Emittance and entropy of electron beams

Collin Erickson

May 2, 2014

Abstract

Creating high-brightness electron beams, which have many practical applications, is done with cathodes in regions with large electric field by field emission. The brightness is high when the current is high and the volume of the beam in phase space is small. An estimate of the phase space volume is the rms emittance. The simulations described in this report show that the beam emitted by a gated diamond field emitter of the type fabricated at Vanderbilt has an exquisitely small emittance, on the order of a few nm. This is probably too small to measure. The rms emittance is generally thought to be constant for a beam or increasing in the presence of aberrations in the beam optics, but simulations show that this is not true. The rms emittance rises and falls according to the geometry that surrounds the beam.

The entropy of an electron beam can be calculated, and should be constant if done in six dimensional phase space, and approximately constant in four dimensions. Calculations show that the entropy is not constant in two nor four dimensions. Rather, the entropy changes over time with a shape similar to how the emittance changes, which includes a local maxima at the entrance to an aperture and minima on either side. It is not clear if the fault is with the calculations or theory. To our knowledge, this is the first time that the entropy of a particle beam has been computed and used in a quantitative fashion. Future research should explore the applications and limitations of this concept.

Contents

1 Introduction **2**
- 1.1 Background . . . 2
- 1.2 Field emission . . . 3
- 1.3 Geometry . . . 3

2 Simulations and setup **4**
- 2.1 Geometry . . . 4
- 2.2 CPO methods . . . 6

3 Emittance **7**
3.1 Emittance definition 7
3.2 Calculations done for electrons at a given axial position 8
3.3 Calculations done for particles at a given time 10
3.4 Conclusion 12

4 Entropy **12**
4.1 Entropy definition for discrete particless 13
4.2 Entropy definition for a continuous distribution of particles . . . 13
4.3 Calculations of the entropy in 2-D phase space 14
4.4 Calculations of the entropy in 4-D transverse phase space 23
4.5 Calculations of the entropy in 4-D transverse phase space using polar coordinates 25
4.6 Calculation of the entropy in 6-D phase space 28
4.7 Conclusion 28

5 Conclusion **29**
5.1 Summary of results 29
5.2 Suggestions for future research 31

6 Appendices **32**
6.1 Simulations of gated cathode emittance 32

7 References **37**

1 Introduction

1.1 Background

High-brightness electron beams are an important tool for many applications and have the potential to make revolutionary advances in a broad range of devices. One important use of high-brightness electron beams is in x-ray free-electron lasers (FELs) [1]. Recently, high-brightness electron beams have been proposed for use in high-spectral-brilliance x-ray sources [2]. High-brightness electron beams are also used in scanning transmission electron microscopy to provide resolution down to 50 picometers [4]. Improvements over previous methods used by conventional transmission electron microscopy allow for the same resolution to be obtained with better efficiency. This means that damage from radiation can be minized, allowing wider application of the technique [10]. High-brightness electron beams is an area in which there is a lot of current research and many advancements continue to be made in improving emission and finding new sources. Ultimately, the brightness of an electron beam is limited by quantum mechanics and the Pauli exclusions principle, which restricts the density in phase space. One interesting advance has been the use of carbon nanotubes to create high-brightness electron beams, some of which have approached the quantum limit [3] [6]. However, this limit has never been reached, and the beams

under consideration for this project are far from the quantum limit. Therefore, classical mechanics is used to describe the beam and its propagation.

1.2 Field emission

Field emission results when there is a strong electric field near a charged object, called the cathode, that creates a current of free electrons. It occurs when the electric field is greater than about 10^{10} V/m. When the field reaches this value, the electrons inside the metal are able to tunnel out of the metal into free space. For this reason, cathodes are used with sharp tips which cause a stronger field than would otherwise achievable. High fields cannot be reached unless there is a good vacuum when the voltage is applied.

If the distance between the cathode and anode is large, then we would have to apply a large voltage to reach the desired field. In many cases the required voltage is impractical for the laboratory. To avoid this problem we create cathodes with gates. The gate is much closer to the cathode, so the same field can be achieved with a much lower voltage. This means that we can turn emission on/off by turning on/off a 100 V supply instead of 100 kV, which is much easier to do and control in the lab.

The cathodes that we use are fabricated with diamond tips. Pyramidal depressions a few microns in size are etched into a silicon wafer and the surface is oxidized to a depth of about one micron. The oxidized mold is then filled with diamond by a microwave plasma chemical vapor deposition (MPCVD) process. The diamond is then brazed to a molybdenum substrate and the silicon mold is etched away. A thin metallic layer is then evaporated onto the oxide layer to form the gate electrode. Finally, the gate and oxide layers are locally removed to expose the diamond tip. The molybdenum puck is placed onto the cathode holder. This apparatus is then installed into a vacuum chamber. The vacuum chamber is set up with leads connected to the puck and to the gate. This allows us to set the voltage on the cathode. A short distance away from the cathode is a metal surface, called the anode, that is in the direction of the cathode tip. The anode is also connected to a lead which allows us to set the voltage on it. Setting the voltage of the gate allows us to create a high field region near the cathode tip, which will begin field emission. The emission is controlled by the cathode-gate potential. After the electrons leave the gate region they are accelerated to the anode by the high potential. The anode is fabricated with a small aperture through which the electrons can pass to fall on a fluorescent screen as can be seen in figure 1.

1.3 Geometry

The geometry of field-emission electron beams presents a major challenge because of the scale of the problem. Field emission can only occur when the cathode has a small radius. However, the electron rays must travel a distance that is much larger. For the simulations in this paper, the tip radius is 6 nm but the rays travel 2 mm, a five-order-of-magnitude range of scale size. This

difference in scale makes it difficult for simulations to be accurate in the small region and for the simulation to run in reasonable time with reasonable amounts of computer memory. CPO was chosen for the simulations because it is able to simulate the emission and trajectories of the rays accurately over a long distance.

2 Simulations and setup

2.1 Geometry

The geometry comprises a gated cathode emitting into a Faraday cup, with the face of the cup being the anode. This geometry was originally designed for a research project funded by DARPA; a diagram of the apparatus can be seen in figure 1.

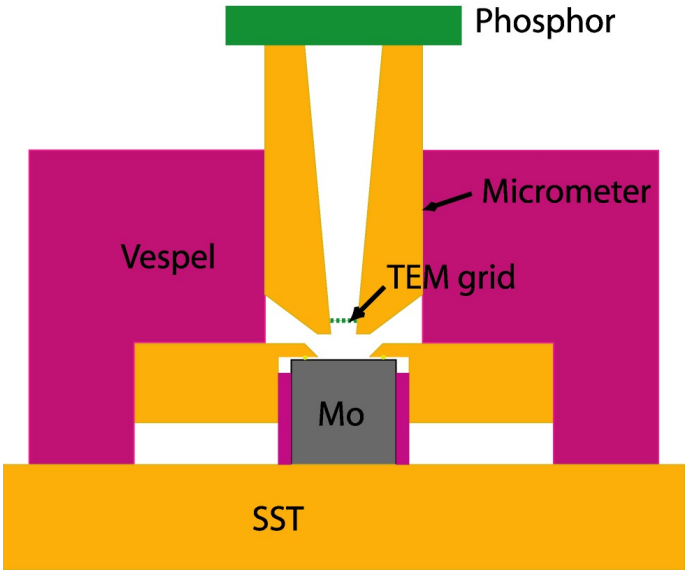


Figure 1: Diagram of the physical object being simulated.

A diagram of the geometry used for the simulations is shown in figure 2. In this picture the cathode is red, the gate is green, the anode is blue, and the rays can be seen traveling from the cathode to the anode.

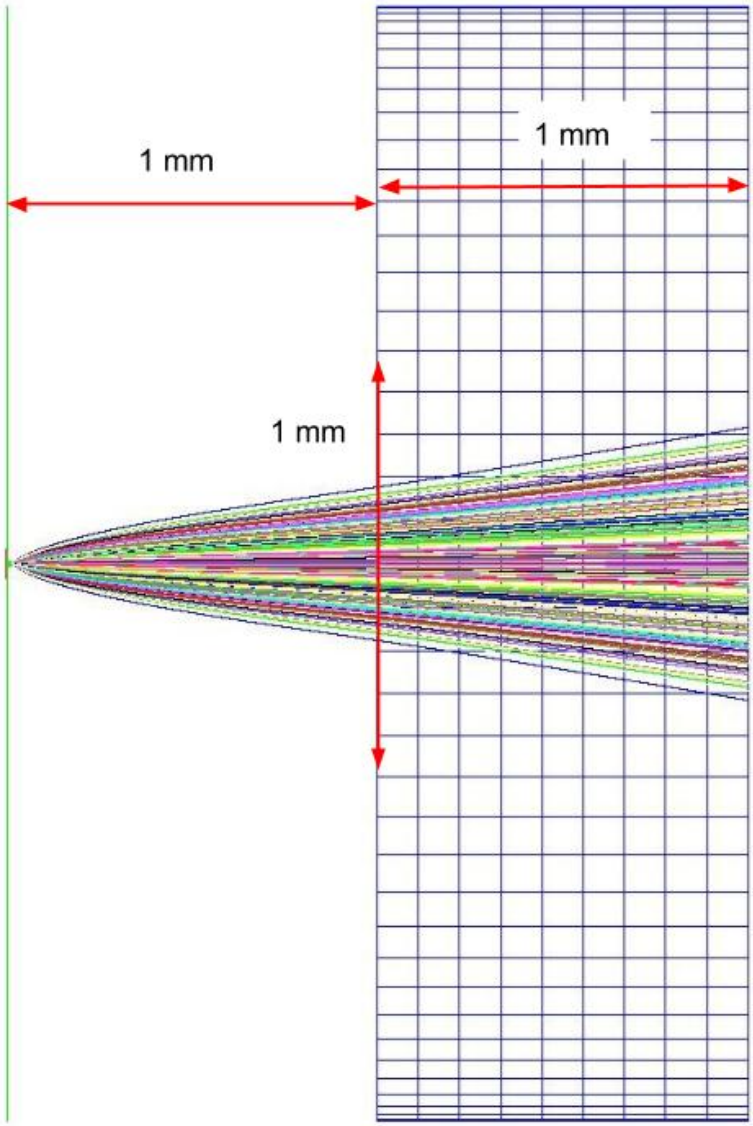


Figure 2: Geometry of simulation.

The cathode is a cone attached to a small base plate on the far left side. The tip radius of the cone is 6 nm, which is typical of those we are able to fabricate in the lab. The cathode is always held at ground potential.

The cathode is surrounded by a gate, as can be seen on the left side of the geometry; figure 3 shows two images of the cathode and gate in more detail. The gate is about 1.5 microns away from the cathode in all places. The gate is

placed at a different voltage than the cathode in order to create a strong electric field at the tip of the cathode. This allows the cathode to field emit according to the Fowler-Nordheim equation. The gate has the very large base plate that can be seen on the left side of the figure. This is necessary to control the field in the region where the rays travel.

