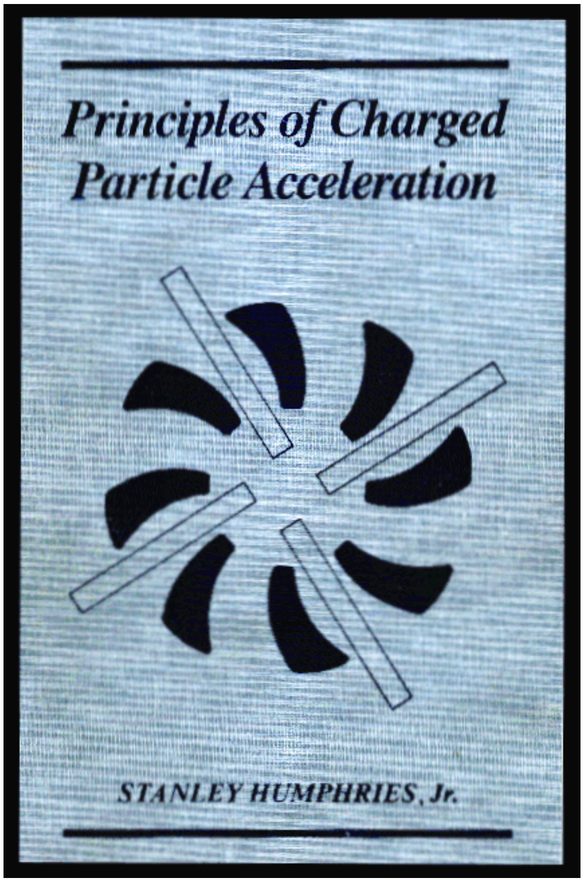


Principles of Charged Particle Acceleration

Stanley Humphries, Jr.

Department of Electrical and Computer Engineering
University of New Mexico
Albuquerque, New Mexico

QC787.P3H86 1986, ISBN 0-471-87878-2

To my parents, Katherine and Stanley Humphries

Preface to the Digital Edition

I created this digital version of **Principles of Charged Particle Acceleration** because of the large number of inquiries I received about the book since it went out of print two years ago. I would like to thank John Wiley and Sons for transferring the copyright to me. I am grateful to the members of the Accelerator Technology Division of Los Alamos National Laboratory for their interest in the book over the years. I appreciate the efforts of Daniel Rees to support the digital conversion.

STANLEY HUMPHRIES, JR.

University of New Mexico
July, 1999

Preface to the 1986 Edition

This book evolved from the first term of a two-term course on the physics of charged particle acceleration that I taught at the University of New Mexico and at Los Alamos National Laboratory. The first term covered conventional accelerators in the single particle limit. The second term covered collective effects in charged particle beams, including high current transport and instabilities. The material was selected to make the course accessible to graduate students in physics and electrical engineering with no previous background in accelerator theory. Nonetheless, I sought to make the course relevant to accelerator researchers by including complete derivations and essential formulas.

The organization of the book reflects my outlook as an experimentalist. I followed a building block approach, starting with basic material and adding new techniques and insights in a programmed sequence. I included extensive review material in areas that would not be familiar to the average student and in areas where my own understanding needed reinforcement. I tried to make the derivations as simple as possible by making physical approximations at the beginning of the derivation rather than at the end. Because the text was intended as an introduction to the field of accelerators, I felt that it was important to preserve a close connection with the physical basis of the derivations; therefore, I avoided treatments that required advanced methods of mathematical analysis. Most of the illustrations in the book were generated numerically from a library of demonstration microcomputer programs that I developed for the courses. Accelerator specialists will no doubt find many important areas that are not covered. I apologize in advance for the inevitable consequence of writing a book of finite length.

I want to express my appreciation to my students at Los Alamos and the University of New Mexico for the effort they put into the course and for their help in resolving ambiguities in the material. In particular, I would like to thank Alan Wadlinger, Grenville Boicourt, Steven Wipf, and Jean Berlijn of Los Alamos National Laboratory for lively discussions on problem sets and for many valuable suggestions.

