THE USE OF ELECTROMAGNETIC PARTICLE-IN-CELL CODES IN ACCELERATOR APPLICATIONS*

Kenneth Eppley
Stanford Linear Accelerator Center
PO Box 4349, Stanford, California, 94309

Abstract

The techniques developed for the numerical simulation of plasmas have numerous applications relevant to accelerators. The operation of many accelerator components involves transients, interactions between beams and rf fields, and internal plasma oscillations. These effects produce non-linear behavior which can be represented accurately by particle in cell (PIC) simulations. We will give a very brief overview of the algorithms used in PIC Codes. We will examine the range of parameters over which they are useful. We will discuss the factors which determine whether a two or three dimensional simulation is most appropriate.

PIC codes have been applied to a wide variety of diverse problems, spanning many of the systems in a linear accelerator. We will present a number of practical examples of the application of these codes to areas such as guns, bunchers, rf sources, beam transport, emittance growth, and final focus.

Introduction

Over the past thirty years plasma physicists have developed many useful numerical simulation techniques for modelling plasmas and electromagnetic fields and implemented them in a large number of codes. These codes are proving to be powerful tools for the design of accelerator components.

Numerical simulation allows us to study systems which are difficult to treat analytically, for example, devices with non-linearities or irregular boundaries. They also provide a more detailed picture of an experiment than it is usually possible to obtain by measurement. These codes have now become more accessible to the non-specialist, both in ease of preparing the input datasets and in availability of computational resources. With the development of supercomputers, three-dimensional simulations are becoming practical for a much wider range of phenomena.

What are PIC Codes?

Electromagnetic Particle in Cell (PIC) Codes represent charged particle species as a statistical ensemble of test particles, while solving Maxwell's Equations on a finite-difference Eulerian mesh. Calculating the electromagnetic interaction through the intermediary of a mesh rather than by direct particle-particle summation makes the number of operations for N test particles scale as N rather than N^2.

These codes have the capability to use fully time-dependent, relativistic, and self-consistent algorithms. For economy, specific applications may utilize steady-state, electrostatic, or magnetostatic approximations.

Many problems involving intense charged particle beams also involve rapid transients, e.g., pulsed beams, or interaction with rf fields, e.g., bunchers, klystrons, or accelerating

structures. High current beams, as in klystrons, have internal plasma oscillations which significantly affect their behavior. These non-linear effects are generally beyond the scope of analytic methods and are not accurately treated by simpler models such as ray-tracing or electrostatics. Fortunately, the time scales involved in these devices are often well fitted to PIC simulation, so the computer time required is not excessive. There do exist classes of problems with large separation of space and time scales (e.g., an optical FEL), that are not well suited to conventional PIC methods.

Numerical Algorithms

The first step in the numerical algorithm is to define a mesh on which we locate the sources, currents, and fields. Independent of the mesh is a set of macroparticles, each with a given charge, mass, position, and velocity.

We represent Maxwell's equations as a coupled, first order system of partial differential equations. We replace the space and time derivatives by finite differences, staggered in space and time (the "staggered leapfrog" differencing scheme, used widely because it is stable and second-order accurate) The relativistic equation of motion for the particles is replaced by a finite difference equation.

The initial data is stepped forward in time. First the field quantities advance the particle trajectories. The new particle data then generates the sources. These advance the fields. All this occurs over a finite timestep Δt. (Figure 1.)

To calculate the charges and currents, we must interpolate between particle and field positions. Interpolation makes the particles behave as though spread out over a finite region. Several interpolation schemes have been used, but the most common is linear area weighting, which gives a good compromise between accuracy and efficiency. The particle charge or current is divided among the four neighboring mesh points in the proportion that the area of a cell centered on the particle overlaps each of these nodes. This is called the "scatter" of the particle data to the mesh. The reverse interpolation is done to "gather" the field values at the neighboring points to the particle position. Scatter-gather routines use most of the computing time in typical PIC simulations, and often use assembly language routines hand-coded for speed.

We have represented the mesh by rectangular cells. One may also wish to work in curvilinear coordinates. Many codes have the capability to use $R - Z$ and $R - \theta$ coordinates, and some can use more general orthogonal coordinates. A few codes such as ISIS can use fully general coordinates. Generally finite differencing is done in an internal rectangular mesh and a Jacobian transformation is done to produce the correct equations of motion. A few codes have the capability to use triangular zones, which allows much more flexibility in representing curved boundaries at the cost of more complicated bookkeeping.

