Title: Advanced Modeling of High Intensity Accelerators

Author(s): Robert D. Ryne, Thomas P. Wangler, LANSCE-1
Salman Habib, T-8

Submitted to: DOE OFFICE OF SCIENTIFIC AND TECHNICAL INFORMATION (OSTI)

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. The Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
Advanced Modeling of High Intensity Accelerators

Robert D. Ryne*, Salman Habib, and Thomas P. Wangler

Abstract

This is the final report of a three-year, Laboratory Directed Research and Development (LDRD) project at the Los Alamos National Laboratory (LANL). The goals of this project were three-fold: (1) to develop a new capability, based on high performance (parallel) computers, to perform large scale simulations of high intensity accelerators; (2) to apply this capability to modeling high intensity accelerators under design at LANL; and (3) to use this new capability to improve our understanding of the physics of intense charged particle beams, especially in regard to the issue of beam halo formation. All of these goals were met. In particular, we introduced split-operator methods as a powerful and efficient means to simulate intense beams in the presence of rapidly varying accelerating and focusing fields. We then applied these methods to develop scalable, parallel beam dynamics codes for modeling intense beams in linacs, and in the process we implemented a new three-dimensional space charge algorithm. We also used the codes to study a number of beam dynamics issues related to the Accelerator Production of Tritium (APT) project, and in the process performed the largest simulations to date for any accelerator design project. Finally, we used our new modeling capability to provide direction and validation to beam physics studies, helping to identify beam mismatch as a major source of halo formation in high intensity accelerators. This LDRD project ultimately benefited not only LANL but also the U.S. accelerator community since, by promoting expertise in high performance computing and advancing the state-of-the-art in accelerator simulation, its accomplishments helped lead to approval of a new DOE Grand Challenge in Computational Accelerator Physics.

Background and Research Objectives

The United States is now involved in efforts to develop accelerator-driven technologies to solve several problems of national importance, which have highly beneficial environmental and economic implications. The technologies include: (1) next-generation spallation neutron sources for materials science and biological science research, (2) accelerator transmutation of waste (ATW), (3) accelerator-based conversion of plutonium (ABC), (4) accelerator production of tritium (APT), and (5) accelerator-driven fission

*Principal Investigator, e-mail: ryne@lanl.gov
energy production (ADEX). These projects utilize accelerators that have energies of 1-2 GeV and average currents of 10-200 mA. The highest power comparable accelerator in the world is at LANSCE and has an average current of only 1 mA. This large increase in average current makes controlling beam loss, which can lead to unacceptably high levels of radioactivation, a major issue for the success of these projects, since reliability, availability and maintainability are crucial.

Very high resolution modeling, far beyond that which has ever been performed by the accelerator community, is necessary to reduce cost and technological risk, and to improve accelerator efficiency, performance and reliability. Though such a modeling capability was not available when this project began, it has now been achieved through the development of software aimed at high performance computing platforms and making use of advanced numerical methods and algorithms. For example, prior to the start of this project, most linac simulations were done with 10,000 to 100,000 simulation particles. Using our new capability, we routinely perform simulations with 1 to 10 million particles, an increase in problem size by a factor of 100. Furthermore, prior to this project, nearly all linac simulations were performed using a two-dimensional (2-D) space charge routine; as a result of our effort, we now perform fully three-dimensional (3-D) simulations on a routine basis. Thus, not only have we dramatically increased the size of linac simulations, we have also developed codes that have increased accuracy and a greater domain of applicability.

Importance to LANL’s Science and Technology Base and National R&D Needs

This project is of direct importance to two of LANL’s core competencies: (1) Theory, Modeling, and High-Performance Computing, and (2) Nuclear Science, Plasmas, and Beams. In regard to the former, the main areas of accelerator simulation, namely computational beam dynamics and computational electromagnetics, share much in common with computational aspects of plasma physics, astrophysics, molecular dynamics, and structural mechanics. The development of new software and algorithms in accelerator modeling is often pertinent to these other fields as well. The converse is also true: the harnessing of advances in computational methods in these other fields has led to a rapid increase in the state-of-the-art in accelerator simulation. In regard to the latter core competency, this work has considerable impact in several areas of importance to the LANL, especially as pertains to DOE programs in Energy Research and Defense Programs. Accurately modeling intense beams is important to the accelerator driven transmutation technologies (ATW, ABC, ADEX), and to APT, with resulting impacts in environmental waste management, energy self-sufficiency, and national defense. In the area of basic and applied research, accelerator modeling is crucial to the design of next-generation spallation
neutron sources, and to accelerators for high-energy physics research such as the Large Hadron Collider and the International Linear Collider. Finally, accelerator simulation is also of importance to Science Based Stockpile Stewardship, as it pertains to performing state-of-the-art simulations in support of proton radiography.

