Exploring focal and aberration properties of electrostatic lenses through computer simulation

This article has been downloaded from IOPscience. Please scroll down to see the full text article.
2008 Eur. J. Phys. 29 1165
(http://iopscience.iop.org/0143-0807/29/6/005)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 159.149.103.9
The article was downloaded on 21/02/2012 at 17:07

Please note that terms and conditions apply.
Exploring focal and aberration properties of electrostatic lenses through computer simulation

Omer Sise\textsuperscript{1}, David J Manura\textsuperscript{2} and Mevlut Dogan\textsuperscript{1}

\textsuperscript{1} Department of Physics, Science and Arts Faculty, Afyon Kocatepe University, 03200 Afyonkarahisar, Turkey
\textsuperscript{2} Scientific Instrument Services, Inc., Ringoes, NJ, USA

E-mail: omersise@aku.edu.tr

Received 26 June 2008, in final form 4 August 2008
Published 1 September 2008
Online at stacks.iop.org/EJP/29/1165

Abstract
The interactive nature of computer simulation allows students to develop a deeper understanding of the laws of charged particle optics. Here, the use of commercially available optical design programs is described as a tool to aid in solving charged particle optics problems. We describe simple and practical demonstrations of basic electrostatic lens properties, such as the determination of focal points, magnification and aberration coefficients, as well as illustrate basic theorems (the Helmholtz–Lagrange law and Snell’s law) and plotting typical lens curves. Throughout this investigation, students are encouraged to make connections between light and charged particle optics.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Charged particle optics is usually introduced in a graduate or final year undergraduate level specialist course. The topic generally covers electrostatic and magnetic lens systems for efficient transport of particles and analyser systems for the separation of charged particles by energy, mass and momentum. The theory of particle optics provides a valuable source of information for students working on physical or biophysical problems as well as biologists who need to understand the mode of image formation in electron microscopes [1]. Perhaps, more importantly, the student who has completed the general physics course may gain greater understanding of the similarities and differences between light and charged particle optics. Understanding of charged particle optics is also important for intelligent operation of other optical instruments, such as cathode ray tubes, accelerators, time-of-flight mass spectrometers, electron/ion guns and energy analysers [2]. Practical problems in charged particle optics are usually solved by means of computer simulations.
Do computer simulations help students learn charged particle optics? We have indeed a number of teaching tools such as demonstration experiments, analytical solutions and computer simulation. Though experimental apparatuses can secure a place in any teaching laboratory for demonstration experiments [3], many of these experiments require intricate setup that can be costly, and some measurements on a system can even be inaccessible. Analytical models help tremendously in understanding, but few problems in charged particle optics admit analytical solutions, and this lack leads to the omission of interesting and important problems in the conventional physics curriculum. Often one then turns to computer simulation [4, 5]. Computer simulations make it possible to explore physical situations where conducting the real experiment is impractical. Students have a direct visualization of the entire path of a particle through an electrostatic field, and they can ‘measure’ what happens to a particle when the potential is varied. Furthermore, by adjusting the boundary values of the computer model the student may gain insight into deviations from analytical expectations.

An excellent example for demonstrating the treatment of a conceptually simple problem of charged particle optics is an electrostatic lens system [6, 7]. The term electrostatic lens is used because it exhibits analogous behaviour to light lenses with which students are more familiar. The optical properties of electrostatic lens systems are often simulated by various charged particle optics packages, such as LENSYS [7], SIMION [8] and CPO [9], and determined by the way in which the electric field affects the trajectories of the particles. Therefore, the focal points, lens magnification and aberration coefficients can be determined from direct ray-tracing results.

In this paper we will use one of these simulation programs, SIMION (because it is readily available to authors), and illustrate its use as an educational tool with just a few of the many optical principles of electrostatic lens systems. SIMION allows the starting position and velocity of the particle to be defined in a flexible way, and the optical information can be extracted by investigating the trajectories of a chosen beam of particles. We propose that SIMION-based activities are well suited to learn charged particle optics. In a university course, SIMION might not be the primary topic but rather it may assist the course in various ways. For example, the students might operate SIMION in a lab or assignment, or the instructor might show screenshots or operation of an example that illustrates a topic currently being studied (see, for example [10, 11]).