Utility of PIC Codes

The space and time scales of our simulation must be fine enough to resolve the physics, or else inaccuracy or even instability will result. The time step must satisfy the Courant limit ($c\Delta t/\Delta x < 1$), and must resolve the plasma and cyclotron frequencies ($\omega_p\Delta t < .5$, $\omega_c\Delta t < .5$). Exceeding these limits will cause exponentially growing instabilities. A number of researchers are now working on implicit codes which may alleviate some of these restrictions, at the cost of greater complexity.

There are other sources of numerical error. Accuracy requires an adequate number of particles, both to properly define the distribution and to reduce the numerical noise which results if there are too few particles per cell. The spatial zoning must be fine enough to resolve the wavelengths of interest (generally with at least eight zones per wavelength). Inadequate zoning can cause instabilities such as aliasing, which results from interaction of particle modes with non-physical modes of the grid. Zoning larger than the effective Debye length (defined in terms of total particle energy rather than temperature) will result in artificial heating of the particles. Even with fine zoning, the numerical algorithm cannot exactly preserve Poisson's equation, and many codes periodically correct the solution to force conservation of charge.

There are many more numerical instabilities than we can discuss here (see the literature[1,2,3] for details), but they can generally be removed by sufficiently fine zoning and an adequate number of particles. Accuracy can also be improved by use of higher order differencing and interpolation schemes, at some cost in complexity. Most codes in general use today still rely on relatively simple weighting methods.

A user of PIC codes needs some knowledge of possible numerical pitfalls as well as the physics of the problem. A zoning study is an important part of any PIC calculation. An improperly zoned problem may produce answers which appear reasonable but are highly inaccurate. Most algorithms conserve momentum exactly but not energy. Thus the accuracy of energy conservation is often a good diagnostic of overall numerical accuracy. Code results should also be benchmarked with results from analysis, experiment, and other codes.

An important practical aspect of PIC codes is the specification of boundary conditions. Coding to handle the various types of boundaries generally takes up a large part of the program. Even boundaries coinciding with the mesh endpoints can have a wide variety of types: metal walls, periodic surfaces, symmetry axes, electromagnetic ports, injection and emission surfaces, etc. Most interesting problems also require irregular internal boundaries, with more complicated bookkeeping. Internal boundaries can be modelled by stairsteps, or, in more elaborate codes, by triangles or trapezoids. The internal structures defined by these boundaries can be perfect conductors, real conductors, dielectrics, and in some codes can have more detailed material properties permitting secondary emission, particle tracking with energy loss, or chemical reactions.

Input and output also involve a disproportionately large part of these codes. The specification of arbitrary boundary surfaces in particular is a very time consuming task in setting up problems. Some of the newest codes are developing automatic mesh generators such as have been used in eigenvalue codes like URMEL and SUPERFISH.

Making sense of output from PIC simulations is another problem. These codes can easily generate voluminous quantities of printout. Extracting the relevant information requires careful edits, as what is of interest in one problem may be irrelevant in another.

Two and Three Dimensional Simulations

Until the advent of supercomputers, only 2-D calculations were practical. For many systems, 2-D simulation is quite adequate. 2-D codes involve less data to specify, and the input and output problems are much less severe. Computational requirements of 3-D are often much larger. In some situations, however, the extra complexity and expense of 3-D is well justified. If the boundary conditions do not have 2-D symmetry (e.g., in a cavity with an external waveguide), or if the external fields are asymmetric by design or through error, then there may be inherent 3-D effects. Even with 2-D symmetry, there may be 3-D modes which can develop out of noise and produce important effects. The present generation of supercomputers is only marginally capable of running 3-D PIC codes, but the next generation should be able to handle these problems much more easily. Fortunately, PIC simulations are natural candidates for highly parallel processing.

Existing PIC Codes

Let us list some existing PIC codes. This list is far from exhaustive but is a good sample of general-purpose PIC codes in wide use.

Table 1. Two Dimensional Codes .

