smoothing of blunt needle tips*

Blunt needle tips in tungsten are formed mechanically. Tungsten wire stock (150 μm diameter) is loaded into a Dremel rotating tool with a specially designed collet. After careful alignment to prevent wobbling, the tool is turned on and the rapidly spinning wire stock is shaped to the desired tip radius with several degrees of increasingly fine sandpaper. Using the rotating Dremel tool insures symmetry of the resulting needle tip. This process has been used to create needles with radii of curvature at the tip from hundreds of microns to several millimeters. Tips with radius of curvature larger than the radius of the wire stock have rounded edges with a smaller radius of curvature than that of the flat tip. All of these mechanically formed tungsten tips are smoothed *in situ* by electron bombardment heating.

We have also made blunt tips from square wire stock of yttrium metal, but with much less control over the final shape than we have using tungsten. Yttrium is not as safe to work with as tungsten, so needles are initially prepared by hand with sand paper using glove bags (heavy industrial rubber gloves attached to a clear plastic bag allowing wire stock to be handled and shaped within an airtight enclosure). It has been difficult to keep the mechanically formed shape through the electron bombardment heating and smoothing process. The melting point of yttrium at 1800 K is easily reached by our electron bombardment process, and because the heating is unstable at high temperatures, complete melting of the end of the needle is common.

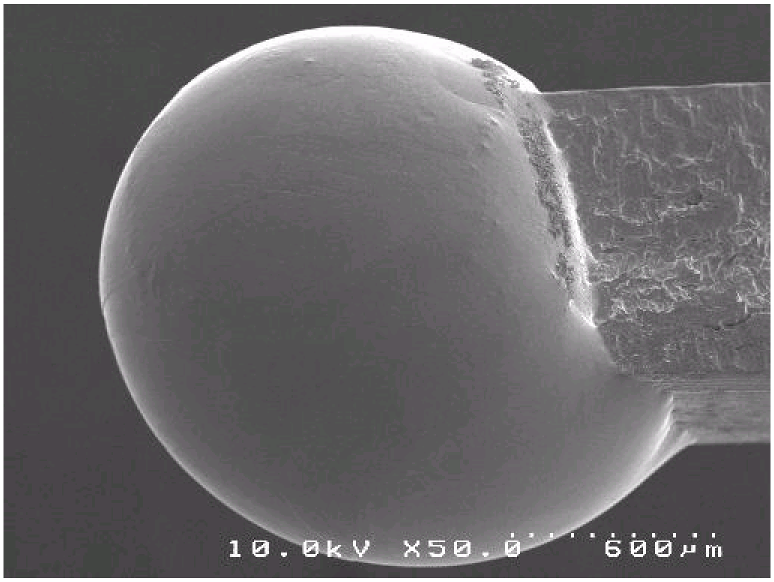

Figure 4-19. Yttrium cathode with a tip radius of ~700 μm formed by melting the end of square yttrium wire stock.

The instability in the electron bombardment heating is qualitatively understood. When electrons strike the yttrium surface and heat it, some amount of yttrium is removed from the surface of the needle. To the degree that this yttrium is deposited back onto the tungsten filament used in the heating process, the workfunction of the filament surface is reduced. As a result, the electron emission from the filament is increased until the deposited yttrium evaporates or sublimates from the filament surface. The increased emission, however, causes more yttrium to be removed from the needle and deposited on the filament. This positive feedback loop increases the electron current into the yttrium needle until the current limit of the positive HV supply is reached. At that point, the voltage on the needle drops, and with it the temperature, shutting off the removal of yttrium from the surface until the current is reduced below the limit of the supply. The supply then begins to raise the voltage of the needle and the cycle repeats.

By carefully controlling the current in the tungsten filament, this runaway process can often be avoided, but the existence of the instability makes careful heating and smoothing of yttrium needles extremely difficult. Often, the end of the wire melts into a ball with a tip radius on the order of 1 mm. Despite the difficulty in controlling the melting process, the blunt cathodes produced in this way have made interesting photocathodes. A scanning electron microscope (SEM) of a cathode formed in this way is shown in Figure 4-19.

### *Photoemission from blunt tungsten needles using 5th harmonic of an Nd:YAG laser*

Photoemission experiments were performed using the 5th harmonic from a Polaris II laser (New Wave Research) with an aperture installed inside the resonator cavity to select the TEM00 mode. The laser provided ~4 ns pulses of 213 nm photons with a total pulse energy of up to 50 μJ, at a repetition rate of 20 Hz. A two-lens beam expander system increased the radius of the beam to a diameter of ~1 cm before it was focused onto the needle tip by a final lens with a focal length of 400 mm. Based on individual transmission measurements of the lenses (manufactured by Del Mar Photonics, Inc.) and reflection measurements of the mirrors (CVI Laser, LLC) at 213 nm, the optical transmission of the system was estimated at 40%.

To estimate the photon density at the needle, we need to know the laser spot size there. Unfortunately, the laser energy required to trigger our power meter gives enough power at the laser focus to destroy the razor blades used for beam profile measurements. We can perform knife edge measurements at a much higher f/number and calculate the

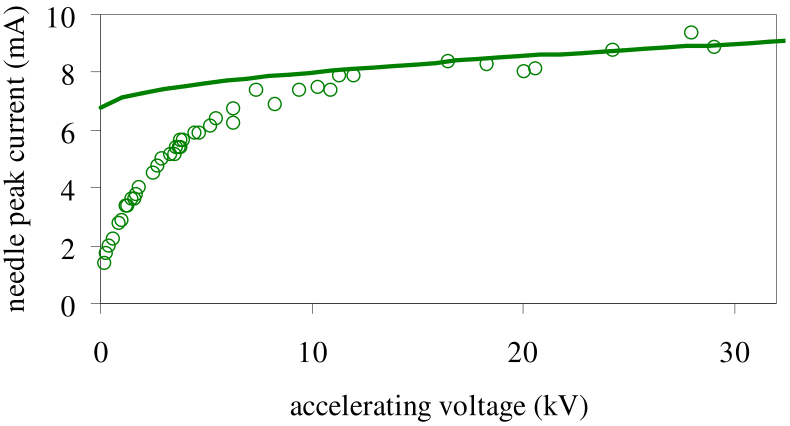

Figure 4-20. Photocurrent from a blunt tungsten needle (tip radius 270 µm) illuminated with laser light at 5.81 eV as a function of needle voltage.

beam propagation factor $M^2$ defined by

$$M^2 = \frac{\pi w \theta}{\lambda} \tag{4-33}$$

where $w$ is the measured beam radius, $\theta$ is the convergence half-angle, and $\lambda$ is the wavelength of the laser light. At low $\theta$, we find $M^2 \approx 5$, where $M^2 = 1$ for a diffraction-limited beam. We extrapolate to our f/40 optics (400-mm final lens, 5-mm expanded beam radius) to get a spot size of 80 µm at the needle. The estimated maximum laser power density is $6\times10^{11}$ W/m$^2$. The Polaris II laser has a built-in attenuator, so we need not change the spot size at the needle to control the power at the needle tip.

Photoemission data from a 270-µm radius tip are shown in Figure 4-20. The data show that the current rises more slowly with voltage than the longer-wavelength photoemission from sharp needles. For blunt needles, the QE changes very little with

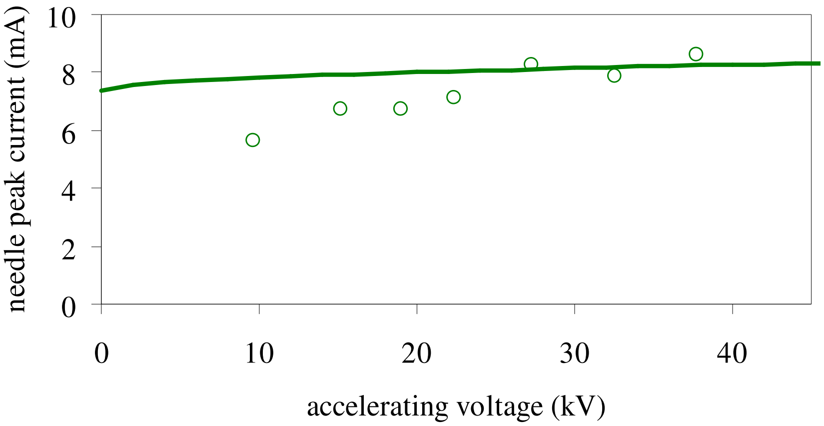

Figure 4-21. Photocurrent from a blunt yttrium needle (tip radius 700 μm) illuminated with laser light at 5.81 eV as a function of needle voltage.

voltage, simply because the surface fields are much smaller. The emitting area is also constant for this case because the photons are well above threshold for all surface fields. The failure of the fit at low voltages is attributed to the limitation on the total current by longitudinal space charge.

### *Photoemission from blunt yttrium needles using 5th harmonic of an Nd:YAG laser*

The same Polaris II laser has been used in photoemission experiments with blunt yttrium cathodes. The workfunction of yttrium is 2.9 eV, and the $K_{PE} \times P$ inferred from the literature is higher than tungsten. Both factors contribute to a photoyield that is higher than that of tungsten photocathodes. As is the case for blunt tungsten needles, the emitting area and QE for blunt yttrium needles is relatively insensitive to voltage, except where the current is reduced by space charge at low fields. Data for these experiments is presented in Figure 4-21.

# Summary

We have discussed both field emission and photoelectric emission from needle cathodes under various conditions. We have analyzed the effect of the needle geometry using prolate spheroidal coordinates to model the electric field at the needle surface, and calculated the effective emitting areas and source divergence of electron beams produced by each process. Our model for field emission uses the Fowler-Nordheim theory, and we construct a very simplified version of Spicer's three-step model for photoelectric emission. The results for field emission allow us to estimate the radius of curvature at the tip of the needle in good agreement with radii measured from SEM pictures of the needle. In the absence of data on hot electron lifetimes within the cathode, we can still reproduce the dependences of electron current on voltage in a variety of situations with sharp and blunt needle photocathodes made of tungsten or yttrium metal.