So far we have described the role of accelerators in major projects for solving problems of national importance (waste transmutation, energy production, tritium production), and their role in advancing U.S. science and technology through such projects as spallation neutron sources and colliders. But accelerators also have an impact on national science and technology in a wealth of other ways. Quoting from the well-known accelerator physicist, Prof. Alex Dragt:

“Major existing applications [of accelerator physics theory and technology] include electron microscopy, microprobes, charged-particle-beam lithography, ion implantation, isotope production, particle beams for precision irradiation therapy, superconducting magnets and medical magnetic resonance imaging, neutral-beam heating of plasmas, synchrotron light sources, x-ray lithography, and free-electron lasers. Just this past year synchrotron light was used in the field of biology to determine how meters of DNA can be coiled and managed in cells, to determine the long-sought structure of bacteriorhodopsin, and to determine the largest x-ray crystal structure to date, that of the bluetongue virus made of more than 1000 separate proteins.”

In summary, particle accelerators, and the accelerator simulation needed to design and optimize them, have a profound impact on U.S. science and technology, and great relevance to the missions of the Department of Energy.

Scientific Approach and Accomplishments

General

The most widely used method for modeling intense charged particle beams, as well as plasmas and gravitationally interacting systems, is the particle simulation technique [1,2]. In this approach, one integrates the equations of motion for individual particles taking into account both external fields (if applicable) and the collective field of the system of particles. This collective field is obtained using, for example, the Particle-Mesh method or a Fast Multipole method. These particle simulation techniques have been developed over many years, mainly by members of the computational plasma physics, astrophysics, and cosmology communities. During the past two decades great strides have also been made in the field of linear and nonlinear beam dynamics (also called magnetic optics), the study of charged particle motion in the presence of electromagnetic elements that focus, bend, and accelerator particle beams [3]. However, most of that progress has been in the area of single particle dynamics, i.e. beam dynamics without space charge, the self-fields that are an essential aspect of the physics of intense beams.
Our approach combines aspects of both particle simulation techniques and magnetic optics, by making use of symplectic integration algorithms of the late 1980s and early 1990s [4,5,6]. In particular, so-called “split operator” methods, applied to the Hamiltonian rather than the equations of motion, allow us to combine the best features from decades of research in these two areas. By implementing this approach on high performance computers, we are able, for the first time, to model a beam with very high resolution, taking into account both high order magnetic optics effects and intense space-charge effects. Examples of our work follow.

2-D Particle-in-Cell Code

In the first year of this project, we developed and used a 2-D Particle-In-Cell (PIC) code to model long beams in a variety of focusing systems, including constant focusing, interrupted solenoid focusing, and quadrupole focusing. Our code treats a variety of initial beam distributions, such as Kapchinskij-Vladimirskij (KV), Gaussian and Waterbag. We also generate more general distributions using the rejection method. Our code uses Hockney's convolution algorithm to compute the space charge fields of an isolated bunch of charge, implemented using Fast Fourier Transform routines. To optimize our code’s performance, we implemented the charge assignment and field interpolation algorithms of Ferrell and Bertschinger [7]. One of our early applications of this code was to study beam halo formation in mismatched charged particle beams. Early studies of mismatched charged particle beams showed that such beams could undergo emittance growth and develop a large halo surrounding the beam core [8,9]. Beam mismatch is now believed to be a major source of halo formation. A popular model used to study beam halo is the particle-core model of halo evolution [10-15]. In this model halo particles interact with a beam core that is assumed to oscillate because of an initial radial mismatch; the fields inside the core are roughly proportional to the radial coordinate r, while they are inversely proportional to r outside the core (in the 2-D case). Thus, a halo particle moving in and out of the core sees a strongly time-dependent nonlinear field superposed on a linear external-focusing field. As one might imagine such a system exhibits a variety of dynamical phenomena including chaos. In the particle-core model the core is assumed to have a radial density profile that does not change except in a root-mean-square (rms) sense. In other words, the beam is always a KV distribution, or always a Gaussian, etc., but the rms size of the beam is allowed to change in accord with the rms envelope equations. In addition to this approximate treatment of the core, the model is not self-consistent since the halo particles do not affect the motion of the core. However, as will be shown below results from the particle-core model are in excellent agreement with results from high-resolution particle simulations.
A useful way to study this model is to make a stroboscopic plot, in which test particles are plotted once each cycle of the core oscillation. (This technique was first used in halo studies by Lagniel [11,12].) As an example, consider the case of a mismatched KV beam. Figure 1 shows a stroboscopic plot of 32 test particles that were initialized with 16 on the x-axis and 16 on the p_x-axis. The main features of the plot are: (1) a central region that has an extent somewhat larger than the core radius; (2) a large amplitude region where particles exhibit betatron motion perturbed by the core space charge; (3) a period-2 resonant region associated with the fixed points to the left and right of the central region; and (4) a separatrix with an inner branch that encloses the central region and outer branches that separate the period-2 resonance from the betatron-like trajectories. The period-2 resonance is a parametric resonance corresponding to the fact that resonant particles have an oscillation frequency, which is one half the envelope frequency, as has been shown analytically by Gluckstern [13].