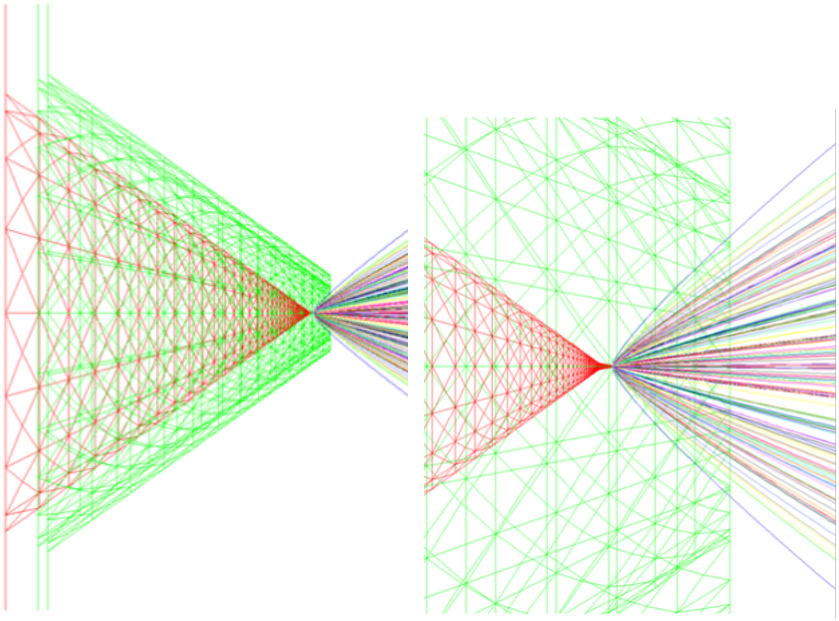


Figure 3: Close-up image of the cathode and gate.

The other part of the system is the anode, which is the Faraday cup. The opening to the cup, called the aperture, is not the entire face, the radius of the opening is half of the radius of the cup. The beam passing through the aperture corresponds to the beam passing through the TEM grid and falling on the screen in the experiment. Due to computer limitations, the simulations only let the rays go 1 mm beyond the aperture, whereas the distance from the TEM grid to the phosphor is much larger as shown in figure 1.

The coordinate system used has z along the direction of travel, which is left to right in the picture. x and y are obviously perpendicular to z. CPO uses symmetry so there is no real difference between x and y. x will be used arbitrarily when only one of the two is necessary.

2.2 CPO methods

The simulation software CPO is used to do these simulations. CPO, which stands for Charged Particle Optics, uses the boundary element method, also

known as the surface charge method, to simulate electron optics. This method is well suited to problems with a large range of scale sizes. Instead of solving Poisson's equation on a grid, in which the grid size must vary over five orders of magnitude in the present case, CPO segments the electrodes and adjusts the surface charge on each segment to satisfy the boundary conditions. In addition, CPO has the Fowler-Nordheim theory for field emission built in. This greatly simplifies the calculations.

To run a simulation on CPO, the geometry and other beginning parameters must be specified. CPO does not simulate electrons individually. Instead it simulates up to 200 rays. A ray is essentially a particle that represents a group of electrons, so each ray has a weighted charge. The emitting cathode is split into a given number of subdivisions, usually between 20 and 200, and each subdivision emits one ray which is tracked by the simulation. With field emission, each ray begins similarly spaced, so the rays where the field is the greatest have the most charge. The ray with the greatest charge has a little less than ten times the charge of the ray with the least charge. The rays used for the calculations in this paper are a composite of four simulations of 48 rays. In total 187 rays are used in the calculations (some did not get emitted). Each simulation starts the rays at the same initial position, the center of each subdivision, so in the beginning there are four rays at each position. However, the rays are emitted with a small random momentum predicted by the Fowler-Nordheim theory. The different transverse momenta cause them to have significantly different positions after only a few time steps.

The first step is to create the geometry, which was done by adapting the setup of a previous simulation carried out by the author, as described in appendix 6.1. The setup must be specified by giving the boundary planes, called electrodes, with sizes, coordinates, voltage, and the number of subdivisions on each electrode. This step is very difficult to do in CPO because the software has many requirements for the electrodes: the size and relative size of the subdivisions on each electrode must be a certain range, neighboring electrodes must have similar subdivision size, etc. Once the geometry is set it is necessary to specify the rest of the parameters for the simulation, such as the step size, the number of rays, etc. The choices for these parameters affect the quality of the simulation and the time it takes for the simulation to run. We have run multiple simulations with varied parameters to determine how our choices of parameters affected the results of the simulation.

3 Emittance

3.1 Emittance definition

In the electron beams under consideration, the interactions between electrons can be ignored. This includes both individual electron-electron collisions and collective space-charge effects, and both are ignored in the simulations. Therefore, the motion of the electrons in the 6-D phase space (x, p_x, y, p_y, z, p_z) is com-

pletely described by the Hamiltonian of (x, p_x, y, p_y, z, p_z). For well-collimated beams, however, it is often sufficient to know the extent of the beam in the 4-D transverse phase space (x, p_x, y, p_y). Since the system has axial symmetry, the disctribution of particles in the (x, p_x) phase plane is the same as that in the (y, p_y) phase plane.

The emittance of an electron beam is a measure of the volume of the beam in phase-space. In other words, emittance measures the amount of variation in the position and momentum of the particles in the beam. Because the electron beam has no sharp boundary, the volume is not well defined. Therefore, the emittance used for our calculations is the normalized rms emittance, which uses momentum instead of direction cosines, because it works when the beam is not paraxial.

The definition of the two-dimensional rms emittance is [8]

$$\epsilon_N^2 = \frac{< x^2 >< p_x^2 > - < xp_x >^2}{m^2c^2}$$

According to Liouville's theorem, the volume of the beam in 6-D phase space should be constant for a Hamiltonian system. However, the rms emittance is not actually the volume and thus need not be conserved. The emittance is similar to the size of the beam in phase space, so the emittance should be approximately constant for our simulations.

Depending on the application or code being used, the emittance can be computed by considering the position and momentum of the electrons when they are all at the same z position or at a certain time. Experimentally the emittance is measured using a screen or similar device as shown in figure 1, meaning that the emittance is measured at a given axial position. Liouville's theorem only applies to particles at a given time, so it does not apply when the calculation is done this way.

3.2 Calculations done for electrons at a given axial position

In an electron-optical system with linear focusing forces, the rms emittance is preserved. However, when the optical system has nonlinear focusing, such as spherical aberration, the phase area of the beam is distorted in a way that generally increases the rms emittance. While the actual phase area of the beam remains constant, as mandated by Liouville's theorem, the effective area is increased by the distortion. This effect is called filamentation because filaments, like arms in a galaxy, can form that increase the effective beam size without actually changing the area [11].

The emittance computated in our simulations of a simple dc beam is not found to be generally constant as expected. Even more surprising, while the emittance increases over the course of the simulation, there is a slight dip in the emittance where the beam passes through the aperture. This same general shape is found for simulations with different number of rays and step size as can be seen in figure 4.

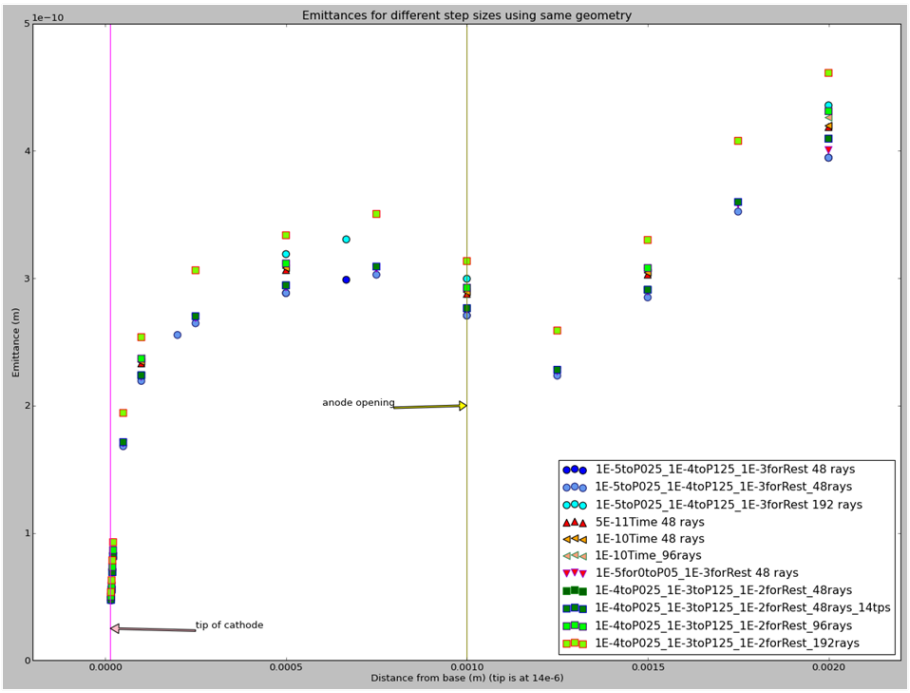


Figure 4: Emittances for different simulations with the same geometry.

At first we thought that this result might be an error in the trajectory simulations or the subsequent numerical analysis. After double-checking the simlulations, calculations, and providing an analytical calculation we came to accept the results. The analytical calculation is an estimate of the variation of the emittance at the point where the beam enters the anode region. To obtain an analytical result, the field is expanded in a Taylor series near the axis and the paraxial approximation is used. This simple model shows that the emittance should decrease briefly, and it is within a factor of 3 of what the simulations predict, which is an affirming result.

It is much easier to get the output from CPO when the rays are at a certain z position than after a certain amount of time has passed. For this reason the emittance is calculated using the position and velocities of the rays when they all have a certain z position. Emittance is often done this way, but Liouville's theorem only applies when the positions and velocities of rays are used after a certain amount of time has elapsed. Normally these values are very similar because a group of rays emitted at the same time tend to be at a similar z position at any given time.

However, after looking at the position of the rays at specific times we find that there is a significant difference between considering the rays at times and positions. This can be seen in figure 5. The output CPO simulation shows a block dot where each ray is after 5 ns intervals during the simulation. The

curvature of the beam means that the central rays have a greater z coordinate than other rays, so the assumption that the calculations could be done with the position of rays at specific times produces different results.

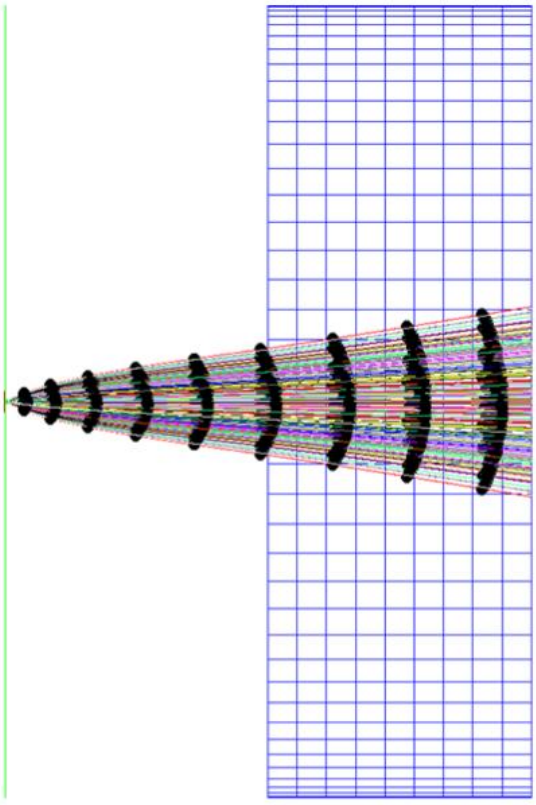


Figure 5: Trajectories of rays for a single simulation with markings after equal time intervals.