2. Calculation of potential and trajectories

The most important function of a computer simulation program is the determination of the electric potential \( \phi(x, y, z) \) and field for all points in some region of space between electrodes. Since the electric potential and field in a lens are not available in analytic form (except some special cases), it is necessary to obtain these components from the electrode shape and potential through a numerical approximation.

It is known from classical electrodynamics that in the absence of space charge the scalar potential satisfies Laplace’s equation [12]:

\[
\nabla^2 \phi = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} = 0,
\]

with the boundary conditions imposed by specifying the values of the potential \( \phi \) on the electrodes (i.e. Dirichlet conditions) and the value of its normal gradient \( \frac{\partial \phi}{\partial n} = n \cdot \nabla \phi \) at the open region borders (i.e. Neumann conditions) [13]. Laplace’s equation simplifies considerably for systems with rotational or plane symmetry since the number of independent
dimensions falls from three to two. In particular, for an axially symmetrical field, in which $\partial \phi / \partial \theta = 0$, the equation becomes

$$\frac{\partial^2 \phi}{\partial z^2} + \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \phi}{\partial r} \right) = 0$$

(2)

in the cylindrical polar coordinates, with the $z$-axis coinciding with the axis of symmetry.

Laplace’s equation is an elliptic partial differential equation with physical applications beyond electrostatics. The study of the solution of Laplace’s equation with boundary conditions is one of the most fundamental topics in intermediate and advanced courses on electromagnetism [14], but the techniques for numerical analysis of the potential distribution are usually not encountered by students until they have had the full calculus sequence as well as a complete introduction to ordinary differential equations. A wide variety of methods of obtaining the potential distribution are available, and most of these have been applied in charged particle optics. The most successful involves replacing Laplace’s equation by finite-difference or finite-element equations, obtaining the charge distribution on the boundaries, or matching the boundary conditions to a suitable functional expansion. For a comparison of the performance of different methods, the reader is referred to [7].

The specific method used within SIMION is a finite-difference technique which can be especially easy to implement [15]. In this method, the Laplace’s equation is converted into a set of linear, simultaneous finite-difference equations. To do this, the area of interest enclosed by the electrode geometry and other boundary conditions is divided into a grid of successive nodes that is usually evenly spaced. For a two-dimensional system, the lower-left vertex of each node is characterized by the index $i$ and $j$ which indicates the coordinates $x_i$ and $y_j$. The three-point approximation is used for the derivatives $\partial^2 \phi / \partial x^2$ and $\partial^2 \phi / \partial y^2$ in equation (1):

$$\frac{\partial^2 \phi}{\partial x^2} \approx \frac{\phi_{i+1,j} - 2\phi_{i,j} + \phi_{i-1,j}}{h^2}, \quad \frac{\partial^2 \phi}{\partial y^2} \approx \frac{\phi_{i,j+1} - 2\phi_{i,j} + \phi_{i,j-1}}{h^2},$$

(3)

where $h$ is the grid spacing. Substituting equation (3) into equation (1), the explicit numerical approximation for the potential at a point $(i, j)$ is [7]

$$\phi_{i,j} = \frac{1}{4} (\phi_{i+1,j} + \phi_{i-1,j} + \phi_{i,j+1} + \phi_{i,j-1}).$$

(4)

The beauty of equation (4) is that it is intuitive for the beginning student to understand that the potential at the centre of four symmetric points is given by the average of the potentials of the points. What is not so obvious to the student is that when we generalize this algorithm to a lattice of more than five points it is necessary to iterate the lattice and to judge the convergence of the potential values. This fact motivates the so-called relaxation methods for the potential determination. One simply performs a point-to-point scan over the entire grid, updating the values of $\phi_{i,j}$ at each point by the average. The objective is to obtain the best estimate of the potentials for those points within the array that depict non-electrode regions. The problem is solved by iteratively calculating the potential at each node from its nearest neighbours until the residual after each calculation becomes small enough. It should also be noted that even using the maximum of grid points, the edges of the curved electrodes can never be modelled precisely, which leads to the errors in the estimated electric field distribution inside the lens.