I am grateful to Francis Cole of Fermilab, Wemer Joho of the Swiss Nuclear Institute, William Herrmannsfeldt of the Stanford Linear Accelerator Center, Andris Faltens of Lawrence Berkeley Laboratory, Richard Cooper of Los Alamos National Laboratory, Daniel Prono of Lawrence Livermore Laboratory, Helmut Milde of Ion Physics Corporation, and George Fraser of Physics International Company for contributing material and commenting on the manuscript. I was aided in the preparation of the manuscript by lecture notes developed by James Potter of LANL and by Francis Cole. I would like to take this opportunity to thank David W. Woodall, L. K. Len, David Straw, Robert Jameson, Francis Cole, James Benford, Carl Ekdahl, Brendan Godfrey, William Rienstra, and McAllister Hull for their encouragement of and contributions towards the creation of an accelerator research program at the University of New Mexico. I am grateful for support that I received to attend the 1983 NATO Workshop on Fast Diagnostics.

STANLEY HUMPHRIES, JR.

University of New Mexico
December, 1985

Contents

1

Introduction

This book is an introduction to the theory of charged particle acceleration. It has two primary roles:

1. A unified, programmed summary of the principles underlying all charged particle accelerators.

2. A reference collection of equations and material essential to accelerator development and beam applications.

The book contains straightforward expositions of basic principles rather than detailed theories of specialized areas.

Accelerator research is a vast and varied field. There is an amazingly broad range of beam parameters for different applications, and there is a correspondingly diverse set of technologies to achieve the parameters. Beam currents range from nanoamperes (10^{-9} A) to megaamperes (10^{6} A). Accelerator pulselengths range from less than a nanosecond to steady state. The species of charged particles range from electrons to heavy ions, a mass difference factor approaching 10^{6}. The energy of useful charged particle beams ranges from a few electron volts (eV) to almost 1 TeV (10^{12} eV).

Organizing material from such a broad field is inevitably an imperfect process. Before beginning our study of beam physics, it is useful to review the order of topics and to define clearly the objectives and limitations of the book. The goal is to present the theory of accelerators on a level that facilitates the design of accelerator components and the operation of accelerators for applications. In order to accomplish this effectively, a considerable amount of potentially interesting material must be omitted:

1. Accelerator theory is interpreted as a mature field. There is no attempt to review the history of accelerators.

2. Although an effort has been made to include the most recent developments in accelerator science, there is insufficient space to include a detailed review of past and present literature.

3. Although the theoretical treatments are aimed toward an understanding of real devices, it is not possible to describe in detail specific accelerators and associated technology over the full range of the field.

These deficiencies are compensated by the books and papers tabulated in the bibliography.