After finding this, the methods of doing calculations changed to doing the calculations with the position and velocities of the rays at specific times instead of z positions.

3.3 Calculations done for particles at a given time

Drastically different results for the emittance are found using the new calculation method. The figures below show the emittance over the course of the simulation when calculated as a function of time. Figure 6 shows the emittance as a function of time, while figure 7 shows the emittance as a function of the z position of the furthest ray at the time when the emittance is calculated; this allows us to get an idea of where the rays are relative to the geometry, keeping in mind that

the outer rays are further behind the central ray. The red line in 7 shows the location of the aperture.

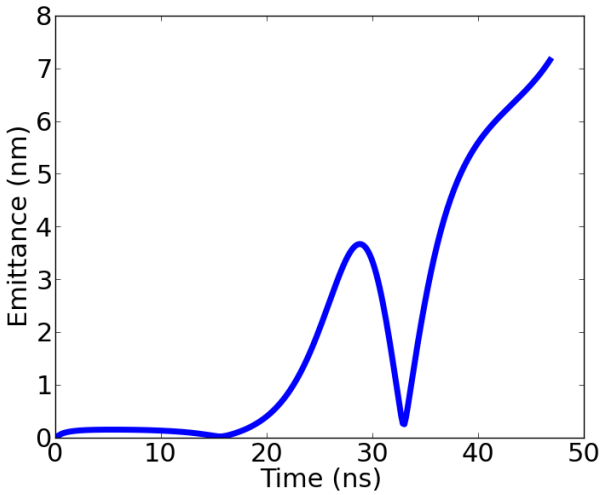


Figure 6: Emittance over the course of a simulation as a function of time.

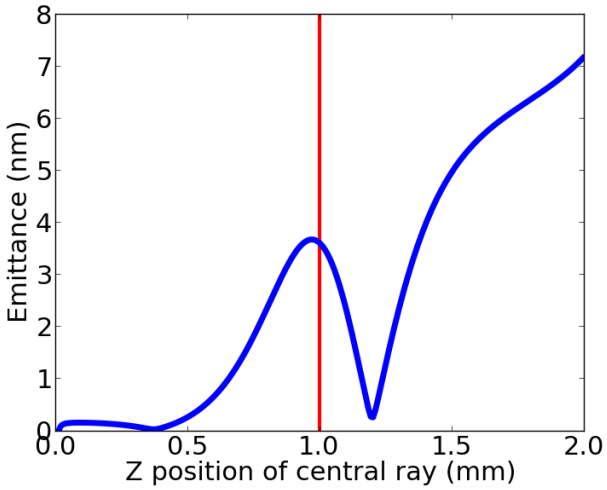


Figure 7: Emittance over the course of a simulation as a function of time, but plotted by the Z-coordinate of the central ray. The red line shows where the aperture is located.

The emittance grows and then decreases slightly in the beginning, followed by a larger growth and a decline back to near its previous level, and finishing

with a large increase again. The first minimum is interesting because it is not exhibited by the emitance computed for the particles at a fixed z position. The calculations done by axial position had a simlar shape of increasing as the rays approach the aperture, decreasing at the aperture, and then increasing to the end. However, the decrease is significantly larger in this case, reducing the emittance to a fraction of its previous maximum. There is also a large change in the maximum value of the emittance compared to the previous calculation: the emittance in this case is about ten times larger. This value is still very small, and is much lower than expected for other types of cathodes.

The main reason that the shape of the emittance plot is so different when the calculation is done by time is that the rays experience much different fields when considered this way. The rays near the center experience the difference in the electric field before the outer rays because they are further along at any given time. This is what causes the emittance the grow significantly the first time. The emittance decreases back to its previous level once the central rays reach the anode opening and the outer rays experience the different electric field. The emittance then grows again when the central rays enter the anode aperture and spread out.

3.4 Conclusion

The emittance of an electron beam is a measure of the volume the beam occupies in phase space. An estimate of this value is the rms emittance. The simulations show that the beam emitted by a field-emitter tip is exquisitely small, on the order of a few nm.

The emittance of an electron beam simulation done using the positions and momenta as as a function of time rather than of axial position provide drastically different results. The emittance is an order of magnitude larger when done as a function of time, but the value is still very small. Over the course of the simulation the emittance varies greatly with a sharp decrease near the center of the geometry, where the rays enter the Faraday cup, followed by a sharp increase back to its previous level after entering the field-free region.

Although the analytical model and tests that were done to check the computer simulations both support the simulation results, specifically the dip in the emittance, it would be useful to have another way check the accuracy of the simulation. One way to do this is to compute another quantity that, unlike the rms emittance, is known to be conserved.

4 Entropy

One way to check this result for emittance is to calculate the entropy for the simulation. While the rms emittance is not constant for Hamiltonian systems, the entropy is. If the calculation for the entropy is not constant, then there must be a problem with the trajectory simulation or with the subsequent calculations of emittance or entropy of the beam. Otherwise if the entropy is constant then

it gives credence to the claim about the change in emittance. This calculation of entropy to give a check to the emittance calculation is explored in the next section.

Entropy is an important measure in physics that can reveal a lot about a system of particles. In statistical physics the single most crucial measure is entropy which in general measures the disorder of a system. In this section entropy is introduced in the context of an electron beam with a finite number of rays. It will provide a check on the emittance calculations of the previous section and will hopefully reveal more about the electron beam.

4.1 Entropy definition for discrete particless

The concept of entropy for beams was first introduced by Lawson et al. in 1976 [7]. For N particles, the state of the system is completely described by the distribution of the particles in 6-D phase space. First the phase space is partitioned along each axis. This creates boxes with volume ΔA, and every electron is in one of these boxes. The number of electrons in the i^{th} box is n_i, with $N = \sum_i n_i$ being the total number of electrons. The multiplicity, or the number of ways the electrons can be placed in these boxes without changing n_i, is $W = \frac{N!}{n_1! n_2! \cdots}$. Then, as usual, the entropy is $S = k \ln W$. This definition of entropy obviously depends on the box size. If the box sizes are large, then all of the electrons are in the same box and the entropy is zero. At the other extreme, if the box sizes are small, then all of the electrons are in their own box and the entropy is $S = k \ln N!$.

4.2 Entropy definition for a continuous distribution of particles

This notion of entropy can be extended to continuous distributions of electrons. For large values we can use Stirling's approximation

$$\ln N! \approx N \ln N - N$$

Then the entropy becomes

$$S = k(N \ln N - N - \sum_i (n_i \ln n_i - n_i))$$

If we let the normalized density of the continuous distribution be $\rho_i = \rho(x_i, p_{x_i})$, then

$$n_i = N \rho_i \Delta A$$

In thermodynamics, the entropy is defined only to within a constant. In the present context, the arbitrary constant is $\ln \Delta A$. In addition, the entropy is an extrinsic variable, proportional to the size system, N. To avoid the dependence of entropy, S, on ΔA and N, the normalized entropy, S_0, is used.

The normalized entropy is written as [7]

$$S_0 = \frac{S}{kN} - \ln \frac{N}{\Delta A} = -\int \rho \ln \rho dA$$

As shown by Struckmeier, the entropy of a bunch of electrons is a strictly conserved quantity for non-interacting particles [9]. This means that the entropy might be used as a test of the accuracy of the computer simulation of the propagation of a particle beam. However, in this context there are problems trying to use entropy this way. First of all, the simulations launch all of the particles at a single instant in time. For a DC beam, the bunch is inifitely long. Short sections of the beam can be considered as a bunch, but the simulations only give an infinitely short bunch. This means that the 6-D density and the entropy are not defined. One way around this is to find how the infinitely short bunch can be related to a continuous beam. In fact, since the physical beam is time-independent, a finite bunch can be simulated by following the infinitely short bunch over some interval of time. It can be shown that in the limit of a very short interval, the 6-D density is related to the 5-D density by $\rho_6 \rightarrow \rho_5/v_z$, where v_z is the axial velocity. Another method that can be used is to consider the entropy of the beam in 4-D transverse phase space (x, p_x, y, p_y), ignoring the axial dimensions (z, p_z). In this reduced phase space the density is the projection of the 6-D density into the transverse space. If all the particles at the point (x, p_x, y, p_y) in transverse phase space have the same axial position and momentum at time t, then they all obey the same Hamiltonian $\mathcal{H}(x, p_x, y, p_y, t)$, where t represents the coordinates $z(t)$ and $p_z(t)$ in $\mathcal{H}$. In this case Liouville's theorem holds, and the 4-D entropy is conserved. The spread of the bunch in the z and p_z directions is small because the energy spread of the beam is very small at the cathode. Thus, the errors in the conservation of 4-D entropy should be small. For this reason, this approach is adopted for these simulations. The entropy of the particles in the 2-D phase space (x, p_x) is also computed.

The second problem is that the entropy is defined as a function of a continuous distribution, but the simulations use discrete rays. These discrete particles must be converted to a continuous distribution to do the calculations. This is a problem because the simulations used less than 200 particles. This can create significant errors, especially in four or more dimensions.

4.3 Calculations of the entropy in 2-D phase space

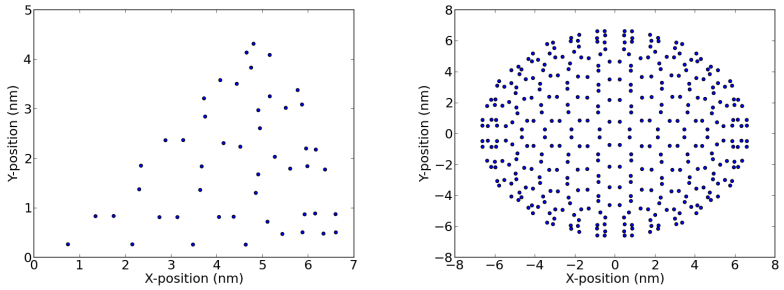
The simulations used for calculations comprised less than 200 rays, which is not a continuous distribution as is required for entropy calculations. The rays are converted into a continuous distribution by spreading out each ray using exponential distributions. The n^{th} ray, located at x_n and v_{x_n} is smoothed out to a Gaussian of the form $e^{-(x-x_n)^2-(v_x-v_{xn})^2}$. The integral over all space of this smooth distribution is normalized to 1. Since the rays in CPO are weighted, this is done by multiplying by the charge of the i^{th} ray, q_i, and dividing by the

total charge, $q = \sum_i q_i$. Thus, the charge distribution for the i^{th} ray is

$$\rho_i(x, p_x) = \frac{1}{2\pi\sigma_x\sigma_{p_x}} \frac{q_i}{q} e^{-\frac{1}{2}(\frac{x-x_i}{\sigma_x})^2 - \frac{1}{2}(\frac{p_x - p_{x_i}}{\sigma_{p_x}})^2}$$

Then ρ, the charge distribution of the entire beam, is obtained by summing over each exponential, of which there is one for each ray.