The trajectories of charged particles are calculated by integrating the system of Newton’s equations. First, electrostatic forces must be calculated based on the current position and velocity of the particles. Specifically, the negative gradient of the potential at any point equals the electric field at that point, $E = -\nabla \phi$, which in turn defines the force/acceleration of a charged particle at that point. These forces are then used to compute the current particle acceleration, which is then used by numerical integration techniques to predict the position and velocity of the particle at the next time step. In SIMION, a highly modified fourth-order
Runge–Kutta technique, which is a popular computational method for solving higher order differential equations, is used for numerical integration of the particle’s trajectory in three dimensions (3D).

3. Electrostatic lens systems

3.1. Optical analogy

The formalism used in particle optics comes from the analogy of light optics. Therefore, the name ‘electron optics’ or more generally ‘charged particle optics’ derives from the fact that the concepts, the formalism and the trajectory calculations have many features in common with the classical techniques that apply to the passage of light rays through glass lenses and prisms [2]. Because of this and because optics formalism has been highly developed, it is customary in charged particle optics discussions to make use of the same terminology and formulae [6, 7]. Terms such as focal length, linear and angular magnifications, aberrations (spherical, chromatic and off-axis), principal planes and object and image distances have a one-to-one correspondence between the two fields.

In light optics, the law for the refraction of a ray of light in passing across a plane separating two different regions is defined by Snell’s law. For charged particle optics there is an analogue of Snell’s law. Consider a plane that defines a boundary between regions of potentials $V_1$ and $V_2$. In practice it is impossible to have a sharp change in potentials like the case as in light optics since it corresponds to infinite fields at the plane; nevertheless, we can approximate this with a gradual potential change over a very short distance using two ideal grids as shown in figure 1(a). Now, if a particle in the region of potential $V_1$ crosses the boundary at an angle $\alpha_1$ normal to the boundary, it enters the region with potential $V_2$ at an angle $\alpha_2$ normal to the boundary, according to the following relationship: $\sin(\alpha_1)/\sin(\alpha_2) = (V_2/V_1)^{1/2}$. The square root of voltage could be interpreted as a refractive index [16].

However, there are some significant differences between particle and light optics. In light optics there is one refractive surface when the ray passes to another region; however, in particle optics, there are an infinite number of equipotential surfaces which deviate the beam of charged particles at different regions. The other difference is the effect of space charge, due to the mutual repulsion of the charged particles, on image formation. An excellent summary of this and other limitations of the analogy between light and particle optics is provided in [17].

The focal and imaging properties of an electrostatic lens are described analogously by the positions of the focal points and the intersections of the principal planes with the axis. A thick lens representation of the cardinal points and the focal and mid-focal lengths of a lens is shown in figure 1(b). Here, $R$ is the reference plane that is usually chosen to be the mechanical symmetry plane of the lens, $H_1$ and $H_2$ are the first and second principal planes, $F_1$ and $F_2$ are the first and second principal foci and $P$ and $Q$ are conjugate object and image distances, respectively. The focal lengths, $f_1$ and $f_2$, are the distances from the principal focal points to the principal planes. The linear magnification, $M$, relates the size of the image to the size of the object. From figure 1(b) it can be seen that the linear magnification is given simply by the ratio of the final to the initial beam diameter in the radial axis, $r_2/r_1$.