The values of $K_{PE} \times P$ used to fit our experimental data suggest orders of magnitude change in the transport probability of hot electrons from the point of photoexcitation to the surface of the cathode over the range of photon energy we explored. It is plausible that the mean free path for excited electrons changes drastically enough between 3.5 and 5.8 eV above the Fermi level to explain these results[11]. We expect $K_{PE}$ to be essentially constant over this range, with the variability in the photoexcitation mostly absorbed into the quadratic dependence of QE on energy above the barrier. Detailed measurements of hot electron lifetimes using time-resolved two-photon photoemission experiments would be possible, and could be used with these data to set bounds on the variability of the transition matrix elements and density of states information that is absorbed into $K_{PE}$ in our simple model.

We have demonstrated that modeling the electric field at the needle surface is necessary for interpreting photoemission results for sharp needles. The emitting area can vary greatly for photons with energies near the workfunction. For photon energies well above the workfunction, the emitting area is much greater than the corresponding area for field emission, which is always on the order of the tip radius. A thorough understanding of the actual emitting area and the beam divergence of field and photoemitted beams is necessary to interpret recent experimental work on developing sharp needle cathodes as drivers for free-electron lasers[20].

# References

1 R. H. Fowler and L. Nordheim, Proc. Roy. Soc. Lond. A **119**, 173 (1928).

2 C. N. Berglund and W. E. Spicer, Phys. Rev. **136**, 1030 (1964).

3 C. A. Brau, *Modern Problems in Classical Electrodynamics* (Oxford University Press, New York, 2004).

4 K. L. Jensen, D. W. Feldman, and P. G. O'Shea, J. Vac. Sci. Technol. B **23**, 621 (2004).

5 C. D. Child, Phys. Rev. (Series I) **32**, 492 (1911).

6 I. Langmuir, Phys. Rev. **2**, 450 (1913).

7 I. Langmuir and K. B. Blodgett, Phys. Rev. **24**, 49 (1924).

8 N. V. Smith and W. E. Spicer, Phys. Rev. Lett. **23**, 769 (1969).

9 M. Cardona and L. Ley, in *Photoemission in Solids*, edited by M. Cardona, and L. Ley (Springer-Verlag, Berlin, 1978), Vol. 1, p. 1.

10 T. Srinivasan-Rao, J. Fischer, and T. Tsang, J. Appl. Phys. **69**, 3291 (1990).

11 V. P. Zhukov, F. Aryasetiawan, E. V. Chulkov, et al., Phys. Rev. B **65**, 115116 (2002).

12 R. Haight, Surf. Sci. Rep. **21**, 275 (1995).

13 R. Menzel, *Photonics* (Springer-Verlag, Berlin, 2001).

14 C. Tomas, J. P. Girardeau-Montaut, M. Afif, et al., Applied Physics A: Materials Science & Processing **V64**, 467 (1997).

15 C. Hernandez Garcia and C. A. Brau, Nucl. Inst. Meth. Phys. Res. A **483**, 273 (2002).

16 J. P. Barbour, W. W. Dolan, J. K. Trolan, et al., Physical Review **92**, 45 LP (1953).

17 C. Hernandez Garcia and C. A. Brau, Nucl. Inst. Meth. Phys. Res. A **429**, 257 (1999).

18 J. D. Jarvis and J. Newport, unpublished data, 2002-2004

19 C. Hernandez Garcia, 2001, *Photoelectric field emission from needle cathodes induced by CW and pulsed lasers.* PhD Thesis. Vanderbilt University.

20 R. Ganter, et al., J. Vac. Sci. Technol. B **24**, 974 (2006).

# CHAPTER V

## ELECTRON BEAMS FROM METAL PHOTOCATHODES

This chapter will introduce the concept of electron-beam quality as described by the trace space representation of the beam. In trace space, the distance of each particle from the beam center is plotted on one axis and the trajectory angle for that particle on the other. From the effective area that encloses the beam in this projection and the average current density over that area, we can determine the current that can be focused into a given volume in real space using linear electron optical elements.

The electron-beam quality required for a Smith-Purcell Free Electron Laser (SPFEL) is then expressed in the trace-space quantities of emittance and brightness. We discuss the quality of a thermionic electron source in terms of its temperature and that of a photoelectric cathode in terms of the workfunction and photon energy. The initial beam quality is preserved by forces which vary linearly with transverse distance from the beam center, but is degraded by nonlinear forces. Such forces may arise from space-charge forces at the beam focus and spherical aberration in the focusing solenoid. The magnitude of the spherical aberration effect is calculated for various magnet designs, and the beam spot sizes produced by this aberration are found in the absence of other nonlinear effects. The aberration spot size is shown to be much larger than the spot sizes resulting from space charge and source quality.

The electron beam apparatus will be described. The principal electron-optical element is the focusing solenoid. Electron beam spot sizes produced by the solenoid are measured using knife-edge experiments and also from CCD camera images of the phosphor screen. These measurements produce spot sizes larger than those calculated from the spherical aberration of the solenoidal lens. Effects of spherical aberration can be observed, nonetheless, in the beam current distribution near the focus.

## Theoretical description of electron beams

To describe the electron beams produced by metal photocathodes and what happens to the beams as they pass through the various electron optical elements in the beamline, we need some statistical descriptions of the beam as a whole. As discussed below, the most critical parameter for the operation Smith-Purcell Free-Electron Laser (SPFEL) is the amount of current that can be focused within the scale height of the evanescent wave above the grating. The amount of current it is possible to focus into a given area depends on the density of electrons in the transverse trace space of the beam. The beam emittance represents the area of the beam in transverse trace space, and we use the emittance to calculate the density in transverse phase space, called the beam brightness [1].

### *Transverse trace space, beam emittance, and brightness*

Instead of using the more common phase space representation of the beam based on the position and momentum of the electrons, the emittance is based on the uncorrelated area of the beam in the closely related trace space. In trace space, we plot the angle made by the particle trajectory and the beam centroid trajectory as a function of the transverse position of the particle relative to the center of the beam. The emittance is just the uncorrelated area of the beam in this space. Uncorrelated means that we take out the part of the area that results from a general convergence or divergence of all the beam trajectories. We get the uncorrelated area by measuring the beam at the focus or by subtracting a term from the measured area that reflects the correlation between the transverse position and trajectory angle.

A practical definition of the emittance must reflect the uncorrelated area of a simply connected region that encompasses the beam, where by simply connected we mean a region with no holes. The beam density is not constant, however, so the outlying trajectories that would define the edge of such a region represent a negligible fraction of the beam current. As a compromise, we define a region that encompasses some fraction of the beam. Our convention is to discuss the root-mean-square (rms) emittance defined by

$$\varepsilon_y = \sqrt{\left\langle y^2 \right\rangle \left\langle y'^2 \right\rangle - \left\langle yy' \right\rangle} \tag{5-1}$$

where $y$ is the position of a given particle along a transverse axis, $y'$ is the derivative of the position of the particle with respect to its position along the beam axis, and the pointed brackets represent averages over all the particles in the beam. We consider that the average beam position is on the $x$-axis, and the average beam velocity is along that axis. The fraction of the beam contained by a simply connected surface with area given by Equation (5-1) depends on the details of the electron distribution. As discussed briefly above, the correlation term $\left\langle yy' \right\rangle$ represents a divergence of the beam as a whole. (Throughout this chapter, the $x$-axis, the longitudinal direction, and the beam axis all refer to the direction of the average motion of particles in the beam. The $y$- and $z$-axes represent the transverse directions.)

We can use the beam emittance to describe the electron beam at a focal point. In the limit of small aberrations in the focusing optics, the emittance is constant as the beam passes through linear electron optical elements. The contribution to the emittance from the three terms in Equation (5-1) does change however. The average over the trajectory angles changes within electron optical elements, but is constant in drift space by conservation of momentum. Importantly, the correlation term vanishes when the beam is brought to a focus. At the focus, the rms transverse position of the particles is

$$\sqrt{\left\langle y^2 \right\rangle} = \frac{\varepsilon_y}{\sqrt{\left\langle y'^2 \right\rangle}} \tag{5-2}$$

For round beams $\left(\varepsilon_y = \varepsilon_z = \varepsilon\right)$, the rms radius of the beam is

$$\left\langle r^2 \right\rangle = 4\frac{\varepsilon^2}{\left\langle r'^2 \right\rangle} \tag{5-3}$$

where $r = \sqrt{y^2 + z^2}$.

For a beam focus in a field-free region, we approximate the rms trajectory angle by the divergence full-angle

$$\alpha = \frac{A}{f} \tag{5-4}$$

where $A$ is the aperture diameter of the beam at the last lens, and $f$ is the focal length of that lens. The rms radius of the beam at the focus is then

$$\sqrt{\left\langle r^2 \right\rangle} = 4\frac{\varepsilon}{\alpha} \tag{5-5}$$

For this reason, $4\varepsilon$ is sometimes referred to as the effective emittance [1].

When the beam is accelerated, the area in the transverse trace space is reduced as the longitudinal momenta of the electrons are increased relative to their transverse momenta. We can, however, define an invariant normalized emittance that does not change with beam acceleration. The normalized emittance is

$$\varepsilon_n = \gamma\beta\varepsilon \tag{5-6}$$

where $\gamma$ is the relativistic Lorentz factor, and $\beta c$ is the particle velocity.

The beam emittance places a limit both on the depth of field of the focus and on the size of the focal spot that can be obtained from a given electron beam. We can then define a beam density based on the total current of the beam and the transverse size resulting from the beam emittance. This parameter is the electron-beam brightness, defined as the current per area per divergence solid angle

$$B(x,y) = \frac{dI}{dA\,d\Omega} \tag{5-7}$$

where $I$ is the current in the beam, $A$ is the transverse area, and $\Omega$ is the solid angle. The brightness is a local quality of the beam, that is, it varies from point to point in the transverse beam profile. An average brightness is customarily defined as the density of electrons in the four-dimensional volume corresponding to the emittance in the two transverse dimensions. The average brightness is

$$\bar{B} = \frac{\eta I}{\varepsilon_y \varepsilon_z} \tag{5-8}$$

where $\varepsilon_y$ and $\varepsilon_z$ are the emittances in the two transverse directions, and $\eta$ is a parameter which depends on the electron beam distribution. The results to be discussed in this dissertation to not depend critically on the precise distribution of the electron beam nor the exact value of $\eta$, which is of order $\frac{1}{8\pi^2}$ [2].