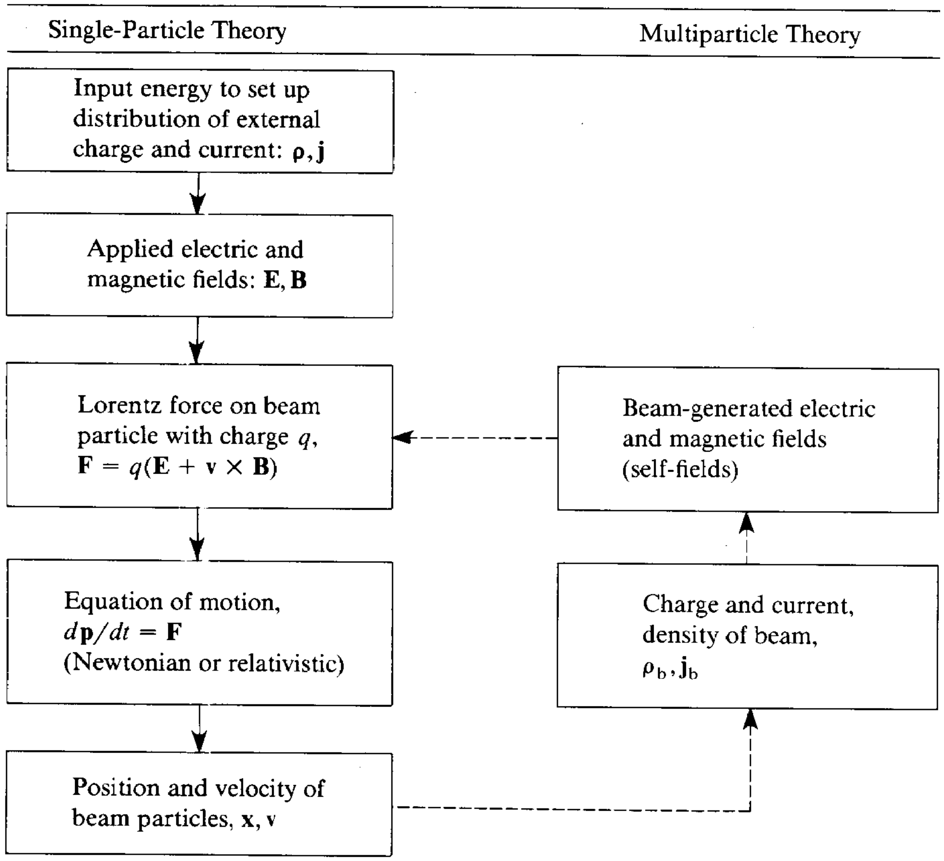
We begin with some basic definitions. A *charged particle* is an elementary particle or a macroparticle which contains an excess of positive or negative charge. Its motion is determined mainly by interaction with electromagnetic forces. *Charged particle acceleration* is the transfer of kinetic energy to a particle by the application of an electric field. A *charged particle beam* is a collection of particles distinguished by three characteristics: (1) beam particles have high kinetic energy compared to thermal energies, (2) the particles have a small spread in kinetic energy, and (3) beam particles move approximately in one direction. In most circumstances, a beam has a limited extent in the direction transverse to the average motion. The antithesis of a beam is an assortment of particles in thermodynamic equilibrium.

Most applications of charged particle accelerators depend on the fact that beam particles have high energy and good directionality. Directionality is usually referred to as *coherence*. Beam coherence determines, among other things, (1) the applied force needed to maintain a certain beam radius, (2) the maximum beam propagation distance, (3) the minimum focal spot size, and (4) the properties of an electromagnetic wave required to trap particles and accelerate them to high energy.

The process for generating charged particle beams is outlined in Table 1.1.. Electromagnetic forces result from mutual interactions between charged particles. In accelerator theory, particles are separated into two groups: (1) particles in the beam and (2) charged particles that are distributed on or in surrounding materials. The latter group is called the external charge. Energy is required to set up distributions of external charge; this energy is transferred to the beam particles via electromagnetic forces. For example, a power supply can generate a voltage difference between metal plates by subtracting negative charge from one plate and moving it to the other. A beam particle that moves between the plates is accelerated by attraction to the charge on one plate and repulsion from the charge on the other.

Electromagnetic forces are resolved into electric and magnetic components. Magnetic forces are present only when charges are in relative motion. The ability of a group of external charged

TABLE 1.1 Acceleration Process



particles to exert forces on beam particles is summarized in the *applied electric and magnetic fields*. Applied forces are usually resolved into those aligned along the average direction of the beam and those that act transversely. The axial forces are acceleration forces; they increase or decrease the beam energy. The transverse forces are confinement forces. They keep the beam contained to a specific cross-sectional area or bend the beam in a desired direction. Magnetic forces are always perpendicular to the velocity of a particle; therefore, magnetic fields cannot affect the particle's kinetic energy. Magnetic forces are confinement forces. Electric forces can serve both functions.