Before the calculations of entropy are performed, there are some changes that must be made to the input data. The first problem is that the CPO simulation uses symmetry and only simulates rays on one octant of the cathode. However, after spreading out, the exponentials extend into neighboring octants. Due to the smearing, the charge density near the center has more contributions from rays in other octants not present in the simulation, while those further from the center and axes are less affected. This can be corrected for by reflecting the rays to occupy all octants. Figure 8a shows the initial emission, and figure 8b shows the emission after the reflection. The tiling that can be seen disappears very shortly after the rays are emitted, so that shouldn't cause any problems. This reflection approximates what it would be like to have to whole simulation done. However, symmetry can be used to reduce the integration effort. Since negative and positive x-values gives the same result, we only need to integrate over positive x-values.



(a) Initial Emission in once octant (b) Initial emission after reflections

Figure 8: Initial emission before and after reflection

Another difficulty lies in choosing σ_x and σ_{p_x}. The purpose of smearing the rays to form a continuous distribution. If the rays are smeared only a very small amount, then the Gaussians do not overlap, so the smeared distribution do not represent the desired kind of continuous distribution. If the rays are smeared by a large amount, then the Gaussians all overlap and spread far beyond where the rays are located, which also does not give the desired kind of continuous distribution.

To determine σ_x and σ_{p_x}, the rms values $\sqrt{< x^2 >}$ and $\sqrt{< p_x^2 >}$ are used as representatives of the size of the distribution in each direction. We use these

values to get σ_x and σ_{p_x} by relating them with a smearing constant, k.

$$\sigma_x^2 = \frac{< x^2 >}{k^2}$$

$$\sigma_{p_x}^2 = \frac{< p_x^2 >}{k^2}$$

Ideally, the smearing lengths σ_x and σ_{p_x} should be small compared with the size of the distribution of ρ in phase space in order to preserve the details of the distribution. Thus we require that $k >> 1$. On the other hand, the smearing lengths must be large compared with the distances between particles to provide a smooth distribution function. This is equivalent to requiring that multiple particles should be found within the smeared region of one particle, meaning that the Gaussian distributions representing neighboring rays should overlap. With N particles, the density in phase space is represented by $N/\sqrt{< x^2 >< p_x^2 >}$. Thus the second requirement is that

$$\frac{N\sigma_x\sigma_{p_x}}{\sqrt{< x^2 >< p_x^2 >}} >> 1$$

which is equivalent to

$$\frac{N}{k^2} >> 1$$

Together these inequalities give

$$N >> k^2 >> 1$$

The geometric mean of these requirements suggests choosing $k \approx \sqrt[4]{N}$. The data being used has 187 rays, so $k \approx \sqrt[4]{187} \approx 4$. This is not much greater than one. While these rough calculations do not make it clear what the optimum value of k is, it is clear that even in two dimensions the number of particles is unfavorably small. In 4-D phase space these same arguments lead to the requirement that

$$N >> k^4 >> 1$$

This suggests $k \approx \sqrt[8]{N} = \sqrt[8]{187} \approx 2$, which is even less satisfactory.

Figure 9 shows a typical 2-D phase-space plot for a simulation. The points nearest the origin represent rays that started nearest the center of the tip, and these rays have more charge than rays that started near the edge of the tip. This means that the calculations are dominated by these rays. Far from the tip the trajectories approach radii originating from a point slightly behind the tip. Therefore, the particles become very spread out along a diagonal line in phase space because the particles further from the center have greater transverse velocity. If the calculations are done using these data, the result is skewed because there is little overlap of the exponentials used to smooth out the distribution. For this reason, the phase space is transformed before performing the calculations.

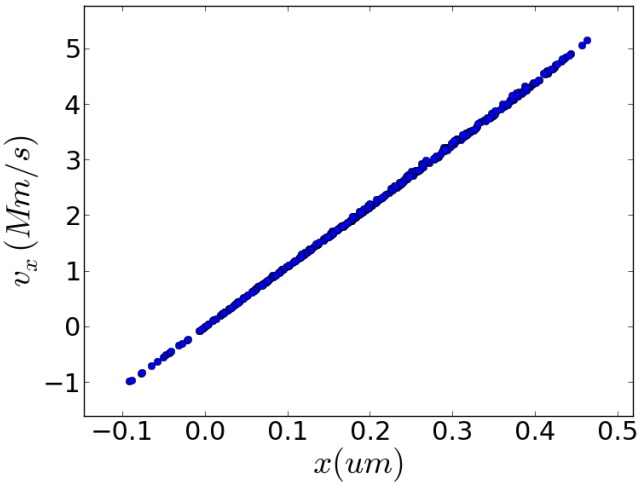
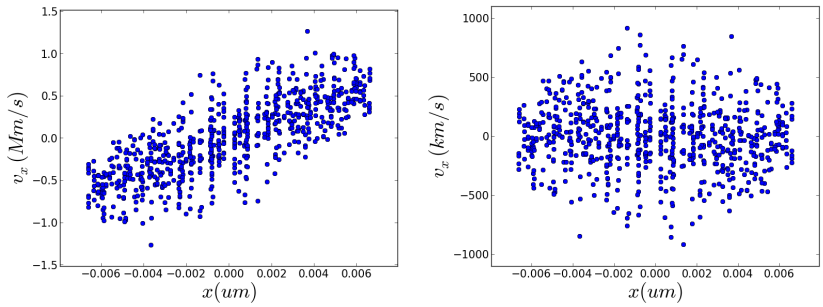


Figure 9: Typical phase space

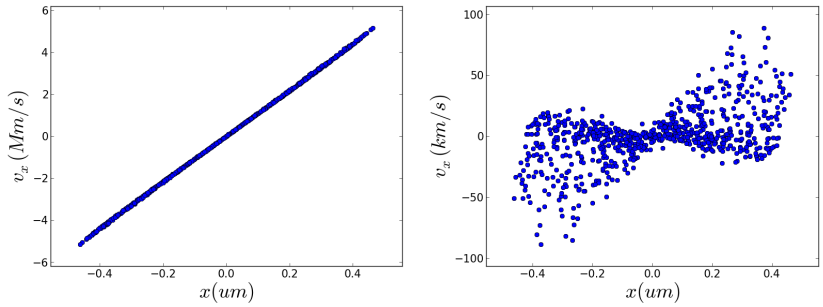
Using a smooth transformation $(x, p_x) \to (x, p_x - f(x))$ on the phase space data does not affect the result of the calculation because the transformation does not change the density of the rays in phase space. The transformation changes the momentum coordinates, not the position coordinates. This is equivalent to having the beam collimated by using a focusing lens. It is hard to tell from the figures, but the phase space of a beam is generally straight with a curvature near the ends. The curvature means that a linear transformation will leave a significant curvature after the transformation. Rotational symmetry means that the graph of the phase space is odd. Thus a transformation of an even power, specifically a quadratic transformation, will be no better than the odd transformation of one power less. For these reasons a cubic regression is used on the data before the entropy calculations are done. The weighted cubic regression is performed on the plot of p_x as a function of x, and the new p_x values are obtained by subtracting the value of the regression line at its x value.

Figures 10, 11, and 12 show the phase space before and after the transformation at three different times. There is a lot of variation in the phase space at time 0, so the phase space before and after the transformation does have much structure, as can be seen in figure 10. At the end of the simulation, which is at time 46.8 ns, the phase space looks very linear; the cubic regression brings the phase space to a bow tie-type shape, as shown in figure 12. In this bow tie shape, the rays are highly concentrated along the axis with outliers extending both upward and downward. After 0.1 ns has elapsed in the simulation, which is about $\frac{1}{500}^{th}$ of the way through the simulation, the phase space and the cubic transformation look nearly identical to that at the end of the simulation, as shown in 11. This means that there is little change to the general shape of the phase space throughout the simulation after the very beginning.



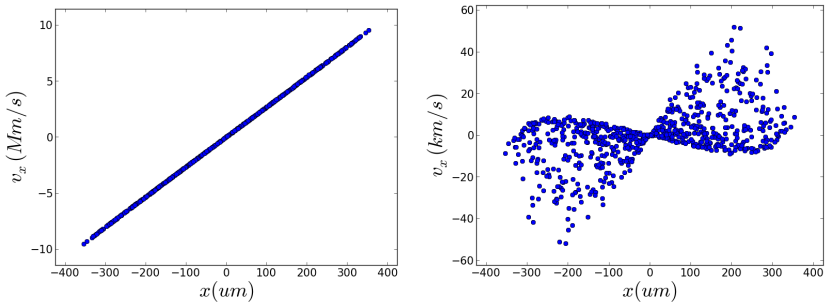
(a) Original (b) Cubic

Figure 10: Phase space before and after a transformation is applied at time 0.



(a) Original (b) Cubic

Figure 11: Phase space before and after a transformation is applied at time 0.1 ns.



(a) Original (b) Cubic

Figure 12: Phase space before and after a transformation is applied at time 46.8 ns.

The graph of ρ, the density of the charge using exponentials after the cubic transformation, 20 ns into the simulation is shown from two angles in figure 13. From this we can infer that the exponentials are most affected by the rays that started near the center of the tip because they have the greatest charge, and that the rays that started far from the center of the tip have less charge.

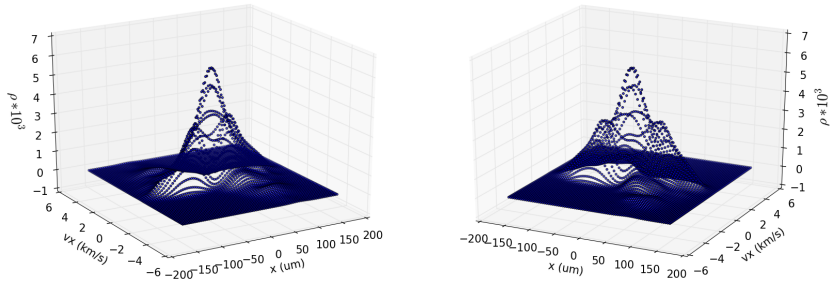


Figure 13: Plots of ρ from two angles 20 ns into the simulation.

More important is the function that will be integrated, $-\rho \ln \rho$. Figure 14 shows what this function looks like from two angles. It appears smooth enough to be easily integrable. The graph of $-\rho \ln \rho$ looks very similar to the graph of ρ. This is because the values of ρ are all small: ρ is less than approximately .25 for the entire simulation, with the maximum at time 0 near .25, and the maximum later on less than .01. The plot of the function $f(x) = -x * \ln x$ appears very linear for $x < .25$, especially for $x < .01$.

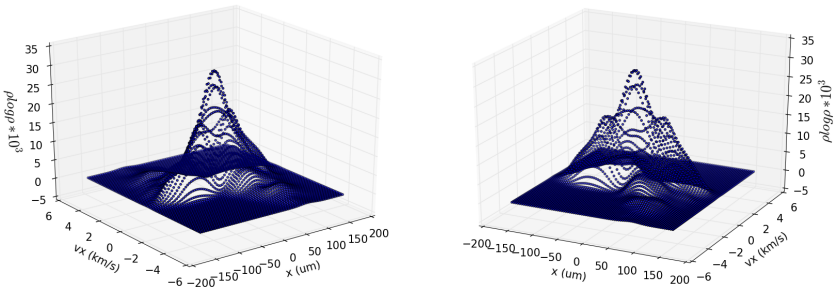


Figure 14: Plots of $-\rho \ln \rho$ from two angles 20 ns into the simulation.