We also define an invariant normalized brightness that does not change with beam acceleration, based on the normalized emittances. The normalized average brightness is

$$\bar{B}_N = \frac{\bar{B}}{\beta^2\gamma^2} = \frac{\eta I}{\beta^2\gamma^2\varepsilon_y\varepsilon_z} \tag{5-9}$$

which, in the non-relativistic limit, is proportional to the average brightness divided by the beam energy.

## Requirements on the electron beam for a Smith-Purcell free-electron laser

The interaction of an electron with an evanescent wave traveling along the surface of a grating decreases with the height of the electron above the grating. The decrease is exponential following the decrease of the field, and applies to both incoherent emission of Smith-Purcell (SP) radiation and startup for a SPFEL. In the case of incoherent SP emission normal to the grating, the characteristic scale height is

$$l_{SP} = \frac{\gamma\beta c}{2\omega} = \frac{\gamma\beta\lambda}{4\pi} \tag{5-10}$$

where $\gamma$ is the Lorentz factor for the electron beam, $\beta c$ is the electron velocity, and $\lambda$ is the free space wavelength of the emitted radiation [3].

For the SPFEL device, different space harmonics of the surface wave have different evanescent scale heights above the grating. The characteristic scale height is

$$l_p = \frac{1}{2\alpha_p} = \frac{1}{2\sqrt{\left(k + pK\right)^2 - \left(\omega/c\right)^2}} \tag{5-11}$$

where $k$ is the wavenumber of the surface wave on the grating, $\omega$ is the frequency, $K$ is the wavenumber of the grating, $c$ is the speed of light, and $p$ is the order of the space harmonic. The factor of 2 appears to calculate the intensity scale height, rather than the amplitude height. For all points $\left(k,\omega\right)$ within the allowed region of the dispersion diagram, and all values of $p$, this expression is real and positive. In the amplifier regime, $0 < k < K/2$, and the longest scale height corresponds to the $p = 0$ spatial component. In the oscillator regime, $K/2 < k < K$, and the spatial component $p = -1$ has the longest associated scale height. When $p = 0$, this formula gives the same result as Equation (5-10).

To efficiently drive the surface wave, the electron beam must be focused within the scale height of the evanescent wave of one of the space harmonics. The restriction on the electron beam is then that it have a spot size at the focus equal to about the scale height, and not expand significantly over the length of the grating. The Rayleigh range is an optical concept, but an analog exists for electron beams. Ignoring the Coulomb repulsion of the electrons, a beam in a field-free region follows the hyperbolic envelope equation

$$w\left(x\right)^2 = w_0{}^2 + \frac{\varepsilon^2 x^2}{w_0{}^2} \tag{5-12}$$

where $w\left(x\right)$ is the diameter of the electron beam, $x$ indicates longitudinal position along the beam axis, $w_0$ represents the diameter at the focus $\left(x = 0\right)$, and $\varepsilon$ is the emittance of the beam. The emittance is the area of the beam in transverse trace space. In the transverse dimension $y$, a trace space plot would be a set of points representing the electrons' coordinates in spatial position and transverse angle $dy/dx$. The emittance is the area within a contour drawn on that plot which encloses some defined fraction of the particles[1]. According to the envelope evolution described by Equation (5-12), the transverse area of the beam doubles at a longitudinal distance

$$x_R = \frac{{w_0}^2}{\varepsilon} \tag{5-13}$$

from the focus. We define this distance as the Rayleigh range for the electron beam.

If we set the Rayleigh range equal to the grating length $Z$ and the waist equal to the scale height as defined in Equation (5-10), then the required normalized brightness is

$$\overline{B}_N \approx \frac{I}{8\pi^2\beta^2\gamma^2\varepsilon^2} = I\left(\frac{x_R}{2\pi\sqrt{2}\beta\gamma {w_0}^2}\right)^2 = \frac{2\pi^2 Z^2 I}{\beta^6\gamma^6\lambda^4} \tag{5-14}$$

For a 10-mm long grating that lases at 300 μm (1 THz) and requires 5 mA start current, the brightness of the electron beam at 50 keV must be

$$\overline{B}_N \approx 10^{11}\frac{\mathrm{A}}{\mathrm{m}^2\cdot\mathrm{ster}} \tag{5-15}$$

This requirement is significant and suggests that SPFELs will require high-brightness electron sources.

## Electron source brightness: thermionic and photoelectric

### *Thermionic sources*

Thermionic emitters are often used in electron microscopy, and the brightness available from them serves as a reference. In a thermionic emitter, the tail of the high-temperature Fermi-Dirac distribution for the electrons extends above the workfunction of the metal and some electrons are directly emitted. We can estimate the thermal emittance from such a source using the Maxwell distribution of particle speeds. The rms velocity in either of the transverse directions (the projection of the velocity in the transverse plane) is

$$v_{rms} = \sqrt{\frac{k_B T}{m_e}} \tag{5-16}$$

where $k_B$ is the Boltzmann constant, $T$ is the electron temperature (here assumed equal to the cathode temperature), and $m_e$ is the electron mass. For uniform emission, the rms position is $\sqrt{\langle r^2 \rangle} = r_c/2$, where $r_c$ is the emitting radius. From these, we can see that the normalized emittance of a beam produced by a thermionic source in either transverse dimension is

$$\varepsilon = \frac{r_c}{2}\sqrt{\frac{k_B T}{m_e \gamma^2 \beta^2 c^2}} \tag{5-17}$$

where $\beta c$ is the longitudinal velocity of the electron beam. The normalized average brightness for a thermionic cathode according to (5-9) is

$$\bar{B}_N = \frac{I m_e c^2}{2\pi^2 r_c^{\,2} k_B T} \tag{5-18}$$

where $I$ is the beam current, and we have approximated the factor $\eta$ by unity.

One standard thermionic emitter is a tungsten hairpin cathode. Drawn tungsten wire around 100 μm in diameter is bent into a point and heated to ~2700 K. These cathodes can produce current densities up to $10^5\ \mathrm{A/m^2}$, and brightnesses up to $10^{12}\ \mathrm{A/m^2 ster}$, but at the temperatures required (~3300K) the cathode lifetime is around 15 minutes due to the evaporation rate [4]. Lanthanum hexaboride is also used as a thermionic emitter. It has a much lower work function than tungsten (2.5 eV, compared to 4.5 eV for tungsten). Somewhat lower total currents are emitted, but from a very small emitting radius (~5 μm) at the end of a pointed rod, resulting in normalized brightnesses that are also as high as $10^{10}\ \mathrm{A/m^2 \cdot ster}$ [4]. In $LaB_6$, this brightness is achieved at 2000 K, where the cathode lifetime is 17 h. Thermionic sources do not provide the beam quality necessary to drive an SPFEL.

*Photoelectric cathodes*

The mechanism for the photoelectric effect in metal cathodes is discussed elsewhere in this thesis. The beams from photocathodes generally are pulsed, with emission pulse length matching that of the drive laser. Semiconductor photocathodes can have very high quantum efficiency (QE), meaning high probability of electron emission for each absorbed photon. The QE can be up to tens of percent for semiconductors due to their high mean free path for carriers. They are often extremely sensitive to surface contamination, however, and therefore difficult to maintain. Metals make more rugged photocathodes, but with a much shorter mean free path for electrons, they have much lower QEs than semiconductors. The QE can be greatly enhanced under high electric field. We have developed needle-type photocathodes, using the geometric enhancement of the electric field at the emitting tip to increase the intrinsically low QE of the metal.

In photocathodes, the thermal emittance is not determined by the cathode temperature, but by the excess energy of the exciting photon above the surface barrier. The normalized emittance for electrons emitted from a circular area of a photocathode surface is

$$\varepsilon = \frac{r_c}{2}\sqrt{\frac{\Delta E}{m_e \gamma^2 \beta^2 c^2}} \tag{5-19}$$

where $\Delta E$ is the photon energy minus the Schottky-lowered workfunction. The normalized average brightness for a photocathode is then

$$\overline{B}_N = \frac{I m_e c^2}{2\pi^2 r_c^{\ 2} \Delta E} \tag{5-20}$$

where, again, we approximate the factor $\eta$ by unity.

For laser spot radii of ~20 μm, the achievable current densities in needle photocathodes are on the order of $10^7 \ \mathrm{A/m^2}$, much higher than thermionic cathodes. Considering $\Delta E \approx 2 \ \mathrm{eV}$, the resulting emittance for a 50 keV beam is $4\times10^{-8}$ m, and the brightness is $10^{11} \ \mathrm{A/m^2 \cdot ster}$, on the order of what is required to drive an SPFEL. To compare with thermionic cathodes, $k_B T \approx 2 \ \mathrm{eV}$ corresponds to a temperature of 20,000 K. Photocathodes give better brightness beams than thermionic cathodes because the achievable current densities are orders of magnitude higher, not because they produce beams with lower effective temperature.

Because they operate at lower temperature, photocathodes do not evaporate like thermionic cathodes. Laser damage can damage the cathode surface, but in general it is contamination of the needle that degrades their performance. The contamination time depends on the vacuum in the chamber. In some cases, we have observed that contamination is removed by laser pulses in the ultraviolet, and this could be very useful for cleaning the needle during operation. Clean, smooth surfaces can also be restored by heating the cathode tips by electron bombardment from a hot tungsten filament. With periodic heat treatment, we have used metal photocathodes several hours a day for a month or more without replacement.

## Other contributions to focal spot size

The emittance and brightness calculations discussed above place limitations on the focal spot size and depth of field, which are both important to SPFEL operation as discussed above. From Equation (5-5), we can calculate the beam spot size resulting from the emittance of the beam. Space-charge forces and lens aberration effects, for example, can change the beam emittance and produce larger beam spot sizes at the focus than predicted from the properties of the beam source. Tracking the evolution of the beam distribution in trace space produced by multiple nonlinear forces can be handled by computer ray-tracing simulations [5]. We can, however, treat these effects in terms of the spot sizes they produce for beams with zero thermal emittance. These spot sizes are then compared to the spot size resulting from the thermal emittance in the absence of the other effects. This approach is particularly useful when one particular effect is dominant, but when several are comparable they can be added in quadrature [1 (p. 226)].

### *Spherical aberration in solenoidal magnetic lenses*

The focusing effect of a solenoidal lens is due to the combined action of the fringe fields and the main solenoidal magnetic field, which runs parallel to the solenoid axis. As the beam enters the lens, an angular momentum is imparted by the fringe fields. Electrons are given a component of velocity perpendicular to the beam axis and then bent toward the beam axis by the main solenoidal field. As the beam exits the lens, the fringe fields act to remove the rotation of the beam, so that the beam leaves the lens field with the same angular momentum with which it entered.