The distribution and motion of external charge determines the fields, and the fields determine the force on a particle via the Lorentz force law, discussed in Chapter 3. The expression for force is included in an appropriate equation of motion to find the position and velocity of particles in the beam as a function of time. A knowledge of representative particle orbits makes it possible to estimate average parameters of the beam, such as radius, direction, energy, and current. It is also possible to sum over the large number of particles in the beam to find charge density $?_b$ and current density j_b. These quantities act as source terms for beam-generated electric and magnetic fields.

This procedure is sufficient to describe low-current beams where the contribution to total electric and magnetic fields from the beam is small compared to those of the external charges. This is not the case at high currents. As shown in Table 1.1, calculation of beam parameters is no longer a simple linear procedure. The calculation must be self-consistent. Particle trajectories are determined by the total fields, which include contributions from other beam particles. In turn, the total fields are unknown until the trajectories are calculated. The problem must be solved either by successive iteration or application of the methods of collective physics.

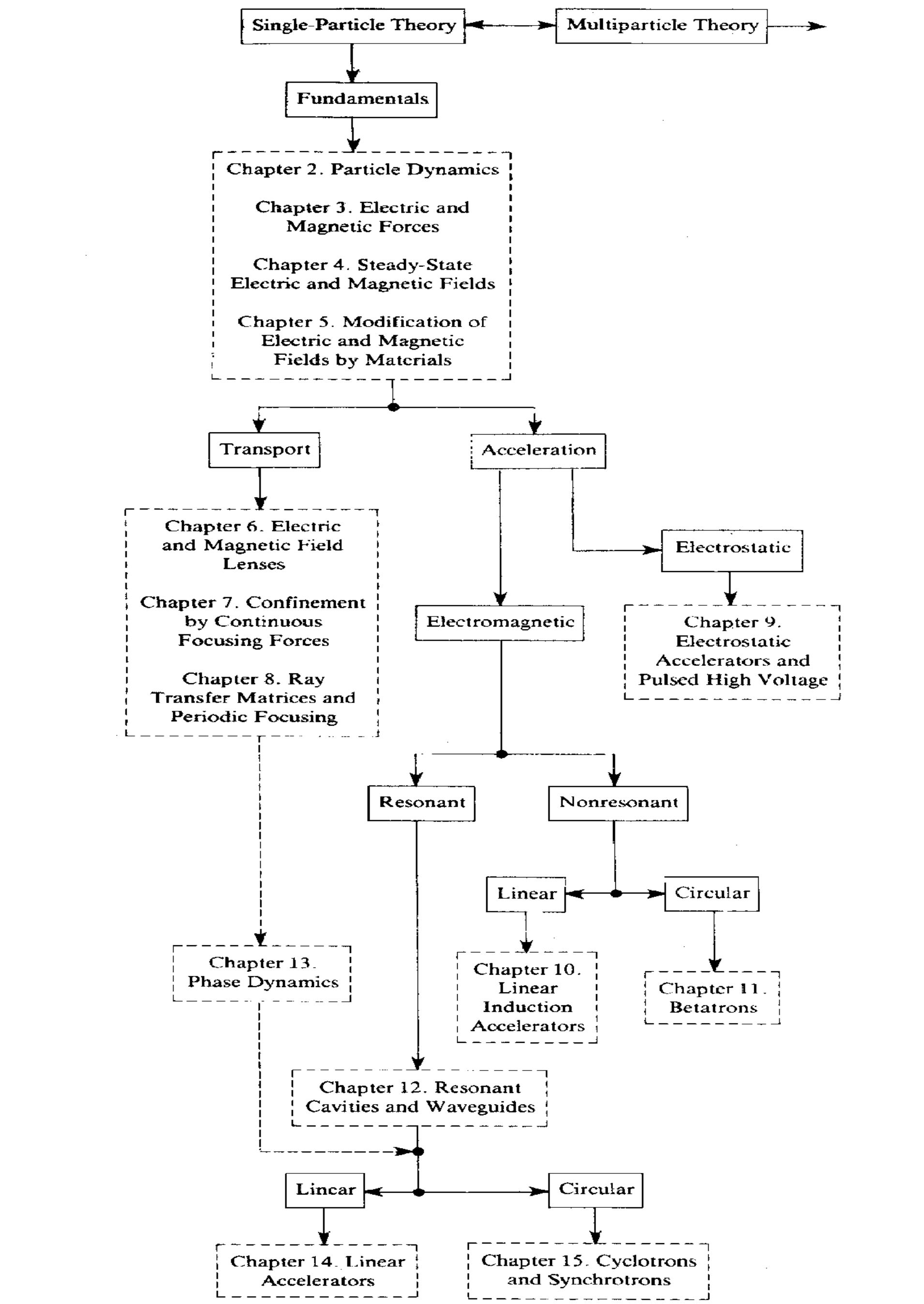
Single-particle processes are covered in this book. Although theoretical treatments for some devices can be quite involved, the general form of all derivations follows the straight-line sequence of Table 1.1. Beam particles are treated as test particles responding to specified fields. A continuation of this book addressing collective phenomena in charged particle beams is available: S. Humphries, **Charged Particle Beams** (Wiley, New York, 1990). A wide variety of useful processes for both conventional and high-power pulsed accelerators are described by collective physics, including (1) beam cooling, (2) propagation of beams injected into vacuum, gas, or plasma, (3) neutralization of beams, (4) generation of microwaves, (5) limiting factors for efficiency and flux, (6) high-power electron and ion guns, and (7) collective beam instabilities.