This integral for S_0 is approximated by a Riemann sum because it is found to be accurate to within one part in a million compared to the more sophisticated computer algorithms when enough subintervals are used. The function to be integrated is very smooth so this makes sense that it would give a close answer.

Figure 15 shows the value for S_0 when varying number of subintervals, $M = 15, 30, 60, 100$, and smearing constants, $k = 2, 4, 6$, are used. The results show that there is little difference when the number of subintervals is increased from 15 to 100, especially when comparing 30, 60, and 100 subintervals. This means that the calculation is stable and the number of subintervals used will not have a significant effect on the calculations as long as it is not too small. The major difference in the values of S_0 calculated is due to changing the smearing constant. A larger smearing constant gives a smaller value of the entropy.

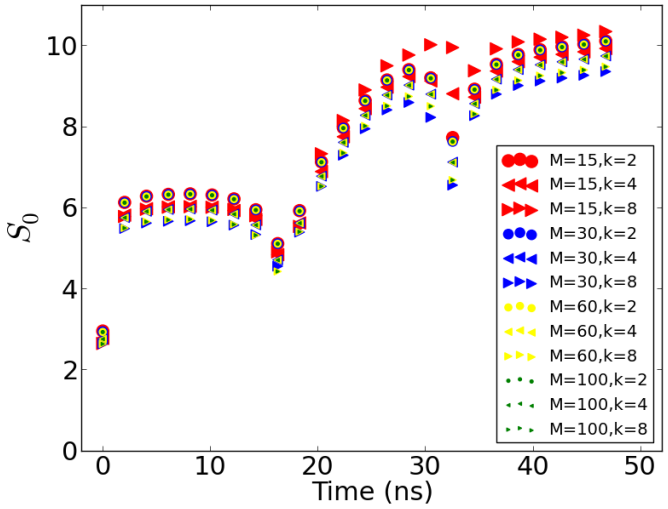


Figure 15: Comparison of S_0 using different number of subintervals and smearing constant.

The fact that the values of S_0 appear to converge to values when increasing M gives confidence that the values are near the true values, but we can also check the integration by seeing what the integral of ρ is when using the same integration technique. Effectively we are checking the normalization of ρ, which should be 1 at all times for all values of k. Figure 16 shows these normalization values. The normalization values are terrible for $M = 15$, and are worse for the larger values of k. However, the normalization for $M = 60$ and $M = 100$ is impeccable, so we have reason to believe that the integration for S_0 is approximately correct for these values.

Ideally the number of subintervals along each axis should be chosen so that the step size is less than the smear of ρ in that direction. For most of the simulation, $\sqrt{< x^2 >}$ is a little less than $k = 4$ at 16 points. It is hard to tell with so few points, but S_0 appears to have a similar shape to the 2-D case: S_0 is generally increasing over the simulation, but there are two local minima near the middle. The graph leaves much to be desired, but we can be confident in concluding that S_0 is not constant over the entire simulation. Half of $\max(x)$, and $\sqrt{< p_x^2 >}$ is about one tenth of $\max(p_x)$. This means that for $k = 4$, σ_x

is about one sixteenth of the range of x and σ_{p_x} is about one eightieth of the range of p_x. Using different values for M in each direction would make the integration more efficient, but this is not implemented for these calculations. Thus, $M = 100$ should give a decent result while the smaller values for M probably miss some of the exponentials completely in the integration.

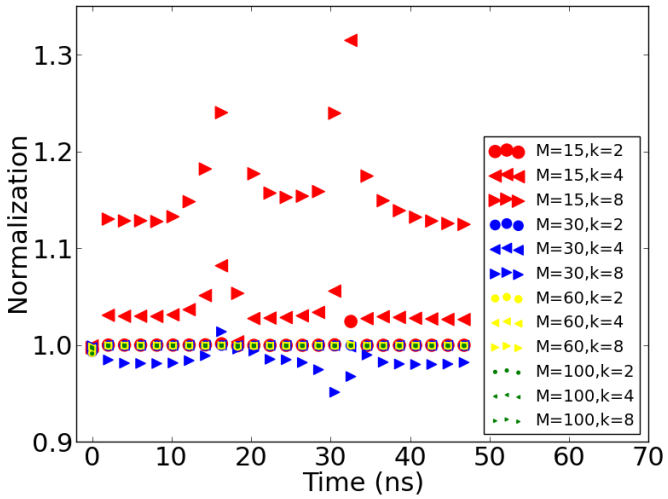


Figure 16: Comparison of the normalization of ρ using different number of subintervals and smearing constant.

Since using a different number of subintervals and smearing constant does not have a large effect on the general shape of the results for large enough M, we can focus on using a single value for M and examine the results. Using $M = 60$ and the same three values for k gives the entropy as a function of time as shown in figure 17. The entropy appears to have a similar shape to the plot of the emittance. There is an initial rise followed by a dip, then another rise and fall, and an increase all the way to the end of the simulation. Figure 18 plots the same but as a function of where the furthest ray is at the time the entropy was calculated, with a line marking where the aperture is. The second peak and second dip appear at the aperture, and immediately after the aperture, respectively, exactly like the emittance.

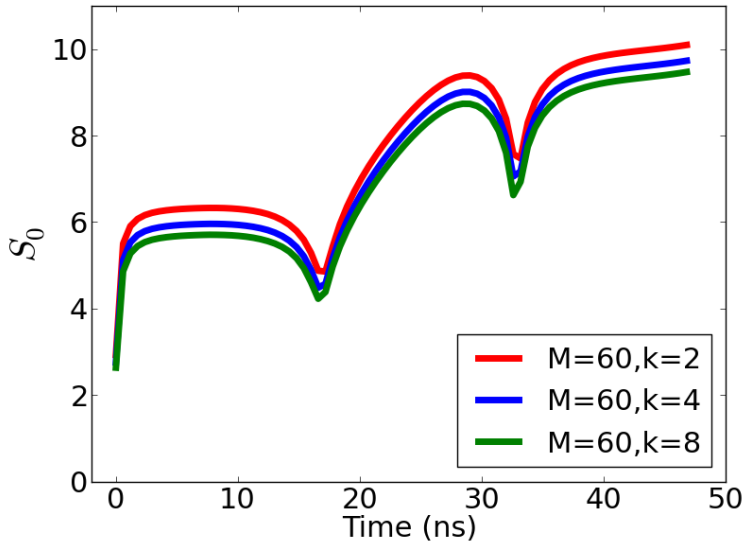


Figure 17: Entropy as a function of time with $M = 60$.

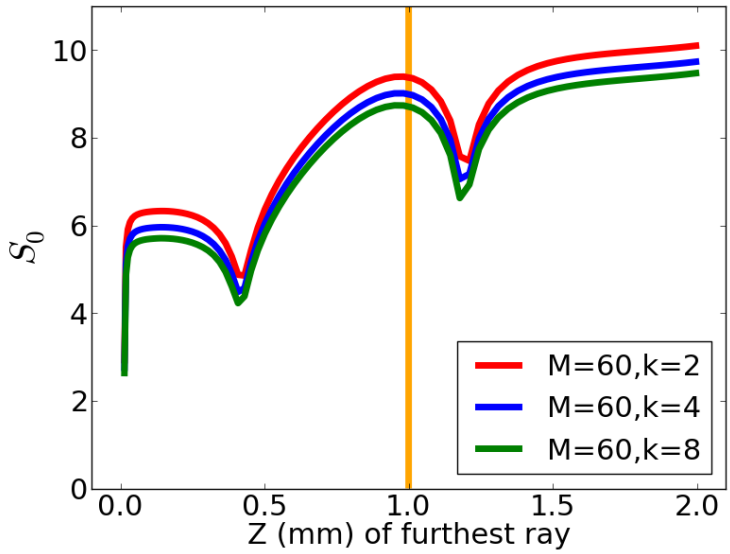


Figure 18: Entropy as a function of time plotted at the z-coordinate of the furthest ray.

Thus, the entropy is not constant when it is calculated using 2-D phase space. Instead it increases and decreases throughout the simulation similarly to the emittance. There was no theoretical reason for expecting the 2-D entropy

to be constant, so this result is not surprising, but it is discouraging that the result is extremely different from a constant function. However, it does give something to compare to the calculations done in four dimensions. Also, the parallel between the shapes of the emittance and entropy graphs might be a clue that there is something else going on yet to be understood.

4.4 Calculations of the entropy in 4-D transverse phase space

The last section did the entropy calculation using only two dimensions: x and p_x. It is better to do these calculations using four (adding y and p_y) or six (adding z and p_z) dimensions. According to theory, the entropy should be constant for six dimensions and approximately constant for the four transverse dimensions, with no such claim for two dimensions.

For four dimensions the charge distribution for the i^{th} ray is

$$\rho_i(x, y, p_x, p_y) = \frac{1}{(2\pi)^2 \sigma_x \sigma_y \sigma_{p_x} \sigma_{p_y}} \frac{q_i}{q} e^{-\frac{1}{2}(\frac{x-x_i}{\sigma_x})^2 - \frac{1}{2}(\frac{y-y_i}{\sigma_y})^2 - \frac{1}{2}(\frac{p_x-p_{x_i}}{\sigma_{p_x}})^2 - \frac{1}{2}(\frac{p_y-p_{y_i}}{\sigma_{p_y}})^2}$$

Again, ρ, the charge distribution of the entire beam, is obtained by summing over each exponential, of which there is one for each ray.

The rays are reflected from one octant to fill the same space as in the previous section, but now the integral over x and y only needs to be over the one octant while p_x and p_y are still over all values in the range. The four-dimensional integral was also done using a Riemann sum. The number of subintervals that could be used is limited because it takes a long time for the computer to do the calculations since it has four dimensions; this means that doubling the number of points in each dimension would make the program take 16 times longer. This is a problem because if M is not large enough, then the integration will miss many of the exponentials causing the results to not reflect what it should be.

The result of the integration is shown in figure 19. The calculation is done with $M = 15$ for $k = 2, 8$ at eight points and $k = 4$ at 16 points. It is hard to tell with so few points, but S_0 appears to have a similar shape to the 2-D case: S_0 is generally increasing over the simulation, but there are two local minima near the middle. The graph leaves much to be desired, but we can be confident in concluding that S_0 is not constant over the entire simulation. While this is not constant like we expected, it is promising that the behavior of the function is nearly the same as in two dimensions. It suggests that the shape of S_0 in both cases are related. Thus it allows us to focus future research on the behavior of S_0 in 2-D, since understanding the behavior of one will most likely help in understanding the behavior of the other.