Name	Contact	Lab	Comments
ANTHEM	Mason	LANL	Implicit
AVANTI	Langdon	LLNL	Implicit
CCUBE	Godfrey	MRC	General purpose
CONDOR	Nielsen	LLNL	General purpose
ISIS	Jones	LANL	General coordinates
MAGIC	Goplen	MRC	General purpose
MASK	Drobot	SAIC	General purpose
WAVE	Lindman	LANL	Fourier field solution
ZOHAR	Langdon	LLNL	Mainly plasma problems

Table 2. Three Dimensional Codes

Name	Contact	Lab	Comments
ARGUS	Drobot	SAIC	General purpose
IVORY	Godfrey	MRC	3rd D by modes
MAFIA	Weiland	DESY	Fields only as yet
QUICKSILVER	Seidel	Sandia	General purpose
SOS	Goplen	MRC	General purpose

2

We have given a single contact person for each code. Space does not permit listing the many individuals who have been involved in the actual development of each.

Examples of PIC Codes in Accelerator Design

Let us consider the sections of a typical LINAC and examine where PIC codes have been used to model various components. In the injectors and the bunchers, short pulse lengths require time dependent simulation to accurately calculate emittance growth. In the klystrons, 2-D PIC simulations estimate gain and efficiency more accurately than 1-D models. In the damping rings, PIC codes can be used to model the effects of wakefields. In transmission through the linac, PIC codes can model emittance growth due to space charge and misalignments. In the final focus, the beam-beam disruption can only be modelled accurately with a particle simulation.

Out of many interesting examples of such applications I have selected several that are representative of the spectrum of problems that can be addressed with PIC codes.

Haber at NRL and Celata at LBL simulated space charge limited transport of an ion beam in a quadrupole focused beam line, for Heavy Ion Fusion research[4]. The objective was to learn how much current could be transported without instabilities which would unacceptably increase the emittance. Analytic calculations identified a number of unstable modes but were unable to calculate at what level they would saturate. The numerical calculations indicated that a space charge limited beam redistributes itself quickly, so that after an initial increase, the emittance is stable or only slowly growing. This finding gave considerable encouragement to the HIF program. Experiments subsequently verified the simulation findings.

The injector gun for the SLC has been modelled electrostatically using EGUN. However, an accurate calculation of the emittance requires that the transient behavior of the pulse be included. Herrmannsfeldt at SLAC used MASK to model the time-dependent pulse[5]. For this simulation, he did not try to model the details of the electrode structure, which would have been expensive (although such simulations have been done on occasion). He used EGUN to calculate an approximate potential distribution which was input into MASK as a boundary condition. This is not a bad approximation for a low perveance gun, where the space charge does not change the potential significantly. Both EGUN and MASK gave a similar emittance for a steady state pulse, but MASK found three times larger emittance when the transients were included.

Hanerfeld at SLAC used MASK to simulate the SLC buncher[6]. The model included the 178.5 MHz subharmonic buncher, the 2856 MHz buncher, and four accelerating cells. The purpose of the simulation was to study emittance growth and bunching characterics of the system. The bunching and accelerating cavities were modelled as rf ports. The propagation of a typical pulse through the device is shown in Figure 2. The radial variation is due primarily to the magnetic field profile. The simulation gave some insight into the effect of the focusing field on bunching and emittance growth.

At the other end of the SLC is the final focus. Chen, Yokoya, and Hollebeek have used PIC simulation to study the disruption of colliding beams and its effect on luminosity[7]. To simulate a very high gamma beam they used a modification of conventional PIC algorithms. The beamstrahlung effect was put in phenomenologically using the Sokolov-Ternov spectrum to calculate the electron energy loss. The beams pinch as the magnetic fields add while the space charge forces cancel. This effect can be used to increase the luminosity. The simulations have explored how the design parameters of a future linear collider would influence the luminosity enhancement in the final focus.

Godfrey and Hughes at MRC have used the 3-D code IVORY to model electromagnetic instabilities in a high current recirculating accelerator, a modified betatron with a strong toroidal magnetic field[8]. There are a number of instablities (e.g., three wave, resistive wall, beam breakup) which can potentially disrupt the beam. Extensive simulations have modelled a number of different field configurations to calculate growth rates and saturation levels for these modes. The results of the simulations have been useful in testing ideas for stabilizing the beam. A 3-D code is needed to perform these calculations. Figure 3 shows an example of a resonant negative-mass instability.