An outline of the topics covered in this book is given in Table 1.2. Single-particle theory can be subdivided into two categories: transport and acceleration. Transport is concerned with beam confinement. The study centers on methods for generating components of electromagnetic force that localize beams in space. For steady-state beams extending a long axial distance, it is sufficient to consider only transverse forces. In contrast, particles in accelerators with time-varying fields must be localized in the axial direction. Force components must be added to the accelerating fields for longitudinal particle confinement (phase stability).

Acceleration of charged particles is conveniently divided into two categories: electrostatic and electromagnetic acceleration. The accelerating field in electrostatic accelerators is the gradient of an electrostatic potential. The peak energy of the beam is limited by the voltage that can be sustained without breakdown. Pulsed power accelerators are included in this category because pulselengths are almost always long enough to guarantee simple electrostatic acceleration.

In order to generate beams with kinetic energy above a few million electron volts, it is necessary to utilize time-varying electromagnetic fields. Although particles in an electromagnetic accelerator experience continual acceleration by an electric field, the field does not require

TABLE 1.2 Organization of Topics



prohibitively large voltages in the laboratory. The accelerator geometry is such that inductively generated electric fields cancel electrostatic fields except at the position of the beam.

Electromagnetic accelerators are divided into two subcategories: nonresonant and resonant accelerators. Nonresonant accelerators are pulsed; the motion of particles need not be closely synchronized with the pulse waveform. Nonresonant electromagnetic accelerators are essentially step-up transformers, with the beam acting as a high-voltage secondary. The class is subdivided into linear and circular accelerators. A linear accelerator is a straight-through machine. Generally, injection into the accelerator and transport is not difficult; linear accelerators are

TABLE 1.3 Modern Accelerator Applications

- Production of short-lived radioisotopies for medical diagnosis
- Intense ion beams to drive inertial fusion reactors
- Electron beam welding
- Pulsed neutron sources for uranium borehole logging
- Cathode ray tubes and fast digitizers
- Electronuclear breeding of fissile fuels
- Measurement of cross sections for atomic physics
- Processing of semi-conductor circuits
- Secondary ion mass spectroscopy
- Generation of synchrotron radiation for solid-state physics research
- Diagnostics of rock formations in oil and natural gas wells
- Elementary particle physics
- Electron microscopy
- Materials testing for controlled thermonuclear fusion reactors
- Drivers for gas lasers and free electron lasers
- Cross-linking of thermoplastics
- Image intensifiers and fast streak tubes
- Pulsed X-ray radiography
- Surface modification of materials by ion implantation
- Nuclear structure studies
- Plasma heating for fusion reactors
- Assay of nuclear materials for safeguard applications
- Sterilization of food products
- Studies of transuranic elements
- Intense pulsed neutron sources for radiography and materials studies
- Generation of X rays and pions for radiation therapy
- Electron and ion surface microprobes
- Analysis of trace elements for biology and archeology
- Calibration of radiation detectors
- Studies of radiation damage in fission reactors

