

Achieving the Highest Accuracy With the BEM

Frank H. Read

Emeritus Professor at School of Physics and Astronomy, University of Manchester, UK.

Abstract

The high accuracy that can be achieved by the Boundary Element Method when it is used to solve the Laplace and Poisson equations for electrostatic systems is discussed. Applications to charged particle optics are described, with the emphasis on the commercial CPO programs [1]. The BEM is a charge-based method and so is ideally suitable for systems that include space-charge and/or cathodes. It can deal easily with electrodes of very different sizes. These and other properties of the BEM are illustrated by a range of benchmark tests. .

Introduction and historical context

Accuracy is not always required in modelling systems since optimizations can often be carried out when known or unknown errors are present, provided that these errors remain approximately constant during the optimization process. The present paper is directed instead at those systems for which the highest accuracy is either indispensable or very useful.

The Boundary Element Method (BEM), also known as the Integral Equation Method, has been used in many fields. The first use of it in charged particle optics was by the Atomic Physics group at Manchester University in 1970, to calculate the focal properties of two-tube electrostatic lenses [2]. This early work culminated in a standard work on electrostatic lenses [3] and has continued since then. In this review we shall concentrate on the use of the software initially developed by that group [1,4].

Outline of the BEM

The BEM, as applied to electrostatic systems, is based on the fact that when voltages are applied to conducting electrodes, physical charges appear on the surfaces of those electrodes. These charges are the only sources of all the potentials and fields in the system. So in the BEM the surfaces of the electrodes are effectively replaced by these sheets of surface charge. Only the surfaces have to be modelled and it is not necessary to create a grid of points in the space enclosed by the electrodes, nor is it necessary to enclose a system.

In the version of the BEM used in the CPO programs 3D electrode surfaces are divided into flat rectangular or triangular segments each of which has a uniformly distributed charge over its surface. The flatness of the segments and the uniformity of their charge densities represent the only significant approximations. The segment charges are adjusted to make the potentials at the centres of the segments equal to the voltages applied to the parent electrodes. This adjustment is carried out by setting up the matrix that connects the mid-centre potential of each segment to the charges on all the segments and then inverting the matrix to obtain the charges that exist for the given set of applied voltages. Potentials and fields can then be calculated anywhere in space.

The errors in setting up the matrix are essentially negligible because the potential distributions of uniformly charged 3D triangles and rectangles are known analytically [5] and also the potential distributions of 2D segments with planar or rotational symmetry can be calculated very accurately [6]. Similarly there is essentially negligible error in inverting the matrix. The errors in calculating the final potentials, which are continuous, are also essentially negligible while the corresponding fields are obtained by numerical differentiation when necessary. The errors in trajectory calculation

are also very small because the fields are uniform. The errors due the flatness and charge uniformity of the segments remain of course and methods of reducing these are discussed below.

The BEM is mathematically and computationally more complex than the traditional Finite Difference and Finite Element Methods (FDM and FEM), but it offers several advantages. Perhaps the most important of these is the inherent accuracy of the method. It was shown some years ago [7,8] that for non-space-charge electrostatic 3D problems the CPO programs are typically one or two orders of magnitude more accurate than two representative FDM and FEM programs for the same computing time (or typically one or two orders of magnitude faster for the same accuracy), although this comparison has not been carried out for the more recent versions of all these methods. A comparison has also been carried out [9] for the space-charge simulation of a CERN laser ion source. A disadvantage is that the BEM tends to integrate trajectories more slowly, but a technique for alleviating this is discussed below. Another slight disadvantage is that the matrix referred to above is fully populated and therefore can take a comparatively long time to invert.

Distributing the segments

The computing time increases with the number of segments, which effectively limits the maximum number that can be used and so the first priority in virtually all simulations is to distribute the segments in the most appropriate way and with the most appropriate sizes. This is particularly important for electrodes that have sharp edges or corners, where the charge densities can become very high. To illustrate this Figure 1 shows the potential contours near the tip of a sharp cone. Gaps have been added between the segments into which the cone is divided, to show that the density of segments has been chosen to be highest at the tip.