Krall at SAIC and Lau, Friedmann, and Serlin at NRL have been studying a compact rf accelerator using CONDOR[9]. This device combines features of an auto-accelerator and a two-beam accelerator in which the modulated driving beam is closely coupled to the accelerating structure. The modulating section differs from a conventional klystron in using a high perveance, hollow beam. Figure 4 shows the simulation geometry. The code calculations agreed with linear theory for low modulating voltages and explored some aspects of nonlinear theory. The accelerating structure has also been simulated. The initial concept involved a photocathode, but the simulations have indicated that a dc emitted beam will be bunched and accelerated as well.

Ko, Drobot, Mankowsky, and Mondelli at SAIC have used the 3-D code ARGUS to do cold-test simulations of a double output gap for the 5045 klystron at SLAC[10]. Figure 5 shows the geometry of the cavity in cut-away, illustrating the number of zones needed for accurate 3-D modelling and the usefulness of good graphic output to understand the results. This cavity has a coupling slot that breaks the symmetry, and can only be modelled with a fully 3-D code such as ARGUS or MAFIA. The code calculations agreed well with experimental tests. When this cavity was initially tested, it worked well with a dummy load but oscillated when connected to the accelerating structure, due to a resonance at one of the higher modes of the cavity. It is hoped that the ability to accurately calculate the modes of these structures will make it easier to avoid such problems.

Petillo at SAIC is using ARGUS to simulate a depressed electron collector, used as a beam dump. The presence of vanes on the collector plates is intended to minimize noise due to radiation and secondary emission. It also requires 3-D modelling. This simulation used an electrostatic approximation, but included secondary particle emission. The goal is to optimize performance by varying the placement of the vanes. Figure 6 shows the geometry of the device. The vanes greatly reduce the number of reflected particles. An ongoing study is attempting to optimize the plate voltage and vane locations.

Chang at SAIC is applying ARGUS to simulate beam transport and quadrupole focusing of medium to heavy ions. A 3-D code is needed to study phase mixing of envelope

oscillations or axial buching and debunching when the beam does not obey the K-V distribution or when space charge is non-uniform. The current simulation study models a system of magnetic quadrupoles set up in a FODO array. The beam is followed by a moving window in the lab frame, which follows the longitudinal motion of the beam centroid. Figure 7 shows a 3-D perspective plot of the beam ion positions at an early time.

I will discuss a typical PIC simulation in more detail, in particular, modelling the SLAC 5045 klystron using MASK. Modelling a klystron with a PIC code is a challenge, due to the wide range of time, distance, and energy scales present. The signal strengths vary between a few hundred watts at the input cavity to tens of megawatts at the output. The filling times of the idler cavities, which are typically not externally loaded, are hundreds of rf cycles. The low voltage cavities are usually detuned by very small amounts, and the gain is highly sensitive to this detuning. The tuning and shunt impedence of the cavities are sensitive to the exact shape of the cavity noses, which are quite small compared to the length of the drift tube. An accurate simulation of the cavities would require extremely fine zoning. Thus it seems that this simulation would be impractically expensive to perform.

Simon Yu and Adam Drobot solved this by replacing the cavity with a port boundary condition[11]. This port model is valid as long as the beam does not enter the cavities. For economy we consider only the steady-state behavior. The only interaction between the beam and the cavity is the power flow across the cavity gap. By using the power balance equations, one can relate the induced current I in the wall to the Fourier component of a volume integral of $\vec{E} \cdot \vec{J}$ in the drift tube. Given any complex cavity impedence Z we seek values of V and I for which $V = IZ$. The non-linear dynamics is contained in the variation of the current I with gap voltage V. We have developed an iterative scheme which finds a self-consistent solution in a few rf cycles.

We have used this algorithm to model the 35 MW XK-5 and the 50 MW 5045 klystrons in use at SLAC. In particular, the simulation was used to track down possible sources of an amplitude modulation which was a serious problem in the 5045. This modulation was in the 10-20 MHz range and was highly sensitive to the settings of the focusing coils, especially near the cathode, and to the drive level. Many changes in the design of the tube were made, some of which reduced the modulation. In particular, one set of tubes had the sides of the cathodes coated with moly-ruthenium, which supresses emission. While these tubes were not totally free of the modulation, they performed better than untreated cathodes. This suggested that we model the effects of emission from the side of the cathode to see what it would do in the simulation.