useful for initial acceleration of low-energy beams or the generation of high-flux beams. In circular machines, the beam is recirculated many times through the acceleration region during the pulse. Circular accelerators are well suited to the production of beams with high kinetic energy.

The applied voltage in a resonant accelerator varies harmonically at a specific frequency. The word *resonant* characterizes two aspects of the accelerator: (1) electromagnetic oscillations in resonant cavities or waveguides are used to transform input microwave power from low to high voltage and (2) there is close coupling between properties of the particle orbits and time variations of the accelerating field. Regarding the second property, particles must always be at the proper place at the proper time to experience a field with accelerating polarity. Longitudinal confinement is a critical issue in resonant accelerators. Resonant accelerators can also be subdivided into linear and circular machines, each category with its relative virtues.

In the early period of accelerator development, the quest for high kinetic energy, spurred by nuclear and elementary particle research, was the overriding goal. Today, there is increased emphasis on a diversity of accelerator applications. Much effort in modern accelerator theory is devoted to questions of current limits, beam quality, and the evolution of more efficient and cost-effective machines. The best introduction to modern accelerators is to review some of the active areas of research, both at high and low kinetic energy. The list in Table 1.3 suggests the diversity of applications and potential for future development.

2

Particle Dynamics

Understanding and utilizing the response of charged particles to electromagnetic forces is the basis of particle optics and accelerator theory. The goal is to find the time-dependent position and velocity of particles, given specified electric and magnetic fields. Quantities calculated from position and velocity, such as total energy, kinetic energy, and momentum, are also of interest. The nature of electromagnetic forces is discussed in Chapter 3. In this chapter, the response of particles to general forces will be reviewed. These are summarized in laws of motion. The Newtonian laws, treated in the first sections, apply at low particle energy. At high energy, particle trajectories must be described by relativistic equations. Although Newton's laws and their implications can be understood intuitively, the laws of relativity cannot since they apply to regimes beyond ordinary experience. Nonetheless, they must be accepted to predict particle behavior in high-energy accelerators. In fact, accelerators have provided some of the most direct verifications of relativity.

This chapter reviews particle mechanics. Section 2.1 summarizes the properties of electrons and ions. Sections 2.2-2.4 are devoted to the equations of Newtonian mechanics. These are applicable to electrons from electrostatic accelerators of in the energy range below 20 kV. This range includes many useful devices such as cathode ray tubes, electron beam welders, and microwave tubes. Newtonian mechanics also describes ions in medium energy accelerators used for nuclear physics. The Newtonian equations are usually simpler to solve than relativistic formulations. Sometimes it is possible to describe transverse motions of relativistic particles using Newtonian equations with a relativistically corrected mass. This approximation is treated in Section 2.10. In the second part of the chapter, some of the principles of special relativity are derived from two basic postulates, leading to a number of useful formulas summarized in Section 2.9.

2.1 CHARGED PARTICLE PROPERTIES

In the theory of charged particle acceleration and transport, it is sufficient to treat particles as dimensionless points with no internal structure. Only the influence of the electromagnetic force, one of the four fundamental forces of nature, need be considered. Quantum theory is unnecessary except to describe the emission of radiation at high energy.