We calculate the focal length of the solenoid from the fringe fields as given by the Maxwell equations. Knowing the axial induction $B_x$ as a function of the position $x$ along the beamline, we calculate the radial field from Gauss's law. If we choose a Gaussian cylinder coaxial with the beam and of very small radius, we assume that the magnetic flux is constant across each cylinder end (the paraxial approximation). Shrinking the length of the cylinder to an infinitesimal, we get

$$B_r(r) = -\frac{1}{2} r \frac{\partial B_x}{\partial x} \tag{5-21}$$

where $B_r(r)$ is the radial magnetic field at the distance $r$ from the axis, and $B_x$ is the axial field at $r = 0$. The radial field gives the electrons an azimuthal velocity

$$v_\phi(x) = \int_{-\infty}^{x} \frac{q_e B_r(x)}{m_e} dx \tag{5-22}$$

Electrons with nonzero azimuthal velocity experience a focusing force and a centrifugal force. The transverse momentum imparted by the solenoid is

$$\Delta p = \frac{q_e^{2} r}{4 m_e v_x} \int_{-\infty}^{\infty} B_x(x)^2 \, dx \tag{5-23}$$

where $q_e$ is the charge on the electron, and $v_x$ is the electron velocity in the longitudinal direction. The focal length of the solenoid lens is defined by

$$\frac{1}{f_0} = \frac{q_e^{2}}{4\gamma^2 m_e^{2} v_x^{2}} \int_{-\infty}^{\infty} B_x(x)^2 \, dx \tag{5-24}$$

where $f_0$ denotes the paraxial focal length. This is a thin-lens approximation as it ignores changes in the radial position of the trajectory inside the lens [6 (p. 99)].

If we enlarge the Gaussian cylinder to look at non-paraxial effects, we must allow for nonuniform flux through the cylinder ends. We allow the axial derivative of the main field to change linearly with radius according to

$$\frac{\partial B_x}{\partial x}(r) = \left.\frac{\partial B_x}{\partial x}\right|_{r=0} + r\frac{\partial}{\partial r}\frac{\partial B_x}{\partial x} \tag{5-25}$$

Combining Ampere's law $(\nabla \times B = 0)$ with Equation (5-21), we get

$$\frac{\partial}{\partial r}\frac{\partial B_x}{\partial x} = -\frac{1}{2}r\frac{\partial^3 B_x}{\partial x^3} \tag{5-26}$$

Equation (5-21) then becomes

$$B_r(x) = -\frac{1}{2}r\frac{\partial B_x}{\partial x} + \frac{1}{8}r^3\frac{\partial^3 B_x}{\partial x^3} \tag{5-27}$$

and the focal length including spherical aberration is

$$\frac{1}{f} = \frac{1}{f_0} + r^2 S \tag{5-28}$$

where $f_0$ is the paraxial focal length defined in Equation (5-24) and

$$S = \frac{{q_e}^2}{8\gamma^2 {m_e}^2 {v_x}^2}\int_{-\infty}^{\infty}\left(\frac{\partial B_x}{\partial x}\right)^2 dx \tag{5-29}$$

$S$ is always positive, so the edge ray (or marginal ray) in any given electron beam is focused by the solenoid with a shorter focal length than the paraxial rays. This creates a curved caustic between the lens and the paraxial focal spot. The smallest spot size occurs not at the paraxial focus, but at the disc of least confusion, shown in Figure 5-1. The disc of least confusion occurs at a point between the marginal and paraxial focus.

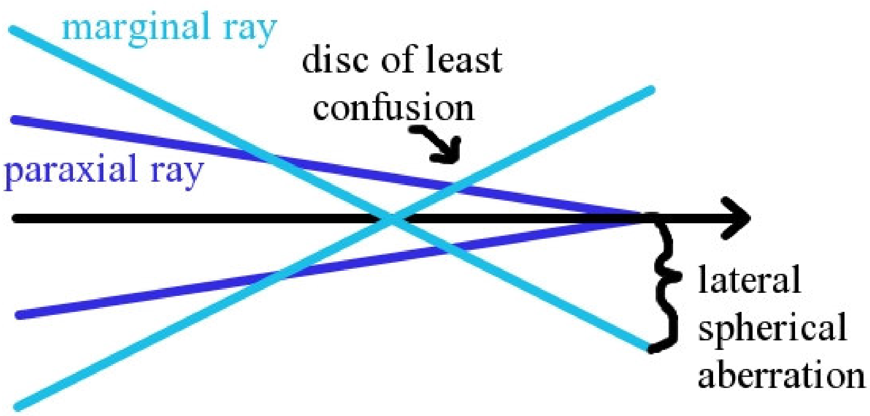

Figure 5-1. Spherical aberration is the focusing of outer rays (marginal rays) with a shorter focal length than those near the axis (paraxial rays). The smallest spot size occurs near the marginal focus at a point called the disc of least confusion.

The ratio of the distance along the axis of the disc of least confusion from the marginal focus to the distance from the paraxial focus is 3:1. The radius of the disc of least confusion is then one-fourth the lateral spherical aberration, defined as the transverse distance of the marginal ray from the axis at the paraxial focus [7 (p. 108)].

We have defined the convergence angle by

$$\alpha = \frac{A}{f} \tag{5-30}$$

where $f$ is the focal length, and $A$ is the diameter of the aperture at the lens. For small convergence angles, the diameter at the disc of least confusion is then

$$d_{SA} = \frac{\alpha}{4}\left(f_0 - f\right) \tag{5-31}$$

or, in terms of $S$ as defined in Equation (5-29),

$$d_{SA} = \frac{1}{4} S f_0 A^3 \tag{5-32}$$

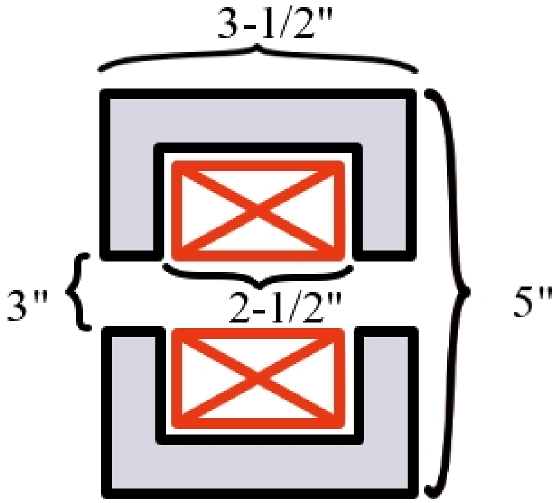

Figure 5-2. Cross-sectional profile of original magnet polepieces with magnet iron in gray and position of current carrying coils marked with red.

$S$ and $f_0$ depend on the current through the solenoid, but their product does not. We can then rewrite Equation (5-32) as

$$d_{SA} = C_{SA} A^3 \tag{5-33}$$

where

$$C_{SA} = \frac{\int_{-\infty}^{\infty} B_x'^2 dx}{8 \int_{-\infty}^{\infty} B_x^2 dx} \tag{5-34}$$

and the prime represents the spatial derivative in the longitudinal direction.

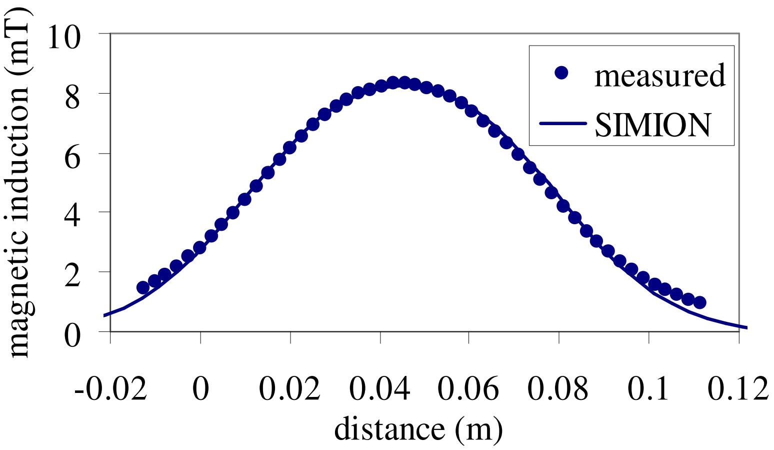

Figure 5-3. Magnetic induction field as measured using original polepieces and portable Hall probe and compared to SIMION simulation. The field in this figure is scaled to represent the field with 1 A current in the solenoid coils.

### *Solenoid field data*

A solenoid was constructed with ~500 coils of copper wire (AWG 18, 1.024 mm diameter). The measured resistance of the coils is 3.4 Ω. The coils are encased in a cylindrically symmetric yoke made of magnet iron, with a gap along the inner face of the cylinder. The dimensions of the original set of magnet pole pieces are shown in Figure 5-2. The on-axis magnetic field was measured with a portable Hall probe, and the polepiece configuration was simulated using SIMION. The field data and the calculation are shown in Figure 5-3. The $y$-axis crossing point on the abscissa represents the edge of the iron polepiece. The full-width half-maximum (FWHM) of the field is 8 cm. The linearity of the on-axis magnetic field with current was confirmed up to 6 A.

A second set of magnetic polepieces was built with a shorter gap (referred to as the “alternate” configuration), and the dimensions of the magnet with these polepieces in

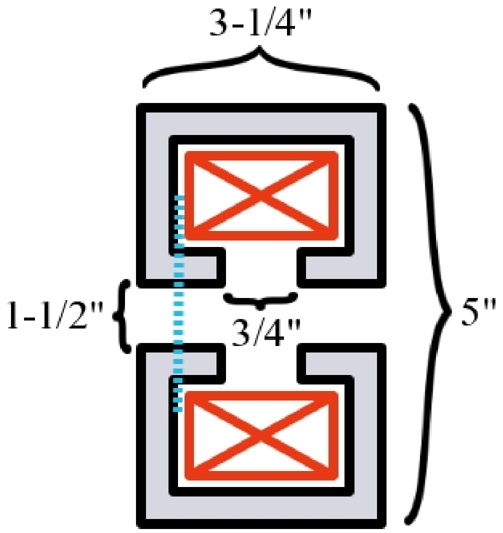

Figure 5-4. Cross-sectional profile of alternate magnet polepieces with magnet iron in gray and position of current carrying coils marked with red. The blue dotted line shows where the polepiece was cut to create the hybrid polepieces.