In Figure 1 the separatrix is actually a narrow chaotic band, and the outer edge of the band has the approximate shape of a peanut. This “peanut diagram” provides a useful picture for describing halo formation: if particles in an initially well defined core reach the separatrix (by transport mechanisms not included in the model), then they will be carried to large amplitudes along the outer branch of the separatrix. Also, since a real beam would not have an exactly uniform core, the injected beam could already have a low density tail that extends into the resonance region, and under the dynamics of the model these particles would be carried to large amplitude and form a halo.

Despite its simplicity, the particle-core model in a constant focusing channel predicts a maximum halo amplitude that is in excellent agreement with high-resolution particle simulations run at the LANL Advanced Computing Laboratory [16]. Figure 2 shows particle simulation results for a mismatched KV beam having the same parameters as Figure 1 (a tune depression of 0.5 and an initial beam size that is 0.62 times the matched value). Though the initial distribution has the property that it is uniform in (x,p_x)-space, it is unstable for the parameters chosen, and the resulting phase space of Figure 2 is highly nonuniform. The first curve bounding the chaotic band in Figure 1 is also shown in Figure 2 for comparison. The parallel simulation results show that for this simple configuration (an axially symmetric beam in a constant focusing channel) the maximum particle amplitudes are in excellent agreement with the amplitude of the separatrix in the particle-core model. The numerical validation of the particle-core model was a major step in the process of understanding halo formation in the dynamics of intense charged particle beams.
The study of beam halo and associated particle loss is being pursued in many locations, including Los Alamos, the University of Maryland, Princeton University, MIT, the University of Indiana, Argonne National Laboratory, and in other countries such as France and the Former Soviet Union. The development of the particle-core model caused concern in the halo community, since it led some to conclude that quadrupole focusing could lead to significant halo formation, even for matched beams. This is an extremely important consideration, since virtually all modern linacs and colliders use quadrupoles.

The application of the particle-core model to a periodic focusing channel is much more complicated than in a constant focusing channel. In a constant focusing channel there is one frequency driving the dynamics, namely the frequency of the oscillating core; the same is true for a matched beam in a periodic transport system. But a mismatched beam in a periodic channel does not normally oscillate at a single frequency, and this means that stroboscopic plots, which were so useful in illuminating the underlying physics in the constant focusing case (e.g. the period-2 resonance) are not as applicable here. It should be noted, however, that if one linearizes the mismatch then it is possible to excite a single “even” or “odd” mode by a careful choice of initial conditions.

We used our 2-D code to study halo formation in matched beams in periodic focusing channels. We found that, for beams with moderate intensity (i.e. a tune depression of roughly 0.5), the envelope flutter is not a source of halo formation. Consider for example a matched KV beam in a quadrupole channel with a zero-current phase advance per focusing period of 70 degrees/period, which is depressed by space charge to 35 degrees/period. Figure 3 shows particle simulation results for an initially rms matched Gaussian beam in this channel after 22 periods. A large amplitude halo is not present. In contrast, Figure 4 shows the situation when the initial beam has the horizontal and vertical rms envelopes too small by a factor of 0.6. Now a significant halo is present. It is worth noting that in the matched case the emittance growth (due to charge redistribution) is only 6%, while in the mismatched case it is approximately a factor of 2.

This can be understood by using an analysis based on the particle-core model in a quadrupole channel. A stroboscopic plot for a collection of test particles in such a system is shown in Figure 5; the data points are recorded at the center of each horizontally focusing quadrupole where the matched horizontal beam size is \( x_{\text{edge}} = 4.6 \text{ mm} \) (i.e. \( x_{\text{rms}} = 2.3 \text{ mm} \)). It is clear that the period-2 resonant structure is not present in this case. Though other resonances and weak chaos are present, they do not provide a path by which particles can be transported to large amplitudes. In closing, we remark that, though envelope flutter in matched beams in periodic channels is not an issue for rf accelerators of the type being designed at Los Alamos and elsewhere, there is evidence that it is an important issue for
beams in the space charge dominated regime (i.e. tune depressions of order 0.1). This
effect might be important, for example, in the design of accelerators for heavy ion fusion.