Useful relationships that can be derived from the lens geometry shown in figure 1(b) include

$$M = -\frac{f_1}{P - F_1} = -\frac{Q - F_2}{f_2}, \quad f_2 = \sqrt{\frac{V_2}{V_1}} f_1.$$  \hspace{1cm} (5)

This equation could then be generalized to the case of more than two lens elements [18].
Figure 1. (a) Illustration of Snell’s law in particle optics. (b) Schematic representation of the four cardinal points (red) and the focal and mid-focal lengths of an electrostatic lens. This example shows the procedure for the computation of cardinal points via ray-tracing. It needs to trace three particles: first (red) starts at the object position \( P \), on-axis and is directed at a small angle relative to the axis. It crosses the axis at the image position \( Q \). This allows the determination of \( Q \).

Second (blue) is the first principle ray. It starts at the object position \( P \), off-axis (positive \( y \)) and is directed parallel to the axis towards the image. It crosses the axis at the second focal point \( F_2 \) and finally reaches the image \( Q \). This allows the determination of \( F_2 \).

Third (green) is the second principle ray. For simulation purposes, we let it start at the image position \( Q \), off-axis (negative \( y \) where first principle ray crossed) and be directed parallel to the axis towards the object. It crosses the axis at the first focal point \( F_1 \) and finally reaches the object \( P \) at the starting point of the first principle ray. This allows the determination of \( F_1 \). (c) Illustration of the Helmholtz–Lagrange law.
Equally important in electron optics is to understand how the angular divergence of an electron beam will change during the image formation process. The so-called angular magnification is then given by $M_\alpha = \alpha_2/\alpha_1$. It is interesting, and also a very important result, that if we multiply together the two equations for the angular and linear magnifications, the product is always equal to $MM_\alpha = (V_1/V_2)^{1/2}$. This is the Abbe–Helmholtz sine approximation to the Helmholtz–Lagrange law described by

$$r_1\alpha_1\sqrt{E_1} = r_2\alpha_2\sqrt{E_2}, \quad (6)$$

where $E_1$ and $E_2$ are the energy values of the particles in the object and image planes perpendicular to the beam direction, while $r_1$ and $\alpha_1$ are the respective transverse displacements and angles (see figure 1(c)). This equation is fundamental for focusing paraxial rays of charged particles and should be accurate to better than 1.5% for angles $\alpha < 17^\circ$.

### 3.2. Electron optical properties

An example of a simple electrostatic lens system is shown in figure 2. This is a two-element lens consisting of two coaxial cylinders of the same diameter $D$ and separated by a distance $G$. Traditionally, the cylinders’ dimensions are measured in units of the inside diameter $D$, and the gap width $G/D = 0.1$. The cylinders, which are made of conductive material, are held at potentials $V_1$ and $V_2$ with respect to a reference potential that corresponds to the zero of particle kinetic energy; i.e., the reference is chosen such that a particle of charge $q$ will have kinetic energy $qV$ when it is in a region of potential $V$. Therefore, the first electrode is placed at the same potential (1 V) with respect to the electrons’ primary energy ($E = 1$ eV). The second electrode is placed at a higher or lower potential, thus providing acceleration or deceleration of the beam.

There are many configurations of electrostatic lenses, from the two aperture lens and the two cylinder lens up to the multi-electrode lenses (see [18, 19] and references therein). One of the most commonly used lenses in charged particle optics is the two-electrode lens, which is often used when the image and object are required to be in space of different potential. The electron-optical properties of a two-element lens can be presented in a diagram like figure 3, showing the image position corresponding to a given object distance, with the acceleration ratio ($V_2/V_1 > 1$) as a parameter (the corresponding data for retarding lenses ($V_2/V_1 < 1$))
Figure 3. Typical $P$–$Q$ curve for a two-element lens. The gap between the two lens elements is $G = 0.1D$, where $D$ is the internal diameter of both cylinders. The object position ($P$) and image position ($Q$) relative to the reference plane ($R$) are computed and plotted for chosen combinations of the lens voltage ratio ($V_{2}/V_{1}$) and lens linear magnification ($M$).