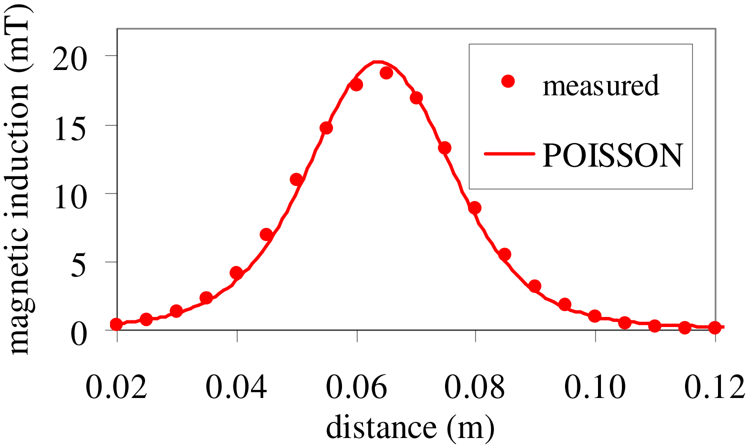

Figure 5-5. Magnetic induction field as measured using alternate polepieces and portable Hall probe and compared to a POISSON calculation by J. D. Jarvis. The field in this figure was measured with 1 A current in the solenoid coils.

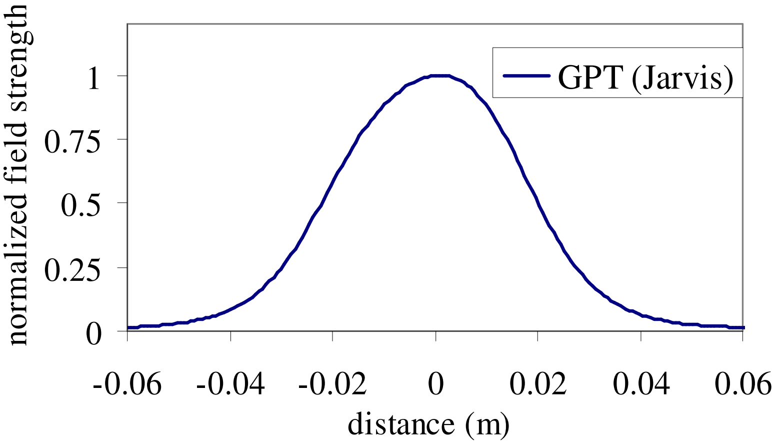

Figure 5-6. Calculated magnetic field strength on-axis in the longitudinal direction for the hybrid magnet polepieces performed by J. D. Jarvis using POISSON. The magnetic induction field is normalized to the peak field.

place are shown in Figure 5-4. The measured values of the field are shown in Figure 5-5 and compared to a POISSON simulation [8] by J. D. Jarvis using the polepiece dimensions. The peak field is multiplied by two compared to the original polepieces and the FWHM is 3 cm. Linearity of the magnetic field with current was confirmed up to 3 A on axis and at a point near the edge of the pole gap.

A third polepiece configuration (the hybrid configuration) was obtained by removing part of the alternate polepieces by cutting along the blue dotted line shown in Figure 5-4. The field for the hybrid configuration was computed by J. D. Jarvis using POISSON and the results are shown in Figure 5-6. The magnetic induction field is normalized to the peak value, but only the shape of the field is needed to calculate the spherical aberration.

The relevant integrals of the magnetic field squared and the derivative of the field squared are given in Table 5-1 along with the spherical aberration coefficient $C_{SA}$ for the

Table 5-1. Focusing properties of the solenoid magnet with two different sets of polepieces. These figures are for 1 A solenoid current and a 50 keV electron beam.

| Solenoid magnet properties | | | |
|---|---|---|---|
| polepieces | $\int B_x^{\,2} dx / I^2 \ \left(\frac{\mathrm{T}^2 \cdot \mathrm{m}}{\mathrm{A}^2}\right)$ | $\int B_x'^2 dx / I^2 \ \left(\frac{\mathrm{T}^2 \cdot \mathrm{m}}{\mathrm{A}^2}\right)$ | $C_{SA}$ $\left(\mathrm{m}^{-2}\right)$ |
| original | $3.8\times10^{-6}$ | $1.9\times10^{-3}$ | 63 |
| alternate | $8.3\times10^{-6}$ | $2.2\times10^{-2}$ | 330 |
| hybrid | – | – | 210 |

three lens configurations. The alternate polepieces have a focal length for a given current and beam energy that is half of that for the original polepieces, but the spherical aberration coefficient is greater by about a factor of 5. The focal strength and spherical aberration coefficient of the hybrid arrangement are intermediate to the original and alternate values.

*Space-charge forces*

The mutual Coulomb repulsion of the electrons within the beam also sets a limit on the spot size that is independent of the emittance. The electric field of the electrons is opposed by the force on the electrons moving in the magnetic field created by the beam, but the outward electric force is always greater in non-neutralized beams. For a uniform electron beam, the net force on an individual electron leads to a transverse acceleration given by [1 (p. 117)]

$$\gamma m_e \ddot{r} = \frac{I q_e^{\,2} r}{2\pi\varepsilon_0 \beta c \gamma^2 R^2} \tag{5-35}$$

where $r$ represents the transverse position of the particle, $I$ is the beam current, $q_e$ is the electron charge, $\varepsilon_0$ is the vacuum permittivity, $\gamma$ is the Lorentz factor of the beam, $R$ is the radius of the entire electron beam, and the dots indicate time derivatives. If we consider the outermost electron, $r = R$, and use the beam velocity to replace the time derivatives with spatial ones, we get the equation for the beam size

$$RR'' = \frac{2I}{\gamma^3 \beta^3 I_0} \tag{5-36}$$

where

$$I_0 = \frac{4\pi\varepsilon_0 m_e c^3}{q_e^2} \tag{5-37}$$

and the primes indicate derivatives with respect to $x$, the longitudinal beam position. Lawson [1 (p. 117)] defines the right hand side of Equation (5-36) as the generalized perveance $K$.

We integrate Equation (5-36) twice and use the generalized perveance to write the result

$$R = x\sqrt{\frac{K}{2}} F\left[\ln\left(R/R_0\right)\right] \tag{5-38}$$

where

$$F\left(x\right) = \left( e^{-x} \int_0^{\sqrt{x}} e^{u^2} du \right)^{-1} \tag{5-39}$$

As also discussed in Lawson, a parallel beam of modest line charge density $\left(I \ll I_0 \beta^3\right)$ focused by a lens with aperture of diameter $A$ comes to a waist at the focus given by

$$d_{SC} = Ae^{-\frac{\alpha^2}{8K}} \tag{5-40}$$

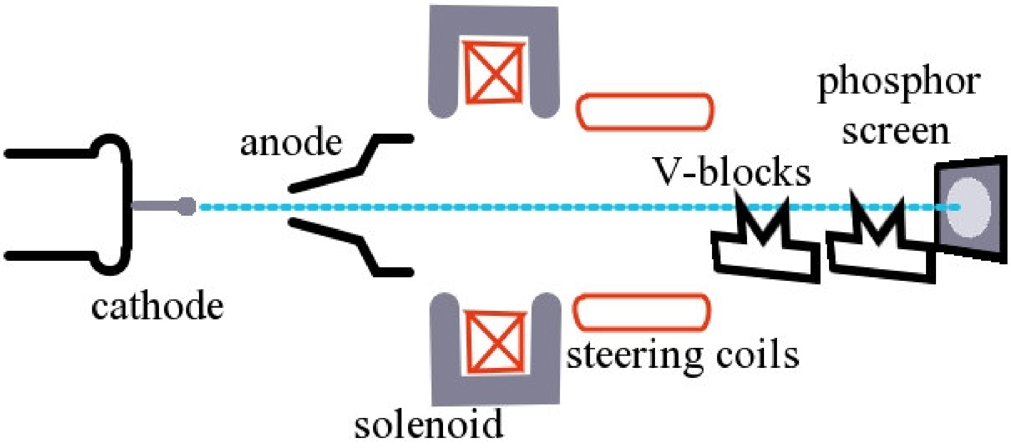

Figure 5-7. Cross-sectional schematic of the electron beam apparatus, with the blue dotted line representing the nominal electron beam path.

where $\alpha = A/f$ and $f$ is the focal length of the lens. $\alpha$ then represents the convergence half-angle of the beam.

## Electron beam apparatus and results

We use the above analysis to interpret measured results from the beamline with a blunt yttrium cathode installed. Data on total current from the cathode are analyzed elsewhere in this thesis. Data on the ratio of beam current through the anode to total needle emitted current are presented, and the effect of space charge on this ratio at high current is discussed. Spot size data are also presented.

*Apparatus alignment*

A schematic of the electron beam setup is pictured in Figure 5-7. The principal optical element in the beam line is the focusing solenoid, shown right after the anode in the figure. In general, spherical aberration in the solenoid is the most significant aberration effect in the focusing system. When the solenoid magnetic axis is not aligned properly (in transverse position and angle) with the beam axis, however, several kinds of aberration become significant. Coma can be observed as well as astigmatic contributions to the spherical aberration effect. Alignment procedures for both transverse position and tilt angle have been developed.