Figure 1 also illustrates another considerable advantage of the BEM – it can deal easily with small electrodes in the vicinity of much larger ones.

For more general simulations an ‘adaptive segmentation’ option is provided which allows the 3D program to find the best distribution empirically and automatically. The objective is to try to equalise the charges on the segments, which has been found to give the best accuracy for a given number of segments. The technique is iterative and uses m stages to reach the required final number of segments N , at each stage subdividing the segments according to their charges. The user only has to state the values of m and N .

Achieving the desired accuracy of field calculations in the minimum time

Although the potential distributions of uniformly charged 3D triangles and rectangles are known analytically [5] the formulations are complicated and so take a comparatively long time to compute.

To give faster computing times the CPO programs use several different types of approximations where appropriate. As a part of this technique the user is asked to specify the inaccuracy ϵ that is desired for the calculation of potentials, which might typically be 10^{-3} in the initial stages of a new simulation or 10^{-7} for the final results. As an example of the technique we consider the calculation of the potential at a distance s from a triangular segment that has a representative length w . Clearly a key parameter is w/s .

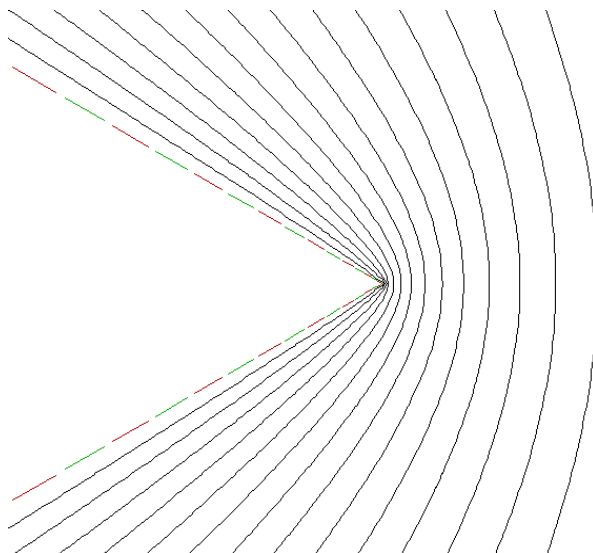


Figure 1. Potential contours near the tip of a cone. These contours are within 0.05% of the analytical values. For display purposes gaps have been added between the segments into which the cone is divided, to show that the density of segments has been chosen to be highest at the tip.

The steps used by the CPO3D program are as follows. In all these steps the criteria for selecting the appropriate approximations have usually been derived analytically but all of them have of course been carefully tested in practice. The relevant parameters for each segment are calculated beforehand and stored.

- (1) If the program judges that w/s is sufficiently small it will use the simplest approximation and treat the segment as a point charge.
- (2) When w/s is not small enough for this level of approximation the program next considers adding the higher-order terms in the expansion of the potential in powers of w/s .
- (3) If this approximation is inappropriate then the program next considers the use of quadrature, using 7 quadrature points for a triangle (9 for a rectangle).
- (4) Finally, if the potential is required at a point that is very near or on the segment then one of the exact expressions (in-plane or out-of-plane) is used for the potential.

In this way the program calculates the potential due to each triangle in the shortest time for the desired inaccuracy. This technique is also used for 3D rectangles. In the case of 2D simulations, in which the segments are circular hoops of finite width, additional types of approximation are available, namely the use of multipole moments (dipole, quadrupole and octopole) and near-axis approximations. Electrostatic fields are also calculated in an analogous way or are obtained by numerical differentiation, using the potentials at the 4 corners of a tetrahedron (in 3D systems) that surrounds the point. Extra care is taken when calculating the field at points that are near to electrode surfaces, if necessary dividing segments into smaller parts.