In order to focus a beam without changing its final energy, a lens with at least three electrodes is necessary. A three-electrode lens where the potential of the first and last electrodes are equal ($V_{3}/V_{1} = 1$) is called an Einzel lens.

In an acceleration or deceleration lens, it is very often desirable to be able to keep the image of a given object fixed when the acceleration/deceleration ratio and magnification are changed. A lens operated in this way is usually referred to as a zoom lens. Three-element electrostatic zoom lenses are widely employed in electron and ion devices. The best-known systems of this type are rotationally symmetric and consist of a series of cylinders [20]. The focusing properties of a typical three-element cylinder lens can be illustrated by a different type of plot. Such lenses are often used with a fixed object distance $P$ and a fixed image distance $Q$, and to maintain these fixed values, $V_{2}/V_{1}$ has to be varied as $V_{3}/V_{1}$ is varied as shown in figure 4. This figure also shows the magnification of the same lens plotted as
Figure 4. Zoom lens curve for a three-element lens consisting of three equi-diameter tubes, where the length of the middle tube is equal to the tube diameter. The gap between the lens elements is $G = 0.1D$.

a function of $V_3/V_1$. The optimization of this type of lens is not as straightforward as in the two-element cases. Harting and Read [21] have evaluated figures of merit that help in evaluating the performance of such lenses and have presented several examples of the type of plot shown in figure 4.

There seems to be no reason to use more than three different electrodes in a basic unit either for focusing or for acceleration/deceleration. If more degrees of freedom are desirable, it is probably better to use some combination of such basic units separated by field-free regions than to design a lens where more electrodes are closely spaced. The properties of the combination of lenses can also be calculated using the transfer matrix method [22].

3.3. Paraxial approximation and aberrations

The basis of light optics theory is the paraxial approximation, which applies whenever the angle and distance between the system’s optical axis and the ray of interest are small. This allows the use of the small angle approximations ($\sin(\alpha) \approx \alpha$, $\tan(\alpha) \approx \alpha$ and $\cos(\alpha) \approx 1$) when tracing the path of the ray through an optical system (such as a lens). In elementary electron optics books, it is generally stated that an electrostatic lens focuses the beam of
electrons that are parallel to the optical axis to a single point defined as the focal point of the lens. Immediately before or after this sentence, it is usually admitted that it is true only in the limit of paraxial approximation. However, when the electrons are not moving close to the axis, then the basic approximation begins to fail and aberrations start to form.

In rotationally symmetric electrostatic lenses, there are six basic aberrations in two classes. Chromatic aberration is energy-dependent, while monochromatic aberrations (spherical aberration, coma, astigmatism, field curvature and distortion) are independent of energy as for light lenses. Spherical aberration is an axial aberration, one that affects the entire field equally, including the centre of the field. This aberration occurs when a change in the angle at which the electron beam enters the lens causes the focal point of the beam to move along the optical axis. The other four aberrations are off-axis aberrations (the chromatic aberration also has off-axis terms), affecting the beam images increasingly towards the edge of the field but not at the centre. Spherical aberration is probably the most important of all the geometric aberrations. It is, for example, responsible for limiting both the screen spot size in a cathode ray tube (oscilloscope) and, along with diffraction effects, the resolving power of an electron microscope. The reason for its importance lies largely in it being the only geometric defect that is present even for axial objects.

A crude light-optical analogy of the electrostatic lens system is sketched in figure 5. From this figure it can be seen that spherical aberration alters the effective focal length of the lens, forming not a point focus, but a spot. Thus the paraxial approximation has to be modified in order that spherical aberration may be taken into account. In addition, we here illustrate the well-known fact that in the presence of spherical aberration the beam is narrowest (approximately by a factor of 4) at a position before the Gaussian image plane. The image blur in this plane is usually called the disc of least confusion. In the electron-optical case, this has the effect that the transmission is highest for an electron energy slightly higher than the paraxially focused one. This is of some importance in the choice of operating potentials. Similar light distributions were obtained by Carpena and Coronado [23] by changing the screen position to illustrate the effect of spherical aberrations in a planar-convex light lens.