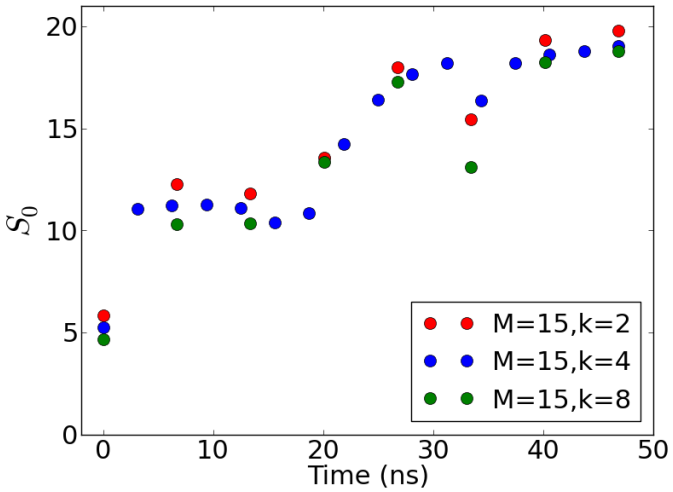


Figure 19: Entropy calculation for 4D.

The value of 15 used for M seem to be very small. However, it is not feasible to use a large number of subintervals because the integration is over four dimensions. This means that doubling M would cause the calculation to take 16 times longer; so if the calculations take 12 hours, which these do, then it would take over a week to do them with $M = 30$. Using $M = 15$ is probably not suitable for precise calculations, but it might be good enough for finding the general shape of the graph. The fact that the shape of the graph is very similar to the calculations in 2-D is also reassuring. The normalization of the density provides a check on the integration step size. Figure 20 shows the results when ρ is integrated using the same parameters as in figure 19. The values are miserable for $k = 4$ and $k = 6$, which makes sense because these have the least spread. The normalization for $k = 2$ is good; this is expected since the smearing is larger relative to the integration step size.

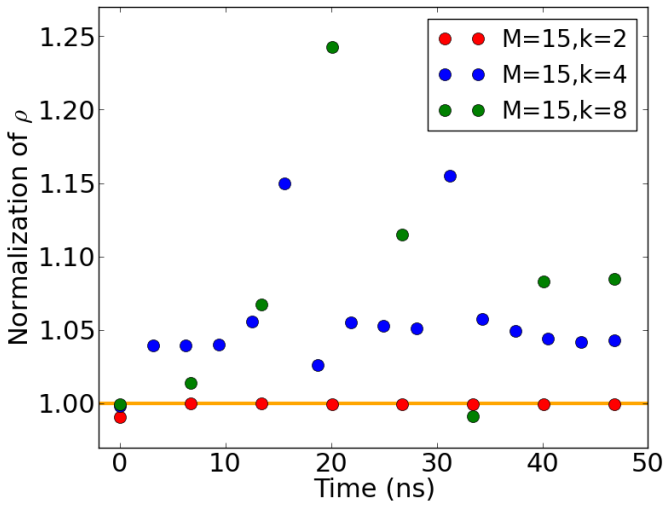


Figure 20: Integral of ρ over all space for the entropy calculations done in figure 19. The orange line is at 1, which is what all the values should be if M were large.

4.5 Calculations of the entropy in 4-D transverse phase space using polar coordinates

The integration in four dimensions is far too time consuming and can be reduced using symmetry arguments. A canonical transformation can be used to switch from x,y,p_x,p_y to r,θ,p_r,p_θ, which is a switch from rectangular coordinates to polar coordinates. The density $\rho(r, \theta, p_r, p_\theta)$ should be independent of θ by symmetry. We can also use the conservation of angular momentum and field emission theory to reduce the integral by another dimension. Conservation of angular momentum means that p_θ is determined upon emission of each ray, and this can be described by the Folwer-Nordheim theory. Thus, the distribution of p_θ is roughly Gaussian and symmetric about $p_\theta = 0$. We can rewrite ρ as

$$\rho(r, p_r, p_\theta) = \rho_r(r, p_r)\rho_\theta(r, p_r, p_\theta)$$

where

$$\rho_r(r, p_r) = \int_{-\infty}^{\infty} dp_\theta \rho(r, p_r, p_\theta)$$

and

$$\int_{-\infty}^{\infty} dp_\theta \rho_\theta(r, p_r, p_\theta) = 1$$

Then the distribution of angular momentum can be approximated with a

Gaussian using the rms angular momentum $< p_\theta^2 >$ so that

$$\rho_\theta(r, p_r, p_\theta) = \frac{e^{-\frac{p_\theta^2}{2<p_\theta^2>}}}{\sqrt{2\pi < p_\theta^2 >}}$$

where

$$< p_\theta^2(r, p_r) >= \frac{\sum_i q_i p_\theta^2 [e^{-\frac{1}{2}(\frac{r-r_i}{\sigma_r})^2} + e^{-\frac{1}{2}(\frac{r+r_i}{\sigma_r})^2}] e^{-\frac{1}{2}(\frac{p_r - p_{r_i}}{\sigma_{p_r}})^2}}{2\pi^2 \sigma_r \sigma_{p_r} q}$$

Now the integration for the entropy is reduced to the double integral

$$S_0 = -2\pi \int_0^\infty dr \int_{-\infty}^\infty dp_r \rho_r log(\frac{\rho_r}{\sqrt{2e\pi < p_\theta^2 >}})$$

The results of doing the integration with different number of subintervals and different values for the smearing constant are shown in figure 21. Once again changing the number of subintervals from 60 to 100 makes only a small difference, so the integral must be stable and approximately correct when M is at least 60. The normalization by integrating ρ over most of the volume for all of these is shown in figure 22. It is clear from this that $M = 15$ is insufficient, but using 60 or 100 for M gives very good results. Changing k from 2 to4 or 8 makes only a small difference, and decreases the value of S_0 by a small amount, just like in two dimensions.

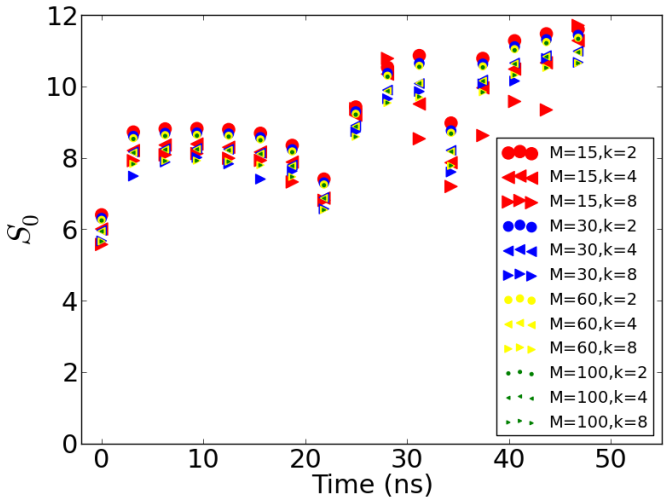


Figure 21: Entropy in 4D using polar coordinates for varying values of M and k.

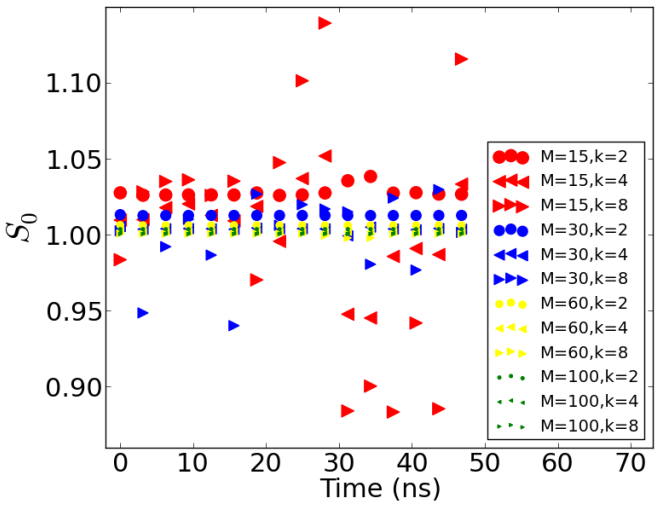


Figure 22: Entropy in 4D using polar coordinates for varying values of M and k.

The entropy calculated at more times using M=60 and varying values of k is seen in figure 23. This plot shows more of the behavior over the whole simulation, the result of which is very similar to the two and four dimensional calculations. Once again the entropy increases for most of the simulation but has two major dips. Unfortunately, these values of entropy do not match up with the calculations done in the full four dimensions. Those values range from 5 to 20, while these values range from 6 to 11, although it is promising that the initial values are similar. The values for S_0 should be the same at all times if the approximations are correct and if the computations are done with enough accuracy.

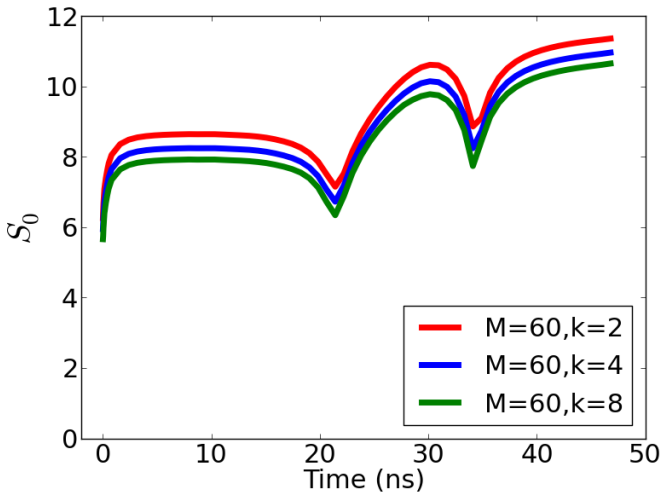


Figure 23: Entropy in 4D using polar coordinates for M=30 and varying values of k.

4.6 Calculation of the entropy in 6-D phase space

Calculating the six dimensional entropy using this method would not work well for a couple reasons. First of all, the amount of time it would take to do the integral in six dimensions would be unwieldy. Doing them in four dimensions is already too difficult to do in reasonable time. Second of all, there are too few particles from the simulation to spread them out smoothly over six dimensions. There would be too much space between the rays to get a useful approximation to the beam distribution. This could be countered by using more rays in the simulation, but this would have consequences. It would take a lot longer to do the computation with more rays. There would already be a problem in calculating a six dimensional integral to high accuracy, even by a simple summation approximation, so adding more rays to the calculation would only make it worse. Symmetry may be used to reduce the integral to four-dimensions, but that would still be difficult to do. However, it may be possible to do it using a computing cluster or by using other integration methods, such as Monte Carlo integration.

4.7 Conclusion

The entropy for electron beam computer simulations has been calculated in order to give a check on the emittance calculation of these simulations. The discrete rays from the simulation are smoothed out using Gaussian distributions to get a charge distribution. The normalized entropy is defined as $S_0 = -\int \rho log \rho$. Using the two-dimensional phase space, the entropy is not a constant over the course of the function. The entropy is generally increasing but has two dips, giv-

ing a shape that is very similar to that of the emittance. We do not expect this to be constant, but it should be approximately constant for four-dimensional phase space. However our calculations show that this integral increases and decreases over the course of the simulation, exactly like in two dimensions. This integration in four dimension is very computationally expensive. This can be avoided by using symmetry and a transformation to polar coordinates, which gives an approximation to the integral but the integration is only over two dimensions. The resulting shape of the graph of S_0 using this method is very similar to the other calculations, but the values do not match up with the four-dimensional integral, which should happen in theory.