Errors due to the flatness and uniform charge densities of the segments

Dealing firstly with the consequences of flatness, when cylindrical electrodes are subdivided into rectangular segments then the segments touch the cylindrical surface at two of their edges but otherwise lie inside the volume of the cylinder. The mean distance of the segments from the axis is therefore less than the radius of the parent cylinder. The same applies when a spherical surface is subdivided into triangles. We refer to this as an ‘inscribing’ error. The programs automatically correct for this by increasing the radii by the appropriate factors.

Errors due to the assumption that each segment carries a uniform charge density obviously decrease when the number of segments is increased but the maximum number of segments is limited in practice by the available computer memory and the computing time. An extrapolation technique is therefore recommended, as illustrated in some of the supplied data files. For example, it is found that when calculated focal lengths of lenses are plotted against $1/N^2$, where N is the total number of segments, the dependence is very close to linear and so that the intercept at $1/N^2 = 0$ effectively gives the value for $N = \infty$. Analogous linear dependences N^p are frequently found for quantities such as potentials, fields, lens parameters or capacitances [10] (see also the CPO inbuilt help), where the empirical value of the exponent p is usually 2, but can also be 1 or 1.5. No theoretical treatment of these exponents is known to the author.

Trajectory integration

Trajectory integration is very accurate with the BEM because the potential and field distributions are continuous. It can however be slow, because all the segment charges are involved. Parallelisation would of course alleviate this problem, but has not yet been implemented. However a ‘mesh’ method is available in which a grid of mesh points is created, as required, in the volume traced out by the beam. This mesh technique is most useful when the number of rays is large and/or when a beam is long and thin.

Other versions of the BEM

Other versions of the BEM also exist. Hawkes and Kasper [11] give a detailed review and extensive references, particularly for 2D systems. Here we add only that attempts to improve the BEM for 3D systems have included using non-uniform charge distributions on the segments (for example [12]) and replacing flat segments with curved ones (for example [13]), but clearly these imply increased complexity and longer computing times. The experience of the author in the early years of the CPO programs was that a given inaccuracy is obtained more easily in a given computing time by simply increasing the number of segments, where this is possible. An exception is in the treatment of systems that include dielectric materials, (see below) since it has been found that only then can the difficult boundary conditions at dielectric interfaces be achieved, and so in such cases the CPO programs use linearly varying charge densities on the relevant segments.

Illustrations of accuracy and flexibility

‘Benchmark tests’ play an important part in most fields of engineering. These are usually defined as tests of systems that have known analytical solutions, so that the accuracy of the results of simulations can be definitively measured. 64 ‘benchmark tests’ are provided by CPO to illustrate the accuracy of the programs. They include an ideal quadrupole mass filter, space-charge repulsion of a beam that initially converges to a point, space-charge limited planar diode with thermal energies, Pierce gun with planar and cylindrical symmetry, stochastic scattering with Boersch energy spreading and field at a circular hole in an infinite sheet.

There are 145 other data files that illustrate the versatility of the programs. They also serve as templates for new simulations. They include automatic iteration to optimise the focus of a 3-cylinder lens, creation of secondary rays with Maxwell and Poisson distributions, user-supplied scattering including absorption and scattering by background gas, sputter ion source, channel electron multiplier with tilted channels, and space charge repulsion in a pulsed beam.

A set of 18 data files illustrate a powerful new ‘users equations’ option has been added recently. This allows users to define the shapes of electrodes by simply typing their own equations into the databuilder. They include: helical coil, corrugated cylinder, ideal quadrupole mass filter, toroidal

surfaces, surfaces with conic sections, elliptical hole in circular electrode, and single or multiple circular or square holes in square or circular plates.