The position of the solenoid in the direction transverse to the beam line is set by tie-down screws that hold a mounting ring in place against the vacuum chamber. To set the position, a LabView [9] program systematically raises and lowers the focusing strength (coil current) of the solenoid and the position is manually adjusted until the centroid of the beam does not move as the focus is changed. For a solenoid that is offset from the beam line by a distance $s$, the difference between the beam centroid position with the lens turned off and the position with the beam focused is

$$d = s\left(1 + \frac{i}{o}\right) \tag{5-41}$$

where $i$ is the distance between the source and the lens and $o$ is the distance between the lens and the focal point of the beam. The ratio $i/o$ is approximately 4 for the electron beam experiments, but the alignment procedure usually involves adjusting the lens by only 10% of its overall focal strength. We estimate that the accuracy in the lens position is therefore of the same order as the detectable movement of the beam centroid on the phosphor screen, which is less than 1 mm. Because of the way the solenoid position

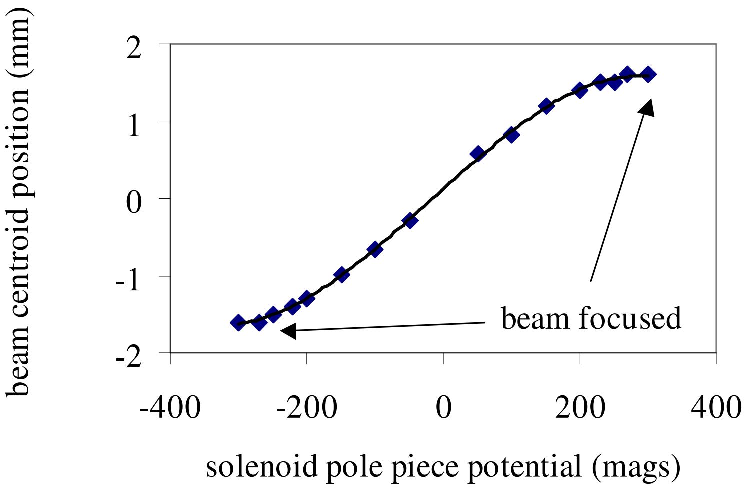

Figure 5-8. Data from SIMION simulation with an electron beam tilted with respect to the solenoid lens. Blue diamonds represent individual simulations and the black curve is a third-order polynomial fit to the results.

adjustment is implemented manually, the smallest practical adjustment is of the same order of magnitude.

The angle made between the solenoid axis and the beam line axis is adjusted independently of the transverse position with three finely threaded screws set into the bronze plate that supports the magnet. SIMION [10] simulations show that a misaligned solenoidal lens leads to a displacement of the beam centroid whose sign changes with the direction of the current in the solenoid. SIMION data for the transverse position of the beam centroid in a fixed observation plane are shown in Figure 5-8. The black arrows in the figure show that the inflection points in the fit curve for centroid displacement occur when the beam is focused in the observation plane. The translational position alignment scheme described in the last paragraph is therefore relatively insensitive to the solenoid tilt. We test the tilt using a different LabView program that periodically switches the direction of the current in the solenoid coils. The tilt alignment screws on the solenoid support plate are adjusted until the beam spot does not move when the direction of the current through the solenoid is reversed. SIMION simulations indicate that a one-degree deflection produces spot motion on current reversal of ~1 mm, and we estimate that the accuracy of the tilt adjustment is within one degree.

### *Electron beam spot size measurements*

Data on the electron beam spot size are taken using the V-shaped blocks before the phosphor screen. Both aluminum V-blocks are mounted on linear actuators that move them perpendicularly to the beam line in the plane of the optical table. They are electrically isolated. The phosphor screen is also electrically isolated and beam current measurements are made there. Comparisons between measurements made between the phosphor screen and an isolated small Faraday cup indicate the phosphor is as good an electron collector to within 5% for electrons at 30-50 keV.

Spot size measurements are made by knife-edge experiments using the V-blocks or by capturing images of the phosphor screen with a CCD camera. The knife-edge experiments themselves are done by either moving the V-blocks and measuring the transmitted current as the blocks occlude the beam, or by using the steering coils to scan the beam across the blocks. Calibration of the steering coils is done by occluding half of the beam with the V-block and measuring the actuator motion necessary to cover half of the beam again after the steering coil current is changed. A sample calibration curve for 40.2 keV electrons is shown in Figure 5-9. Using this method of calibration, the measured spot sizes for spots greater than 100 $\mu$m in diameter taken by the magnetic

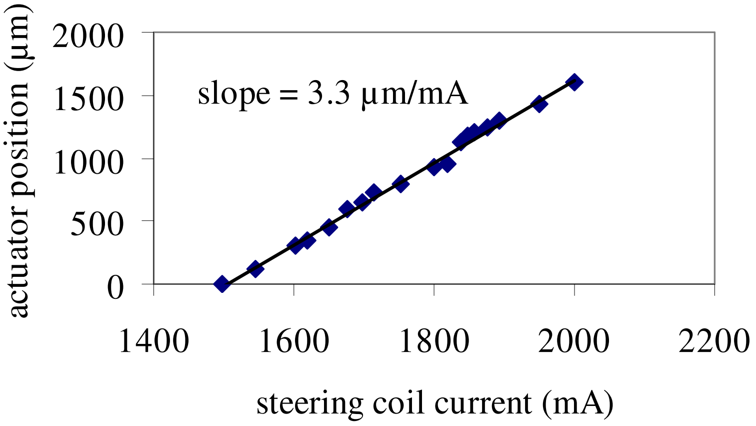

Figure 5-9. Sample steering coil calibration curve for electrons at 40.2 keV. Blue points represent experimental data and the black line is a linear fit with the given slope. Error is estimated at +/- 25 µm and corresponds approximately to the size of the data points.

scanning knife-edge experiments agree closely (<10%) with those made by moving the actuator.

The accuracy in measuring the mechanical motion of the V-blocks with the linear actuator is +/- 25 µm. The programming accuracy of the HP 6624A power supply used to drive the steering coils is ~20 mA, which translates to roughly the same accuracy in displacing the beam (~50 µm). The magnetic steering method is preferred, however, for two reasons: beam position is reproducible (the actuator is only reliable when pushed continuously in the same direction), and the measurements can be automated. LabView programs written to control the steering coils and grab current measurements from the oscilloscope can generate spot-size measurements in less than a minute, compared with two or three minutes (at best) by hand. LabView also randomizes the order of data taking

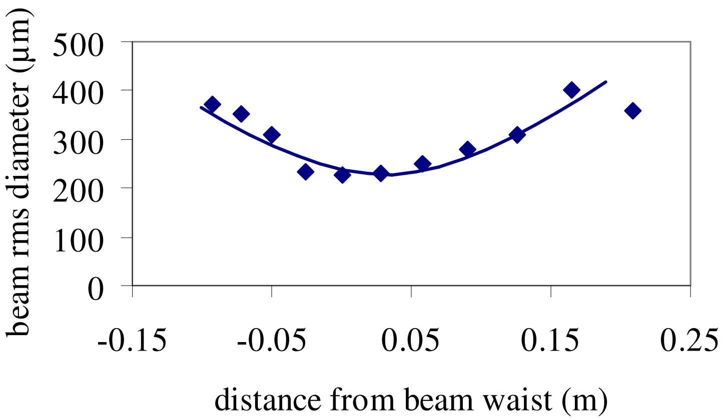

Figure 5-10. Sample solenoid scan of the rms beam diameter near the electron beam focus. Experimental data points from knife-edge experiments are shown as diamonds and a fit to the data using Equation (5-12) is shown on the blue line.

to prevent confusion resulting from drifts in the total current due to slow laser power fluctuation. Randomized data taking is not possible using the actuator.

In both magnetic and mechanical scanning knife-edge experiments, between 8 and 64 current pulses are averaged on the oscilloscope for 10 to 40 positions on the transverse beam profile. Data are fit using Mathcad [11] to an integrated Gaussian with uniform weighting (the data are well fit and tail effects are ignored). The standard deviation of the integrated Gaussian fit is the rms transverse position. The rms diameter, a figure useful for comparison to other experimental data, is $2\sqrt{2}$ times the rms transverse position. A scan of the rms beam diameters is shown in Figure 5-10. The data have been fit using Equation (5-12) with an effective transverse beam emittance of $5\times10^{-7}$ m. Total beam current at the screen for this data set was 1.4 mA.

The knife-edge measurements of the beam size produce transverse spot sizes with much of the nonuniformity in beam intensity washed out by the integration.

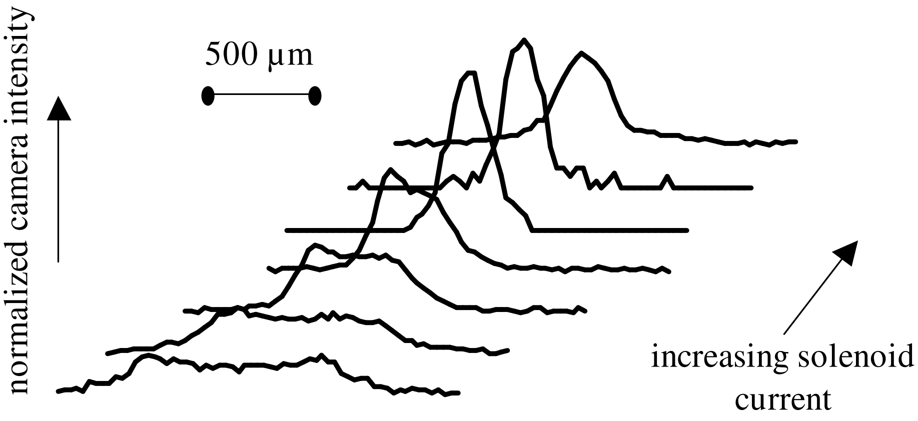

Figure 5-11. Transverse beam profiles obtained from line scans of the CCD image of the phosphor screen.

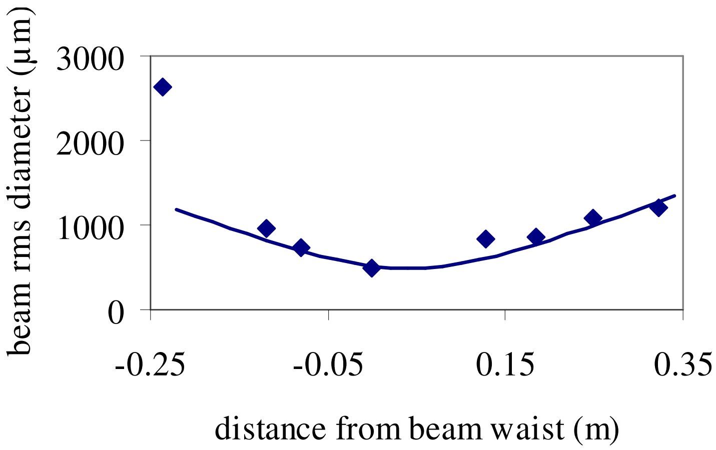

Figure 5-12. Sample solenoid scan of the rms beam diameter for the data shown in Figure 5-11. Experimental data points are shown as blue diamonds and a fit to the data using Equation (5-12) is shown on the blue line.