Therefore our results do not agree with the theory. The entropy of the electron beam is not conserved over time. Thus there must be an error either with the calculations or the simulation software. If there are no mistakes in the programming for the emittance and entropy, then the most likely source of error is the limited number of particles that are simulated using CPO on a desktop computer. Using rays from four simulations gave 187 particles. The statistical fluctuations are expected to be small ($1/\sqrt{187} \approx 7\%$). However, it is possible that details of the distribution function are not adequately represented. For example, the distribution after the cubic transformation is very constricted near the origin, as shown in figures 11b and 12b. This constriction will likely be lost once any smearing is done, which changes an essential feature of the distribution and may significantly affect calculations.

5 Conclusion

5.1 Summary of results

This research has focused on using computer simulations of field-emission electron beams from a gated cathode geometry to study the beam properties. The geometry consists of a cathode surrounded by a gate, and a far away anode with an anode aperture located halfway through the simulation. The simulations were done with the software CPO, which runs simulations with around 100 rays, which are particles with varying amounts of charge. The data used for calculations used 187 rays all emitted from the same octant of the cathode, so reflection is used to represent an entire beam.

The first objective of the project was to compute the emittance of the beam. The emittance is the volume of phase space occupied by the beam, and it is a measure of beam quaility or coherence. Experimentally, the emittance is a very important property of an electron beam. According to Liouville's theorem, this quantity should be conserved over time. The rms emittance is a value that can be calculated for an electron beam that represents the effective volume of the beam in phase space at a given time. It is representative of emittance, although it is not a conserved quantity. Since it is easier to define and compute, it is the most used measure of beam quality. Our results show that field-emitter tips have extremely small emittance, on the order of a few nm, which is much better

than thermionic or photocathode sources. These calculations were done both as a function of the axial position and as a function of time. It is often assumed that these values are approximately the same. However, this assumption is rather poor. The rays that are closer to the center are significantly further along than the rays near the edge of the beam. This makes a large difference in areas where the field is changing because the central rays experience the change in field before the outer rays. The results of the emittance calculation show that the emittance has a small rise and decrease near the beginning followed by a large increase that peaks at the aperture. There is a major decrease immediately past the aperture, but then the emittance increases for the rest of the simulation surpassing the first maxima.

These results are interesting for a couple of reasons. First of all, the value of the emittance is much smaller than can be obtained from other cathodes, such as thermionic cathodes or photocathodes. An emittance on the order of a nanometer is found in the simulations. Second of all, it was not expected that the emittance would have such large increases and decreases. There is no theoretical reason for the rms emittance to be constant, but it is generally thought to exhibit little fluctuation while generally increasing. Standard computer checks and an analytical model support the correctness of this result. One concern is that small emittance is difficult to computer accurately because small computer errors are magnified in the small phase space of the beam.

One way to check this unexpected result is to calculate the entropy of the same beam. For noninteracting particles, entropy is conserved. The entropy of a particle beam was introduced in 1973, but prior to this research the entropy has never been computed or used quantitatively. Many difficulties lie in computing the entropy of a particle beam, both numerical and theoretical. The normalized entropy S_0 is defined to be the integral of $-\rho \ln \rho$ over phase space, where ρ is the normalized charge density. Since the simulations use a discrete number of particles instead of a charge distribution, we convert the rays to a continuous distribution by using Gaussians to smear each ray. Summing all of the Gaussians then gives a charge distribution. There is difficuly in how these rays should be smeared, and it generally will work better with a larger number of rays.

The entropy should be constant when it is calculated in six-dimensional phase space. We began by calculating the entropy in 2-D phase space. The plot of S_0 turns out to be very similar to the plot of the emittance, with two maxima and two minima followed by an increase to the end of the simulation. There is no reason for this 2-D entropy to be constant. Next we calculated the 4-D entropy using a similar method. The 4-D trasverse phase space gives a decent approximation for the 6-D phase space, so this entropy value should at least be approximately constant. Instead we find that the plot of entropy has a generally similar shape to the 2-D entropy. This calculation is very difficult to make because the integral over four dimensions is very cumbersome. Moreover, the process of smearing the relatively small number of rays over four dimensions is not going to give good results. One way this integral can be simplified is by transforming to polar coordinates and using the properties radial symmetry and the Fowler-Nordheim equation. Doing this reduces the integral from four

dimensions to two dimensions. Using this method to calculate the entropy once again gives the same shape as in the 2-D phase space. Also, the values do not agree with those from the 4-D entropy integral, which means there must be an error somewhere. Doing the integral in 6-D phase space would take far too long and would require using many more rays in order to ensure the spreading over six dimensions gives a sensible result.

None of the entropy calculations performed give a constant entropy, as expected for the 6-D phase space entropy. Instead all calculations have a shape similar to the emittance that has multiple maxima and minima during the simulation. It is not clear whether the discrepency is due to the calculations not being done in 6-D phase space or the calculations being done incorrectly. There are three possible sources of error in the calculations. The first is that the coding and analysis could contain errors. The second is that the simulations may not be accurate enough due to small rounding errors having a large effect. A final source of error is in the challenges the analysis presents, namely using a small number of particles to make a continuous distribution. Further research is needed to identify the source of the problems and how they can be remedied. It would be extremely helpful to have an analytical calculation for what the entropy calculation should be in 2-D and 4-D in order to have something to check the results with.

This is a new area of particle beam research. The results presented here are the first known attempt to calculate the entropy of an electron beam in this way, which means that there is a lot of room for improvement and future research.

5.2 Suggestions for future research

The extremely small emittance of field emission cathodes suggests that these cathodes should be developed for various applications. These cathodes should then be used in experiments to check the results of the simulations. This will be difficult since the small emittance is difficult to measure.

The computation of entopy should be continued to be explored. All simulations and calculations done in this paper should be checked to ensure there was not an error. The simulations should be done with more accuracy and with varying parameters to make sure the calculations are stable. The simulations should be done with different software to check the techniques used by CPO. A lot of data analysis and programming was used in this project that easily could have been done wrong, so these should be redone by independent parties. The assumptions used should be checked, and the methods use for smoothing the distribution and calculating the integrals can be improved. It would be very helpful to have an analytical calculation for what the entropy should be in two and four dimensions. Even if this value were known only for time zero it would be sufficient to either disprove or lend credence to calculations. The behavior of 4-D transverse phase space entropy should be explored theoretically. We expect the 4-D entropy to be similar to the 6-D entropy, which is conserved, but we do not know how similar it should be. The magnitude of the errors introduced in simulations can also be studied to provide better perspective on the accuracy

of the calculations.

6 Appendices

6.1 Simulations of gated cathode emittance

Simulations of gated cathode emittance

C. B. Erickson, J. D. Jarvis, and C. A. Brau
Vanderbilt University, March 5, 2012

There is now substantial interest in developing rf linacs to produce electron beams with high brightness and small emittance. Applications for such beams include free-electron lasers [i], electron diffraction imaging [ii], and high-spectral-brilliance x-rays from channeling radiation[iii]. While FELs require high current to exceed threshold, electron diffraction and channeling radiation are linear processes and can use much lower current provided that the emittance is small.

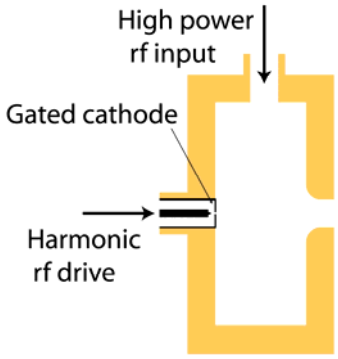


Figure 1. Rf injector with a gated field-emission cathode.

For high-current applications, diamond field-emitter arrays have been suggested as an alternative to photocathodes in rf injectors [iv]. The tips may be gated or ungated. However, to achieve exquisitely small emittance at lower current, it is sufficient to use a single field-emitting tip. As much as 15 μA dc has been observed from such tips[v], with very small divergence from ungated tips [vi].

In an rf injector, the cathode is placed at the back of the cavity, as shown in Figure 1. If a gated cathode is used, the emission can be restricted to a short pulse timed to the optimum phase for injection. First- and third-harmonic drive is applied to the gate by placing the gate at ground potential on the surface of the cavity and placing the cathode on the center conductor of a coaxial rf line. Simulations indicate that the emittance can be preserved through the gun and the accelerator [vii]. If an ungated cathode is used, the emission can be timed by using a gun cavity with first and third harmonic fields [viii].

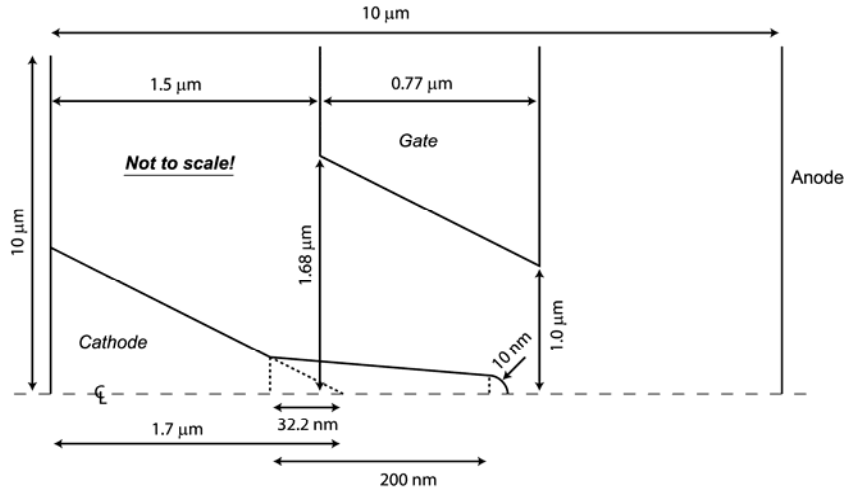


Figure 2. Cathode geometry used in the simulations.

As a first step in the development of ultra-small-emittance electron beams, we have undertaken simulations of a gated diamond field-emission cathode. The cathode geometry used in these simulations is shown in Figure 2 (not drawn to scale). The anode is placed at 10 microns from the base of the cathode. In an rf gun there is no anode; the field outside the gate is

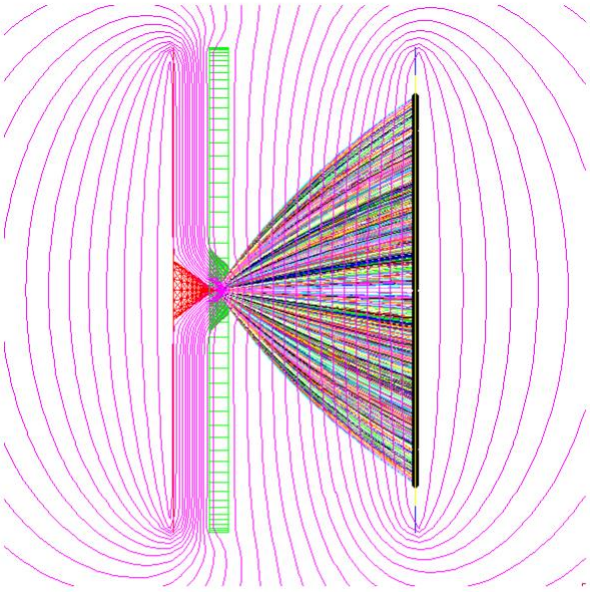


Figure 3. Equipotential surfaces and trajectories for a gated cathode computed using CPO .

provided by a mode of the rf cavity. Therefore, in setting up the computations, the important parameters are the bias on the cathode, relative to the gate, and the electric field above the gate, rather than the anode potential.