Dielectric materials, cathodes and space-charge

As a boundary method the technique is also ideal for dealing with the interfaces that involve dielectric media. The CPO programs use an improvement of the algorithm described by Murata et al [14]. Because the BEM is a charge-based method one its strengths is in dealing with systems that include space-charge and/or cathodes. One of the simplest benchmark tests for cathodes is that of the planar Pierce gun with zero cathode temperature. This is illustrated in Figure 2. Here the cathode is infinite perpendicular to the plane of the figure, while the focus electrode that surrounds the cathode is inclined at the 'Pierce angle' of 67.5° . Ray tracing starts at the edge of a user-defined 'cathode region', using Childs law and simulating the charge density distribution inside that region. Each ray moves in the total electric field as if it were a single electron but carries the charge of many adjacent electrons. After 7 iterations the current converges to within 0.5% of the correct value and the beam is accurately parallel, as required for the Pierce gun. In each iteration the rays move in the total space-charge of the previous iteration, which is saved either by the space-charge tube method [15], as used here, or by the more conventional 'charge in cell' method.

In practice the temperature of a thermionic cathode is not zero: a potential minimum (the virtual cathode) exists in front of the cathode and the Langmuir relationships have to be used. The technique used in the CPO programs for these simulations is described in Ref. [16].

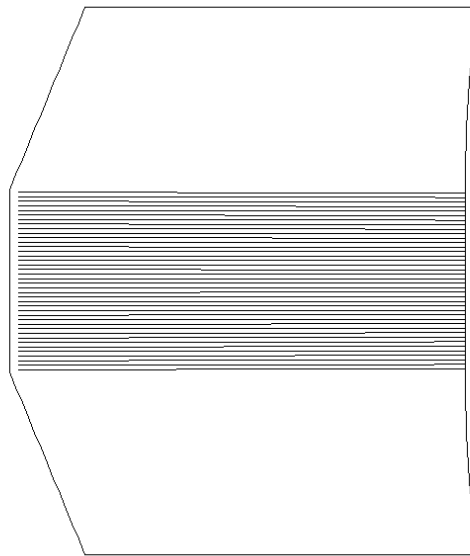


Figure 2. Simulation of a planar Pierce electron gun, $kT = 0$. The cathode is infinite perpendicular to the plane shown. The focus electrode is inclined at the 'Pierce angle' of 67.5° . Ray tracing starts at the edge of a user-defined 'cathode region'.

References:

- [1] Available at www.electronoptics.com, including free demonstration version.
- [2] FH Read, A Adams and JR Soto-Montiel, *J. Phys E: Sci. Instrum.* **4** (1970), p. 625.
- [3] E Harting and FH Read, "Electrostatic Lenses", (Elsevier, Amsterdam), 1976.
- [4] FH Read and NJ Bowring, *Nucl. Instr. Meth. A* **645** (2011), p 273.
- [5] AB Birtles, *Proc. IEEE* **120** (1973), p 213.

- [6] A Renau, FH Read and JNH Brunt, *J. Phys E. Sci. Instrum.* **15** (1982), p 347.
- [7] D Cubric, B Lencova and FH Read, *Electron Microscopy and Analysis 1997*, *Inst. Phys. Conf. Ser.* **153** (1997), p. 91.
- [8] D Cubric *et al*, *Nucl. Instr. Meth. A***427** (1999), p. 357.
- [9] K Hanke *et al*, *Rev. Sci. Instrum.* **73** (2002), p. 783.
- [10] FH Read, *International Series on Advances in Boundary Elements 8*, *Boundary Elements XXII*, eds. CA Brebbia and H Power, (WIT Press, Southampton) (2000), p. 139.
- [11] PW Hawkes and E Kasper in “Principles of Electron Optics” (Academic Press, New York).
- [12] Y Kuno and Y Uchikawa, *IEEE T. Magn.* **21** (1985), p. 2523.
- [13] H Tsuboi, T Takayama and K Yano, *IEEE T. Magn.* **35** (1999), p. 1123.
- [14] H Murata, T Ohye and H Shimoyama, *Nucl. Instrum. Meth. A* **519** (2004), p. 184.
- [15] FH Read, A Chalupta and NJ Bowring, *SPIE Proc. Ser.* **3777** (1999), p 184.
- [16] FH Read and NJ Bowring, *Nucl. Instrum. Meth. A* **531** (2004), p 407.