The spherical aberration can be characterized by the third-order coefficients $C_s$ defined by the relation

$$\Delta r = -MC_s \alpha_0^3,$$  \hspace{1cm} (7)

where $\Delta r$ is the radius of the disc formed in the Gaussian image plane by non-paraxial rays starting from an axial object point with a maximum half angle $\alpha_0$ and $M$ is the linear magnification [7].

While spherical aberration is by far the most prominent, it should also be noted that chromatic aberration can become significant in some cases. The particles in a beam will vary some in velocity, so particles with slightly different energies ($\delta E$) get focused at different image planes, and the focal point becomes blurred. The coefficient of chromatic aberration $C_c$ is defined by [7]

$$\delta r = -MC_c \alpha_0 \frac{\delta E}{E_0}.$$  \hspace{1cm} (8)

According to these equations, both types of aberrations can be minimized by reducing the convergence angle of the system so that the charged particles are confined to the centre of the lenses.

In order to illustrate the aberration effects, we calculated the point spread distribution for the electron beam. These $3 \times 3$ patterns are obtained by plotting the positions of the directly computed rays in the Gaussian image plane (figure 6(a)). If the electrons are started at the
object plane with systematic initial conditions, then a spot diagram similar to that shown in figures 6(b)–(d) can be obtained. In this example, rays are launched from nine points in the object plane with initial angles that describe concentric cones. In a perfect imaging system all rays from a point on the object would converge to a point in the image. The contours around the points in the image show the extent of the chromatic (figure 6(b)) and spherical (figure 6(c)) aberration at that location as a function of electron energy and beam semi-angle $\alpha_0$. If the electrons are launched far from optical axis ($0.01D$) in the object plane, with initial energies and directions, then a scatter plot similar to that shown in figure 6(d) can be obtained. This shows the importance of the off-axis aberrations. In order to find numerically the aberration coefficients $C_r$ and $C_c$, the aberration discs $\Delta r$ and $\delta r$ are recorded as a function of $\alpha_0$ and $\delta E/E_0$, respectively.

Third-order aberrations, though relatively straightforward to correct in light optics, are larger and essentially impossible to correct in rotationally symmetric electrostatic lenses [24].
Figure 6. Typical graphical aberration pattern with systematic initial conditions. (a) Bunches of electrons are traced through the focusing fields, from object plane to image plane. (b)–(d) Spot diagrams of the blur at the image plane for chromatic, spherical and off-axis aberrations, respectively.

Since the Laplace’s equation restricts the distribution of electric potentials, we cannot shape and combine equipotentials in arbitrary fashion to correct for these aberrations. It is therefore important to know good values for the aberration coefficients for a wide variety of electrostatic lenses so that one can select the best lens for a given application.

4. Conclusions

Computer simulations enable the trajectories of the charged particles to be visualized in 3D as they pass through the system. They also show directly the formation of an image and the associated aberrations in a lens, which can be difficult or tedious to evaluate analytically. In this paper, we presented some examples of the electron-optical properties of electrostatic lens systems by using the direct ray-tracing method. We also presented some simulations showing the electron distributions in screens placed at different distances from the lens to illustrate the effect of spherical aberration. This can help the understanding of the limit of validity of the imaging properties of electrostatic lenses.
The questions we want the students to address through simulation of the electrostatic lenses are as follows. How do you minimize the aberrations of the lens given a source located an arbitrary image distance away? How do you increase the imaging quality of the lens? What impact do the lens voltages have on the magnification? How do you calculate the potential distributions and particle trajectories in the lens? What does the magnification mean?

Acknowledgments

The authors would like to thank the referee for offering useful suggestions. This work is supported by the State Planning Organization in Turkey, through grant 2002K120110, and by TUBITAK, through grant 106T722.

References