The EGUN code was used to calculate the trajectories between the cathode and the beginning of the drift tube. For this simulation, a virtual cathode was assumed between the edge of the cathode and the focus electrode, since there is a large area of potential emission surface just behind this space. These extra side-emmitted electrons were tagged separately and used in the MASK simulation of the klystron. They are found to occupy a different part of phase space than the rest of the beam. The simulation showed that for the original magnet settings, at a few db above saturation drive level, these electrons go directly into the penultimate cavity during part of the rf cycle (Figure 8). If one either lowers the drive (lowering the voltage on the penultimate cavity) or increases the bucking coil current (making the beam smaller) these electrons do not enter the cavity. It is easy to devise mechanisms whereby the presence of electrons in the penultimate cavity would reduce the voltage in that cavity, leading to a relaxation oscillation. The filling time for this cavity is a few tens of ns, consistent with the observed 10-20 MHz modulation.

The occasional existence of amplitude modulation for drive levels below saturation was not accounted for by this scenario. However, the code showed that the side-emitted electrons intercepted the output cavity differently at different drive levels. Under some focusing conditions, at saturation they would intercept the drift tube downstream of the output cavity, while below saturation they would intercept the downstream nose. In the first case specular reflection of the beam should be downstream, while in the second case at least some of the beam could be reflected upstream. Even a tiny amount of modulated current reaching the low-power cavities can change their voltage and phase, giving another possible relaxation mechanism.

While the steady-state model cannot fully simulate the time-dependent modulation, the MASK simulations provided a useful clue to the source of the problem, and suggested a number of ways to correct it.

Conclusions

Numerical PIC simulation is a well-established technique for modelling many systems of interest in accelerator design. A large number of codes are in use and are being applied to a wide array of problems in the design and analysis of accelerator components. We have attempted to give an overview of the methods used in such codes. We have given a few typical examples of the applications of these codes to situations of practical interest to accelerator design.

References

1. C. Birdsall and A. Langdon, *Plasma Physics via Computer Simulation,* McGraw Hill, New York, 1985.

2. R. Hockney and J. Eastwood, *Computer Simulation Using Particles,* McGraw Hill, New York, 1981.

3. B. Adler, S. Fernbach, M. Rotenberg, and J. Killeen, eds., *Methods of Computational Physics 16, Controlled Fusion,* Academic Press, New York, 1976.

4. I. Haber, "Numerical Study of Space Charge Limited Transport in a Quadrupole Transport System with 90 Degree Phase Advance," NRL Memorandum Report 4474, Naval Research Laboratory, May 1981.

5. W. Herrmannsfeldt, "Some Applicattions of Particle-in-Cell Codes to Problems of High Intensity Beams," SLAC PUB 3927, Stanford Linear Accelerator Center, April 1986.

6. W. Herrmannsfeldt and H. Hanerfeld, "The Numerical Simulation of Accelerator Components," SLAC PUB 4335, Stanford Linear Accelerator Center, May, 1987.

7. P. Chen and K. Yokoya, "Disruption Effects from the Interaction of Round e^+e^- Beams," SLAC PUB 4339, Stanford Linear Accelerator Center, March, 1988.

8. T. Hughes and B. Godfrey, "Electron Beam Stability in Compact Recirculating Accelerators," MRC/ABQ-R-1040, Mission Research Corporation, March 1988.

9. J. Krall and Y. Lau, "Modulation of an Intense Beam by an External Microwave Source: Theory and Simulation," *Appl. Phys. Lett.* **52**, 431-433, February 1988.

10. A. Mankofsky, "Three-Dimesnional Electromagnetic Particle Codes and Applications to Accelerators," unpublished report, Science Applications Incorporated Corporation, McLean Va.