**Direct 2-D Solver**

Though particle methods are the most widely used techniques for modeling beams
and plasmas, they suffer from the presence of statistical noise in the simulation results.
One approach to dealing with this is to use a direct method instead of an approach based on
particles. In a direct solver, one defines the particle distribution function on a grid in phase
space, and the distribution function at the grid points is itself evolved in time. Prior to the
1990s such simulations were limited to 1-D (or crudely modeled 2-D) systems, since these
simulations have a large memory requirement that grows as some power of the number of
grid points. For example, the memory requirement in a 2-D simulation grows as the fourth
power of the grid size in one dimension. But with the advent of large memory, highly
parallel systems, high-resolution simulations are now feasible in 2-D. As an example,
output from a 2-D Vlasov/Poisson code is shown in Figure 6. The initial distribution was a
Gaussian in four-dimensional phase space that was mismatched into a quadrupole channel.
The simulation utilized a $128^4$ grid for a total of 268 million grid points. In contrast, a 3-D
simulation with a grid size of $128^6$ requires 4 trillion grid points, a resource requirement
that will probably be met by the middle of the next decade.

**3-D Systems**

The first year of this project was devoted mainly to the development of 2-D
simulation codes. The second and third years were devoted almost entirely to three-
dimensional systems. A major milestone was the development of a parallel, 3-D space
charge routine. This routine was compared extensively with the widely used 2-D routine
SCHEFF. We showed that, in the regime for which both approaches were valid (namely,
azimuthally symmetric beams), the two codes were in excellent agreement. An example is
shown in Figure 7, which compares the output from two simulations, one using SCHEFF,
and the other using the parallel 3-D routine. Both simulations used 100,000 particles. The
two results are in excellent agreement. It is worth noting that, if the number of particles is
increased to 1 million, there is a modest increase in the maximum particle amplitude. The
use of a 3-D code and the use of a large number of particles would have been difficult in the
1980s, because of the extremely long run times that would have been required. (In these
types of simulations, the run time is dominated by the space charge calculation.) However,
our new space charge routine, which has been optimized for parallel platforms, makes it
possible to perform fully 3-D linac simulations in a reasonable amount of time. For
example, a simulation of the APT linac with 1 million particles can be performed on 128
nodes of an SGI/Cray T3E in about an hour. Another major milestone for us was to
parallelize the widely used code PARMILA, and to include the new 3-D space charge routine in the parallel version of PARMILA. We also made a 3-D, parallel version of the code LINAC, which is one of the main codes used to model the APT linac.

In addition to making parallel versions of existing codes, we have also developed 3-D, parallel codes “from the ground up,” using modern algorithms. Two major codes have been developed: IMPACT and HALO3-D. IMPACT (which stands for Integrated Map and Particle Accelerator Tracking Code) uses the previously mentioned split-operator methods, including a map-based magnetic optics procedure for an especially accurate treatment of rf accelerating gaps. Because of this approach, IMPACT can simulate very complex dynamics inside an rf gap without having to use a large number of (time-consuming) space-charge kicks normally associated with taking very small time steps [17]. IMPACT has been benchmarked extensively against 3-D envelope equations that were derived and implemented in a 3-D envelope code for this project [18].

In the final year of this project, our emphasis remained on 3-D modeling, but it turned to two specific aspects: (1) modeling 3-D equilibria and (2) longitudinal motion inside 3-D bunches in the presence of nonlinear rf forces. In the former, we were able to initialize and numerically propagate beam equilibria whose existence was found analytically by R. Gluckstern of the University of Maryland. Based on simulations using HALO3-D, we found that longitudinal halo developed more quickly than transverse halo. Following this, we made enhancements to HALO3-D that enabled us to treat nonlinear rf focusing fields. This was used to validate a longitudinal particle-core model developed by Lawrence Livermore National Laboratory.