Alternatively, we consider the image on the phosphor screen as a map of the beam current density. Measurements show that the signal on the monochrome CCD camera is a linear function of the current density at the phosphor screen over the ranges of current density in these experiments (camera gamma is set to unity). The CCD image is captured using LabView image acquisition. The entire map of the current density is available and the transverse width can be calculated from the complete distribution or from line scans. A set of transverse profiles obtained using the CCD camera is shown in Figure 5-11. The measured beam diameters are fit in Figure 5-12, using an effective emittance of $2\times10^{-6}$ m. Total beam current for this data set was 1.0 mA.

The transverse profiles of the electron beam measured from the CCD camera images provide more information about the beam than knife-edge experiments. In Figure 5-11, bright edges in the electron spot are apparent which are washed out by spatial integration in the knife-edge experiments. We observe the bright edges upstream of the electron focus but not downstream of the focus. This is consistent with spherical aberration from the focusing solenoid. Spherical aberration causes marginal rays to be focused with a shorter focal length than paraxial rays. The bright edges in the CCD camera image show the marginal electron trajectories crossing over the paraxial trajectories before reaching the paraxial focus.

### *Space-charge effects at the cathode*

The highest electron density in the electron beam line occurs not at the focus of the beam, but at the source. Since the force on an electron from Equation (5-35) is linear with its transverse position in the beam, space-charge forces do not degrade the beam emittance. To the degree that the beam becomes non-uniform, however, emittance degradation due to space charge occurs at high current densities. In the metal photocathode experiments, space-charge forces are not considered to degrade the

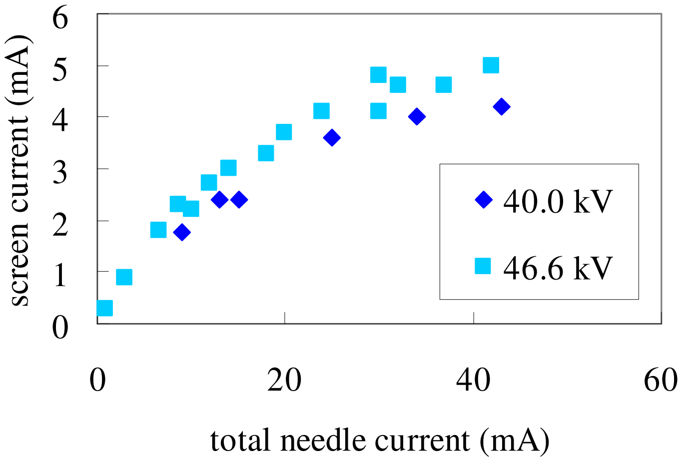

Figure 5-13. Current measured at the phosphor screen as total current is increased (by increasing the laser pulse energy).

emittance significantly, but are shown to cause significant transverse spreading of the beam at the source. This spreading limits the fraction of the beam that passes through the anode at high currents as shown in Figure 5-13. The fraction of the emitted current which enters the solenoid drops from ~50% at very low currents to less than 10% at peak needle currents above 50 mA.

The transverse spreading of the beam at the source due to space charge is somewhat reduced as the beam voltage is increased. The fraction of the beam current transmitted through the anode (measured at the grating) is shown for three different voltages in Figure 5-14. The transmitted fraction is reduced as the total current is increased, but the fraction is higher at higher voltages. Surprisingly, the transmitted fraction seems to scale with the ratio between the total beam current and the longitudinal space-charge limit (as shown in the rescaled Figure 5-15). These data show that the space-charge limit, calculated considering the longitudinal space-charge fields, is relevant

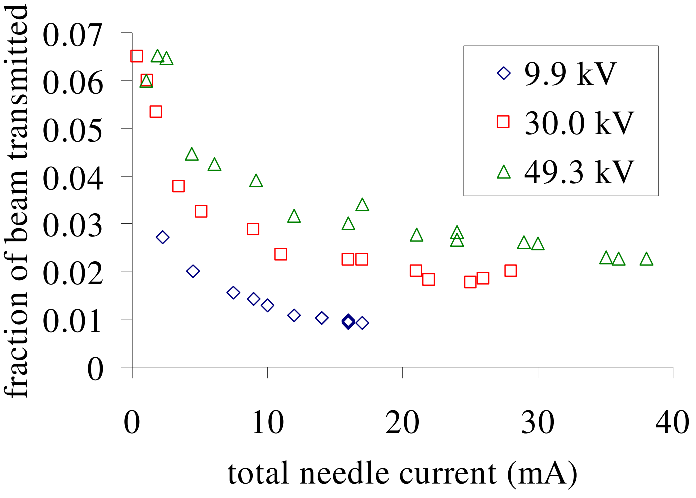

Figure 5-14. Fraction of the electron beam transmitted by the anode as a function of beam current for three different voltages.

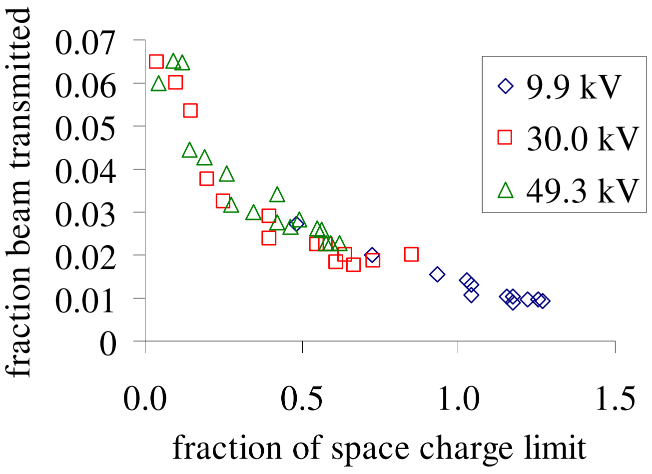

Figure 5-15. Fraction of beam transmitted by the anode as a function of the ratio between the total anode current and the longitudinal space-charge limit.

for estimating the importance of transverse space-charge effects. These effects are very important in needle cathode experiments because of the high current density and low accelerating voltage, but it would be very interesting to apply this scaling law in other environments where beam blowup due to transverse space charge is an issue.

### Interpretation of measured beam waist

The emittances used to fit the beam waist data in Figure 5-10 and Figure 5-12 are much higher than the photoelectric emittance predicted by Equation (5-19). This suggests that the source emittance is not the dominant factor in the evolution of the beam envelope. We can also calculate the predicted beam waist from the photoelectric emittance, and the beam waist produced by spherical aberration and space-charge effects.

From Equation (5-5), we see that the rms diameter of the beam at the focus due to nonzero emittance of the beam is

$$d_E = \frac{8\varepsilon}{\alpha} \tag{5-42}$$

where we have assumed that the emittances in the two transverse directions are equal $\left(\varepsilon = \varepsilon_y = \varepsilon_z\right)$. According to Equation (5-19), the photoelectric emittance of the beam is on the order of $10^{-8}$ m. From observations made with a second phosphor screen inserted directly following the solenoid, we estimate the beam size (the effective aperture) diameter is 12 mm. The distance between the solenoid and the V-block is 45 cm. This corresponds to a convergence angle of $2.6\times10^{-2}$ radians and $d_E = 0.3\ \mu\text{m}$.

The space-charge spot size is given by Equation (5-40) and depends on the convergence angle and beam aperture size, but also on the current through the perveance $K$. For a 1-mA beam, the generalized perveance is $K = 9.5\times10^{-7}$. The space-charge spot size is extremely small, less than 1 nm. It increases exponentially with the current, but even for a 10-mA beam $\left(K = 9.5\times10^{-6}\right)$, the spot-size dominated beam diameter is 1 μm.

The data in Figure 5-10 and Figure 5-12 are both taken with the hybrid set of polepieces, so the spherical aberration coefficient $C_{SA}$ is $210$ m$^{-3}$. The spherical aberration spot diameter is then

$$d_{SA} = 36\ \mu\mathrm{m} \tag{5-43}$$

Spherical aberration appears to be the dominant effect in our experiment. In Figure 5-11, bright edges are clearly observable in the current density profile upstream of the focus, consistent with the marginal trajectories being focused more strongly and crossing over the paraxial rays before reaching the focus. However, the calculated diameter is well below the ~250 μm smallest observed waist in the electron beam experiments.

Several effects could be increasing the effective emittance of the electron beam. Aberration is present to some degree in the steering coils, but experiments using a flexible bellows have not shown that the spot size depends on the steering coil current. Higher-order aberrations in the solenoid could also be significant. The drive laser has a Gaussian profile, not a uniform one, so it is possible that the self-fields of the electron beam at the source produce a transverse expansion of the beam that is not linear in the transverse position, degrading the emittance there. Surface roughness created by repeating melting and freezing of the yttrium metal could be a source of increased emittance [12]. Also, it has been shown that UV laser irradiation on yttrium metal promotes the growth of microstructures on the metal surface [13], and these tiny protrusions could cause a higher source emittance than given by Equation (5-19).

## Summary

In this chapter, we have discussed the electron-beam quality required for an SPFEL and the predicted and measured characteristics of the electron beam produced from a metal photocathode. Electron-beam quality is described in terms of emittance and brightness, the first referring to the size of the electron beam in transverse trace space and second to the electron density in trace space. The strict requirements of the SPFEL are then expressed in terms of beam brightness. The theoretical brightnesses of thermionic and photoelectric cathodes are compared, showing that the thermionic cathode does not produce the required electron beam brightness. Space-charge forces at the beam focus and spherical aberration in the focusing magnet can increase the effective emittance, and therefore the size of the focal spot. For the present electron beam apparatus, spherical aberration coefficients are calculated for various magnet polepieces, and lens aberration seems to be the limiting factor in achieving small spot sizes.

The electron beam apparatus is described. The principal electron-optical element is the focusing solenoid, which is aligned to the beam axis using several simple experiments with the electron beam. Electron-beam spot sizes are measured both by knife-edge experiments and from CCD camera images of the beam on the phosphor screen. These measurements produce spot sizes that are larger than those calculated from the spherical aberration of the solenoidal lens, though spherical aberration is evident in the bright edges of the beam as shown in the CCD camera images.