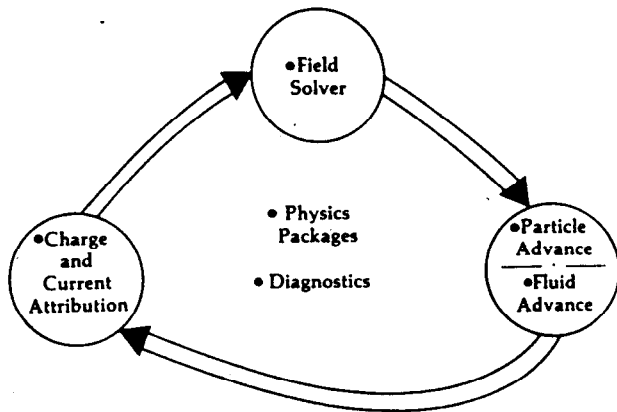11. K. Eppley, "Algorithms for the Self-Consistent Simulation of High Power Klystrons," SLAC PUB 4622, Stanford Linear Accelerator Center, May 1988.

Figure 1. Flow of a classical time-stepping simulation.
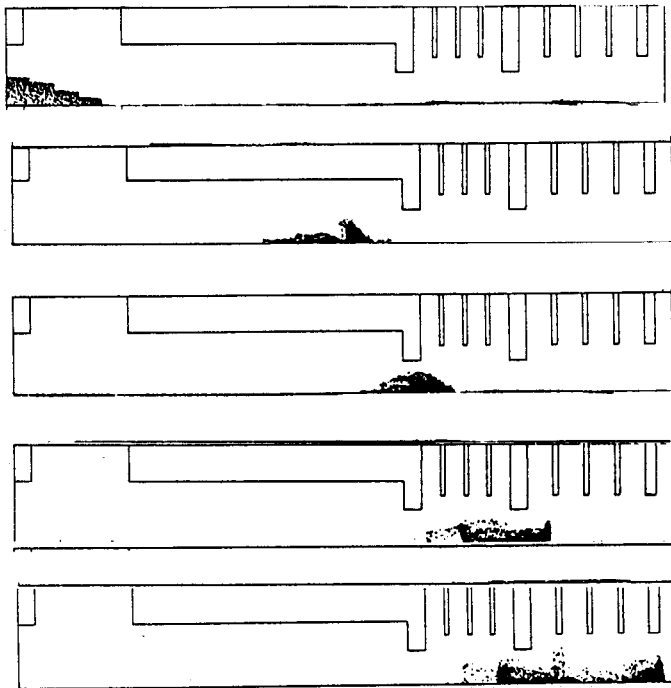
Figure 2. Propagation of a bunch through the SLC subharmonic buncher.
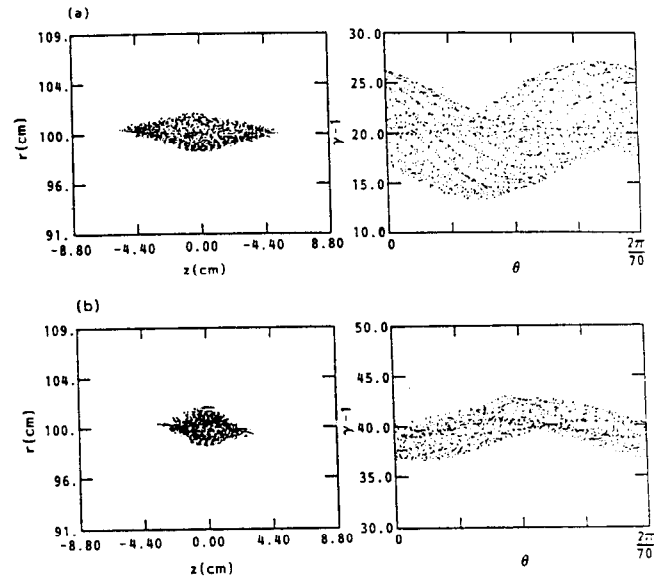
Figure 3. Phase space plots of a resonant negative mass instability in a stellatron.
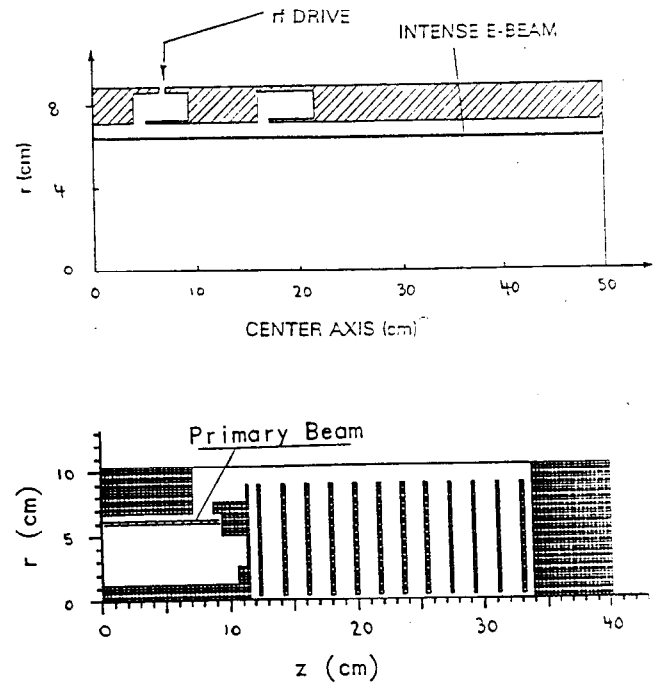
Figure 4. Simulation geometry of a compact accelerator combining RF generator and accelerating structure. The upper figure is a closeup of the modulating section.
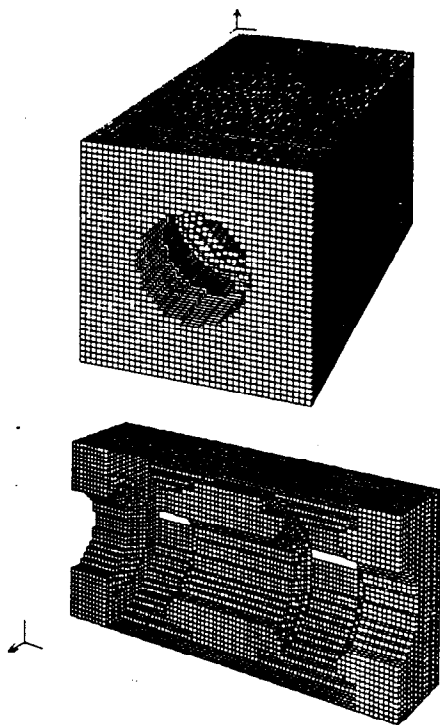
Figure 7. Quadrupole focused heavy ion beam.