In the final year of this project we also developed a 3-D Langevin code to study the effects of noise and damping in accelerators. Such a capability is useful for modeling effects such as rf noise, power supply noise, ground motion, and synchrotron radiation reaction. It can also be used to numerically drive beams to thermal equilibrium, for which an example is shown in Figure 8. The figure shows the trajectories in longitudinal phase space of 6 test particles moving in the self-consistent field of a beam composed of 2 million particles that is in thermal equilibrium. Several features are worth noting in the figure. First, the total force acting on particles is roughly linear near the beam core (as expected), as demonstrated by the elliptical trajectories at the center of the figure. Furthermore, the space charge becomes very important at larger amplitude, as demonstrated by the test particle whose trajectory has become almost rectangular in shape. Finally, the presence of the nonlinear rf focusing is evident, and is responsible for the loss of the outermost test particle.
Throughout this project, we were faced with having to deal with enormous amounts of simulation data. For example, a double precision array of 100 million particles with 3 coordinates, 3 momenta, and 1 element to describe location of particle loss, requires 7 scalars $\times 10^9$ particles $\times 8$ bytes $= 5.6$ Gbytes of memory. Writing all the data at a few hundred locations would require Tbytes of memory, but naively writing a fraction of the data is not useful when the halo fraction is small. For example, if the fraction of halo particles is $10^{-5}$, then we would have to write $>>100,000$ particles to see an appreciable number of halo particles. Therefore, it is essential to intelligently perform “data pruning” inside the code. We have developed a procedure called a “low density cut” that we have implemented in a parallel environment to reduce the volume of output data and help visualize the halo. In this procedure we choose a threshold density just above that in the halo and plot all the particles in the region where the density is below the threshold. In the high-density region, we use undersampling to plot only the fraction of the particles corresponding to the threshold value. This accentuates the halo while reducing the intensity of the often-structureless “burnt in” core region. Figures 9 and 10 show the longitudinal halo of a mismatched beam, without and with the low-density cut, respectively. In figure 10, the details of the halo are clearly visible, in sharp contrast to figure 9.

**Summary**

In summary, as a result of this LDRD project, we succeeded in developing a new capability using high performance computers combined with modern numerical methods and algorithms for modeling intense charged particle beams in accelerators. This new capability has greatly increased the size and speed of the simulations that we can perform (due to the use of parallel computers), and the scope of the simulations (due to the development of a 3-D space charge algorithm). As a consequence, instead of performing 2-D linac simulations with 10,000-100,000 particles, as was the case when this project began, we are now able to perform fully 3-D simulations with 1-10 million particles. We introduced the use of split-operator methods for beam dynamics simulations, an approach that is likely to have a significant impact in the computational accelerator physics community. Finally, our work has benefited much of the U.S. accelerator community since out of this work and other work has grown a nationwide project, namely a new DOE Grand Challenge in Computational Accelerator Physics.
Publications


References


Figure 1. Stroboscopic phase space plot based on the particle-core model (uniform density core, tune depression=0.5, mismatch=0.62).
Figure 2. Beam phase space from a 2 million particle simulation (65,536 points plotted). The outer peanut-shaped set of points were obtained from the particle-core model shown in Figure 1.
Figure 3. Simulation results showing the beam density after 22 focusing periods in a quadrupole channel. The initial distribution is an rms matched Gaussian beam. Zero-current phase shift/period and depressed phase shift/period are 70 degrees and 35 degrees, respectively.
Figure 4. Simulations results showing the beam density after 22 focusing periods in a quadrupole channel. The initial distribution is a mismatched beam, where the initial beam size is 0.6 times the matched size. Other parameters are as in Figure 4.
Figure 5. Stroboscopic phase space plot based on the particle-core model for a matched, uniform density beam in a quadrupole channel. Zero-current phase shift/period and depressed phase shift/period are 70 degrees and 35 degrees, respectively.
Figure 6. Output from a 2-D direct Vlasov/Poisson simulation (4-D phase space) using 268 million grid points. The beam density is shown on a grid of size 128 x 128.
Figure 7. Results from a simulation of the APT linac using 100,000 particles, obtained using two versions of the code LINAC: (1) a serial version based on the 2-D space charge routine SCHEFF, and (2) a parallel version based on a new 3-D space charge routine.
Figure 8. Output from a parallel, 3-D Langevin code. The figure shows the trajectories in longitudinal phase space of 6 test particles moving in the self-consistent field of a beam composed of 2 million particles that has reached thermal equilibrium.
Figure 9. Output showing the longitudinal phase space based on a 3-D simulation with 25 million particles. The initial distribution is a mismatched Gaussian beam.
Figure 10. Output showing the longitudinal phase space based on a 3-D simulation with 25 million particles. The initial distribution is a mismatched Gaussian beam. In contrast to Figure 9, the data in this figure were preprocessed using a low-density cut to accentuate the details of the halo.