The code used for these simulations is CPO3DS [ix]. For the field computations, CPO uses the boundary element method. In this method, the boundaries are divided into segments and the surface charge on each segment is adjusted until the surface forms an equipotential. This has the advantage over codes, such as Poisson, that compute the field on a mesh since the segments can be made to conform to the surface of the boundaries. In the computations, the number of segments was increased until no effect was observed on the fields. Also, CPO uses rays instead of individual electrons; these simulations used 192 rays from the tip. Figure 3 shows the equipotential curves for the case $V_{cathode}$=-150 V, V_{gate}=0, and V_{anode}=150 V.

As a check, the results were compared to computations made with Poisson, using input files from John Lewellen [x]. Two cases were run, as shown in Table I. The most important measure is the field at the tip of the emitter as a function of the cathode bias, relative to the gate, and the field outside the gate. Since the tip field is linear in the cathode potential and the external field, the ratio of tip field to cathode bias and tip field to external field are used for comparison. The agreement is satisfactory.

Table I. Fields computed using CPO and Poisson

$V_{cathode}$	V_{gate}	V_{anode}	$E_{tip}/V_{cathode}$	$E_{tip}/V_{cathode}$	E_{tip}/E_{anode}	E_{tip}/E_{anode}
			CPO	Poisson	CPO	Poisson
-100	0	0	53 /μm	43 /μm		
0	0	100			12	10.2

Next, trajectories were computed using CPO for the case $V_{cathode}$=-150 V, V_{gate}=0, and V_{anode}=150 V. For this case, the external field is 20.4 V/μm, which is on the order of the fields expected in an rf injector. For these parameters, the tip field is 8200 V/μm. The total current from the cathode computed by CPO using a work function of 4.5 eV is 71 μA. This is higher than we expect to get in experiments, so the emittances computed are probably conservative. The trajectories are shown in Figure 3.

The phase space of the electrons at the anode is shown in Figure 4. The curvature of the distribution due to spherical aberration of the gate field is apparent. It should be pointed out that due to the wide angles of the outer electrons relative to the axis, the paraxial approximation is not valid. It is therefore

more convenient to compute the normalized emittance from the transverse momentum using the formula

$$\varepsilon_N{}^2 = \frac{\left[\sum_i q_i x_i{}^2\right]\left[\sum_i q_i p_{x_i}{}^2\right] - \left[\sum_i q_i x_i p_{x_i}\right]^2}{4m^2c^2\left(\sum_i q_i\right)^2}$$

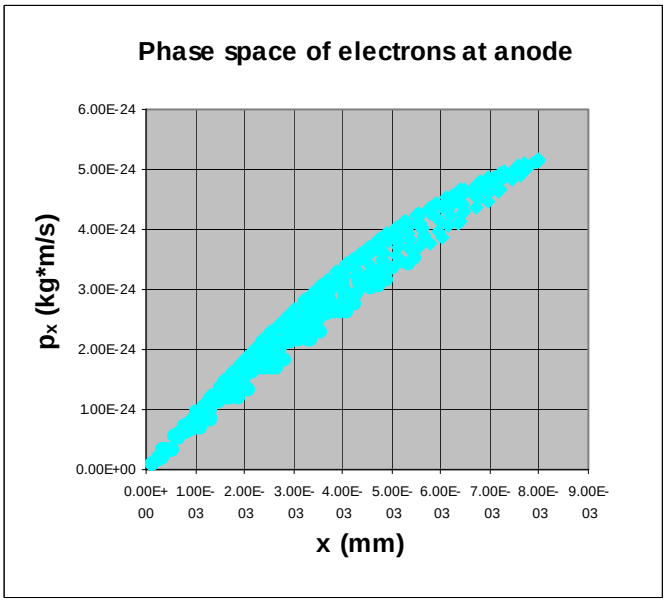


Figure 4. Phase space of the electrons at the anode of the gated cathode shown in Figure 2.

This equation, in which m is the electron mass and c the speed of light, gives the emittance, ε_N, in terms of the momentum p_{x_i} and coordinates x_i of the electrons at the anode, and the current q_i associated with each trajectory. The normalized emittance computed in this way is 2.1 nm at the anode. Calculations of the emittance using the code IMPACT-T give 2.7 nm [xi].

For comparison we have also calculated the emittance of a field emission cathode with no gate electrode. For this case the cathode voltage is increased to -362 V and the anode voltage is increased to 362 V so that the current is 8.99 μA, which is comparable to the previous case. The fields and trajectories are shown in Figure 5. The emittance is found to be 0.736 nm, which is about a factor of three smaller than the emittance computed for the gated cathode of Figure 2. The difference is presumably due to the spherical aberration of the gate field.

To take advantage of the smaller emittance of an ungated tip, we have computed the emittance of a gated cathode in which the tip lies above the level of the gate. For this case we lowered the surface of the gate by 500 nm, compared with Figure 2, but the rest of the geometry is the same. The cathode voltage was lowered to –141.9 V, and the anode voltage was lowered to 141.9 V so that the current

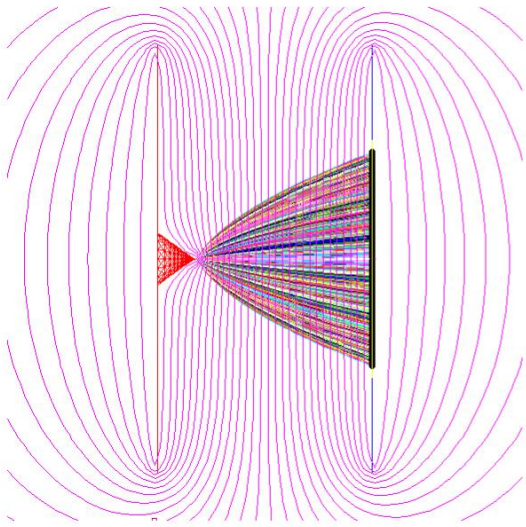


Figure 5. Electron trajectories and potential contours computed using CPO for the case with no gate.

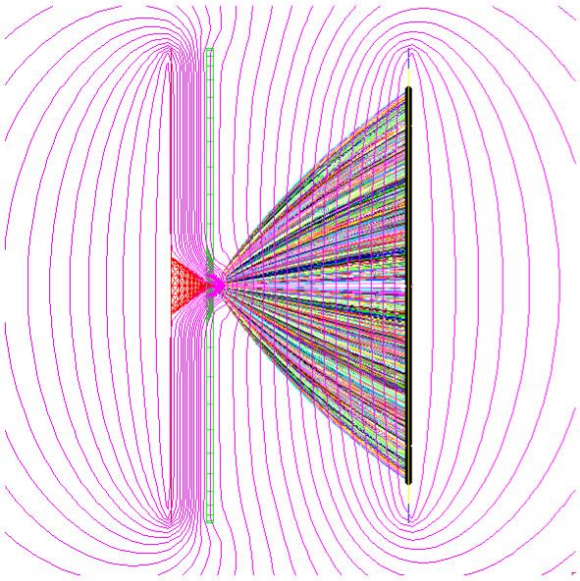


Figure 6: Electron trajectories and potential contours computed using CPO for the case with the gate lowered 500 nm.

is 8.98 μA, which is comparable to the previous cases. The contours and ray trajectories are shown in Figure 6. The emittance is found to be 2.24 nm.

---

[i] C. A. Brau, *Free-Electron Lasers*, Adv. Electronics and Electron Phys. Suppl. 22 (Academic Press, Boston, 1990).
[ii] P. Musumeci, J. T. Moody, C. M. Scoby, M. S. Gutierrez, and M. Westfall, Appl. Phys. Lett. **97**, 063502 (2010).
[iii] C. A. Brau, B.-K. Choi, J. D. Jarvis, J. W. Lewellen, and P. Piot, Synch. Rad. News **25**, No. 1, 20 (2012).
[iv] J. D. Jarvis, H. L. Andrews, C. A. Brau, B. K. Choi, J. Davidson, W.-P. Kang, and Y.-M. Wong, J. Vac. Sci. Technol. B 27, 2264 (2009).
[v] J. D. Jarvis, H. L. Andrews, B. Ivanov, C. L. Stewart, N. de Jonge, E. C. Heeres, W.-P. Kang, Y.-M. Wong, J. L. Davidson, and C. A. Brau, J. Appl. Phys. **108**, 094322 (2010).
[vi] J. D. Jarvis, B.-K. Choi, A. B. Hmelo, B. Ivanov, and C. A. Brau, J. Vac. Sci. Tech. B (to be published).
[vii] P. Piot and D. Mihalcea, private communication.
[viii] J. Lewellen and J. Noonan, Phys. Rev. ST Accel. Beams 8, 033502 (2005).
[ix] CPO Ltd. Charged Particle Optics programs, www.electronoptics.com.
[x] J. W. Lewellen, private communication.
[xi] P. Piot and D. Mihalcea, private communication.

7 References

References

[1] Akre, R. et al. *Commissioning the Linac Coherent Light Source Injector.* Phys. Rev. ST Accel. Beams 11, 030703. 12 March 2008.

[2] Brau, C. A., B.-K. Choi, J. D. Jarvis, J. W. Lewellen, and P. Piot. *Channeling Radiation as a Source of Hard X-rays with High Spectral Brilliance.* Synchrotron Radiation News, Volume 25, Issue 1, 2012, pages 20-24.

[3] de Jonge N., Y. Lamy, K. Schoots, T. H. Oosterkamp. *High brightness electron beam from a multi-walled carbon nanotube.* Nature. 2002 Nov 28; 420(6914):393-5.

[4] Erni, Rolf, Marta D. Rossell, Christian Kisielowski, and Ulrich Dahmen. *Atomic-Resolution Imaging with a Sub-50-pm Electron Probe.* Physical Review Letters 102, 096101 (2009).

[5] Huang, Zhirong and Kwang-Je Kim. *Review of x-ray free-electron laser theory.* Physical Review Special Topics - Accelerators and Beams, 10, 034801 (2007).

[6] Jarvis, J. D., H. L. Andrews, B. Ivanov, C. L. Stewart, N. de Jonge, E. C. Heeres, W.-P. Kang, Y.-M. Wong, J. L. Davidson and C. A. Brau. *Resonant tunneling and extreme brightness from diamond field emitters and carbon nanotubes.* Journal of Applied Physics, 108, 094322 , (2010).

[7] Lawson, J. D., P. M. Lapostolle, and R. L. Gluckstern. *Emittance, Entropy and Information.* Particle Accelerators. 1973, Vol. 5, pp.61-65.

[8] Rosenzweig, James B. *Fundamentals of Beam Physics.* Oxford University Press, Oxford, 2003. p. 126.

[9] Struckmeier, Jurgen. *Concept of entropy in the realm of charged particle beams.* Phys. Rev. E 54, 830. 1 July 1996.

[10] Wall, J., J. Langmore, M. Isaacson, and A. V. Crewe. *Scanning Transmission Electron Microscopy at High Resolution* Proc. Nat. Acad. Sci. USA Vol. 71, No. 1, pp. 1-5, January 1974.

[11] Wiedemann, Helmut. *Particle Accelerator Physics, third edition.* Springer.