Figure 5. Geometry of the SLAC 5045 double-output cavity for 3-D simulation.
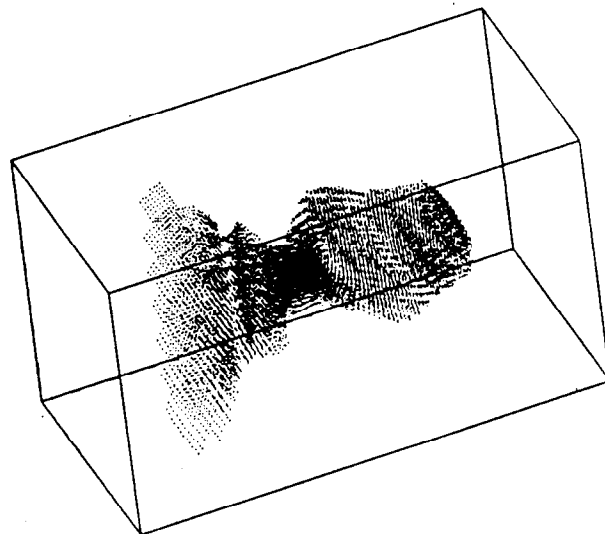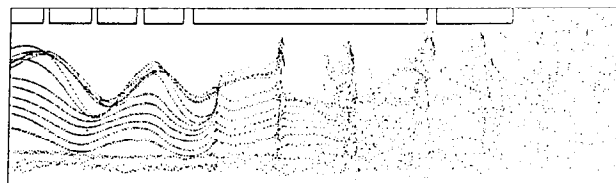

Figure 8. Snapshot of the electron position space distribution for the SLAC 5045 klystron. The horizontal axis is Z and the verticl axis is R. Beam is injected from the left wall. The ports are located at the openings in the blocks in the top wall. The output cavity is just after the final block. Side-emitted electrons enter the penultimate cavity during one part of the rf cycle.

Figure 6. Depressed electron collector structure with vanes.