# References

1 J. D. Lawson, *The Physics of Charged-Particle Beams* (Oxford University Press, New York, 1988).

2 C. A. Brau, in *Physics and Applications of High Brightness Electron Beams*, edited by J. Rosenzweig, G. Travish and L. Serafini (World Scientific Press, Laguna, Sardinia, 2002).

3 J. H. Brownell, J. Walsh, and G. Doucas, Phys. Rev. E **57**, 1075 (1998).

4 A. N. Broers, Journal of Phys. E **2**, 273 (1968).

5 P. Loschialpo, W. Namkung, M. Reiser, et al., Journal of Applied Physics **57**, 10 (1985).

6 M. Reiser, *Theory and Design of Charged Particle Beams* (John Wiley & Sons, Inc., New York, 1994).

7 R. S. Heath, *A Treatise on Geometrical Optics* (C. H. Clay, M. A. and Sons, Cambridge, 1897).

8 J. H. Billen and L. M. Young, *POISSON/SUPERFISH* (2003).

9 N. I. Corporation, *LabView 8.0* (2005).

10 D. A. Dahl, *SIMION 3D Version 6.0* (1995).

11 Mathsoft, *Mathcad 2000 Professional* (1999).

12 Y. Y. Lau, Journal of Applied Physics **61**, 36 (1987).

13 T. Srinivasan-Rao, J. Fischer, and T. Tsang, J. Opt. Soc. Am. B **8**, 294 (1991).

# CHAPTER VI

## CONCLUSIONS AND THESIS SUMMARY

This thesis has explored the Smith-Purcell free-electron laser (SPFEL) as a source of compact, spectrally bright, tunable terahertz (THz) radiation and the development of high-quality electron sources suitable for driving an SPFEL. A new theory of the SPFEL has been created, taking into account the dispersion of evanescent waves on the grating surface. The scale height for the interaction between the evanescent wave and the electron beam places stringent requirements on the quality of the electron beam that are not met by thermionic sources. High-field photocathodes in yttrium and tungsten metal have been made and tested for the purpose of driving an SPFEL, and their performance has been analyzed using a simple model of the photoemission process.

### THz radiation and the SPFEL

There are a variety of ways of making THz radiation, as described in Chapter II, but there remains a lack of compact, spectrally bright, tunable sources. An SPFEL operating in the THz would have great impact on various potential applications for THz light in biology, medicine, and fundamental physics. The THz-SPFEL would be much less expensive and more compact than a conventional free-electron laser, yet still tunable by simply changing the electron accelerating voltage. It would be compact, certainly fitting on a single optical table and perhaps smaller.

The Dartmouth experiments seem to show how to construct an SPFEL and indicate that the electron beam required for lasing is modest: less than 1 mA focused into a 60-μm focal spot, for example. They were able to observe Smith-Purcell (SP) radiation that was nonlinear in the beam current for a variety of gratings and beam spot sizes. We have tried to follow their experiments along two paths: constructing a full theory of the SPFEL to optimize the device parameters, and building a better electron source in the attempt to go beyond their experiments to reach saturation.

We have presented a consistent theory of the SPFEL which takes into account dispersion of the evanescent waves along the surface of the grating that are responsible for bunching of the electron beam. We calculate the wavelength of the evanescent wave whose phase velocity is synchronous with the electron beam and show that it is longer than the longest wavelength in the fundamental SP radiation band. Because the electron bunching occurs at this wavelength, we expect to see coherent SP radiation at integer multiples of the synchronous wavelength.

The Dartmouth experiments showed SP radiation at 90° on the fundamental order that grew nonlinearly with the electron beam current. This radiation is not a harmonic of the evanescent wavelength for their grating as calculated by our theory. In fact, we predict that as the electron bunching grows, fundamental-order radiation at 90° would be suppressed in favor of coherent radiation at harmonics of the evanescent wave wavelength. The Dartmouth group describes the beam-field coupling mechanism as being the same in the spontaneous and the stimulated cases [1], but the SP radiation is not responsible for bunching the electron beam. The bunching is a result of the interaction with the evanescent wave, and therefore should only produce coherent radiation at harmonics of the bunching wavelength.

Our SPFEL experiment was designed to allow us to measure the SP radiation spectrum over a range of angles in the coherent regime. We were not able to measure SP radiation in our bolometer, however, and we cannot resolve experimentally the discrepancy between our theoretical predictions and the Dartmouth results. Another experiment performed with a modified scanning electron microscope (SEM) at the University of Chicago observed a nonlinear signal in their detector, but attributed it to blackbody radiation from their heated grating [2]. The Chicago group found that by using separate power supplies for the lenses and an electrically isolated supply for the Wehnelt, they could achieve 10 mA and beam sizes similar to the Dartmouth group. They observed SP radiation above the blackbody background, but no superradiant emission.

The Dartmouth experiments were performed with less electron beam instrumentation than the experiments at the University of Chicago. Especially if the beam voltage was adjusted with the Wehnelt electrode voltage, it is possible that there were changes in the beam profile as the current was increased. It would be difficult for changes in the beam profile to produce the order of magnitude increase in the SP radiation signal seen by Dartmouth in the nonlinear regime, however. Similarly, though blackbody radiation from a heated grating could have contributed to a bolometer signal in the Dartmouth experiments, it would neither explain the polarization of the radiation nor the power oscillations observed by the Dartmouth group [3].

Though our theory of the SPFEL does not explain the Dartmouth results, it is in good agreement with computer simulations of SPFELs using particle-in-cell codes [4, 5, 6]. The computer simulations clearly show the dominance of the evanescent wave interaction at the wavelength predicted by our theory. Further, the growth rates for the mode in the oscillator regime are in agreement with those calculated by our theory.

## Metal photocathode development

The Dartmouth experiments suggest that it might be possible to reach saturation in an SPFEL with a higher-quality electron beam than the thermionic cathode in the SEM. For this reason, we developed several kinds of metal photocathodes that should produce higher-brightness beams. Needle photocathodes like the ones we have developed are also being considered as electron sources for the next generation of X-ray free-electron lasers [7]. Some of our early experimental results have been analyzed with detailed theoretical models [8], but this thesis demonstrates that we can understand much of the behavior of needle photocathodes with a simple model based on the three-step photoemission theory [9].

Photocathodes are generally characterized by their quantum efficiency (QE), which is the probability that a photon striking the face of the photocathode results in an emitted electron. For flat face cathodes at low field, it is acceptable to characterize the photocathode performance by a constant QE for a given wavelength photon. We have demonstrated experimentally in this thesis that the performance of needle cathodes cannot be described by a single constant QE.

In the three-step emission model we consider the absorption of a photon, the transport of an excited electron to the surface, and the transmission of an electron across the metal-vacuum interface as the important processes in electron photoemission. The absorption depth of the photon depends only on its wavelength and is therefore constant across the surface of the photocathode. The transport of the excited electron from the absorption position to the needle surface depends on the energy of the excited electron above the Fermi level. The energy spectrum of the electrons that can cross the vacuum-metal barrier depends on the surface electric field at the emission point.

Needle photocathodes are more complicated than low-field photocathodes because the surface fields involved result in a significant Schottky lowering of the metal-vacuum barrier. Furthermore, in sharp needles the field (and the barrier lowering effect) changes rapidly across the laser spot. As demonstrated by the experimental data in this thesis, characterizing photoemission from needle cathodes requires considering the effects of both hot-electron transport and Schottky lowering across the surface of the cathode. This requires a model of the electric field variation over the needle surface like the spheroidal-coordinate model used in this thesis.

Describing a high-field cathode by a constant QE leads to incorrect predictions for the change in the emitted current as the laser spot and the shape of the cathode are changed. This thesis has shown that the shape of the cathode critically affects the initial divergence of the beam. The initial divergence is dominated by the needle geometry for sharp needles and by the transverse space-charge expansion for blunt needles. The transverse space charge is important even when the total current is not space-charge limited. Because the initial divergence sets the focusable fraction of the electrons emitted, a model of the photoemission over the surface of the cathode is indispensable for optimizing the electron beam.

## Future work

The SPFEL experiment was designed to reproduce the effects seen at Dartmouth in a pulsed device. The goals for current and spot size were given by the Dartmouth parameters. However, neither our theory nor others that have been proposed describes the Dartmouth results and other experiments have not been able to reproduce them [2]. Our theory indicates that the SPFEL operating in the backward-wave oscillator (BWO) regime should require a higher start current than that observed at Dartmouth. It is possible therefore that Dartmouth saw some type of unexplained enhancement of the SP radiation occurring below the BWO threshold current. In this case, the first SPFEL operating above threshold has yet to be achieved.

The theory developed in this thesis has been borne out by computer simulations of SPFEL oscillators, and experimental confirmation of the theory would be a significant result in terms of the device physics and in terms of THz technology. But if the Dartmouth results are not a benchmark for SPFEL oscillation, we use our theory to provide the goals in terms of start current for a device to function. The metal photocathodes developed are already close to the Dartmouth parameters and should be able to exceed them in total current, but they may not be able to reach the start currents predicted for oscillation by our theory. For that reason, further cathode development in this lab will include field-emitting arrays (FEAs), which, like needle photocathodes, are being considered also for X-ray free-electron laser applications. A FEA-based device could be the first demonstration of the SPFEL in BWO mode, and the experimental test of the theory developed here.

# References

1 J. Urata, M. Goldstein, M. F. Kimmitt, et al., Phys. Rev. Lett. **80**, 516 (1998).

2 O. H. Kapp, Y. Sun, K.-J. Kim, et al., Rev. Sci. Inst. **75**, 4732 (2004).

3 A. Bakhtyari, J. E. Walsh, and J. H. Brownell, Phys. Rev. E **65**, 066503 (2002).

4 D. Li, in *FEL 2006* (www.jacow.org, Berlin, Germany, 2006).

5 J. T. Donohue and J. Gardelle, Phys. Rev. ST-AB **8**, 060702 (2005).

6 J. T. Donohue, and J. Gardelle, in *FEL 2006* (www.jacow.org, Berlin, Germany, 2006).

7 R. Ganter, et al., J. Vac. Sci. Technol. B **24**, 974 (2006).

8 K. L. Jensen, D. W. Feldman, and P. G. O'Shea, J. Vac. Sci. Technol. B **23**, 621 (2004).

9 C. N. Berglund and W. E. Spicer, Phys. Rev. **136**, 1030 (1964).