This book will deal only with ions and electrons. They are simple, stable particles. Their response to the fields applied in accelerators is characterized completely by two quantities: mass and charge. Nonetheless, it is possible to apply much of the material presented to other particles. For example, the motion of macroparticles with an electrostatic charge can be treated by the methods developed in Chapters 6-9. Applications include the suspension of small objects in oscillating electric quadrupole fields and the acceleration and guidance of inertial fusion targets. At the other extreme are unstable elementary particles produced by the interaction of high-energy ions or electrons with targets. Beamlines, acceleration gaps, and lenses are similar to those used for stable particles with adjustments for different mass. The limited lifetime may influence hardware design by setting a maximum length for a beamline or confinement time in a storage ring.

An electron is an elementary particle with relatively low mass and negative charge. An ion is an assemblage of protons, neutrons, and electrons. It is an atom with one or more electrons removed. Atoms of the isotopes of hydrogen have only one electron. Therefore, the associated ions (the proton, deuteron, and triton) have no electrons. These ions are bare nucleii consisting of a proton with 0, 1, or 2 neutrons. Generally, the symbol Z denotes the atomic number of an ion or the number of electrons in the neutral atom. The symbol Z* is often used to represent the number of electrons removed from an atom to create an ion. Values of Z* greater than 30 may occur when heavy ions traverse extremely hot material. If Z* = Z, the atom is fully stripped. The atomic mass number A is the number of nucleons (protons or neutrons) in the nucleus. The mass of the atom is concentrated in the nucleus and is given approximately as Am_p, where mp is the proton mass.

Properties of some common charged particles are summarized in Table 2.1. The meaning of the rest energy in Table 2.1 will become clear after reviewing the theory of relativity. It is listed in energy units of million electron volts (MeV). An electron volt is defined as the energy gained by a particle having one fundamental unit of charge ($q = \pm e = \pm 1.6 \times 10^{-19}$ coulombs) passing

TABLE 2.1 Charged Particle Properties

Particle	Charge (coulomb)	Mass (kg)	Rest Energy (MeV)	A	Z	Z^*
Electron (β particle)	-1.60×10^{-19}	9.11×10^{-31}	0.511	—	—	—
Proton	$+1.60 \times 10^{-19}$	1.67×10^{-27}	938	1	1	1
Deuteron	$+1.60 \times 10^{-19}$	3.34×10^{-27}	1875	2	1	1
Triton	$+1.60 \times 10^{-19}$	5.00×10^{-27}	2809	3	1	1
He^{+}	$+1.60 \times 10^{-19}$	6.64×10^{-27}	3728	4	2	1
He^{++} (α particle)	$+3.20 \times 10^{-19}$	6.64×10^{-27}	3728	4	2	2
C^{+}	$+1.6 \times 10^{-19}$	1.99×10^{-26}	1.12×10^{4}	12	6	1
U^{+}	$+1.6 \times 10^{-19}$	3.95×10^{-25}	2.22×10^{5}	238	92	1

through a potential difference of one volt. In MKS units, the electron volt is

$$\text{I eV} = (1.6 \times 10^{-19}\ \text{C})\ (1\ \text{V}) = 1.6 \times 10^{-19}\ \text{J}.$$

Other commonly used metric units are keV (10^3 eV) and GeV (10^9 eV). Relativistic mechanics must be used when the particle kinetic energy is comparable to or larger than the rest energy. There is a factor of 1843 difference between the mass of the electron and the proton. Although methods for transporting and accelerating various ions are similar, techniques for electrons are quite different. Electrons are relativistic even at low energies. As a consequence, synchronization of electron motion in linear accelerators is not difficult. Electrons are strongly influenced by magnetic fields; thus they can be accelerated effectively in a circular induction accelerator (the betatron). High-current electron beams (~10 kA) can be focused effectively by magnetic fields. In contrast, magnetic fields are ineffective for high-current ion beams. On the other hand, it is possible to neutralize the charge and current of a high-current ion beam easily with light electrons, while the inverse is usually impossible.

2.2 NEWTON'S LAWS OF MOTION

The charge of a particle determines the strength of its interaction with the electromagnetic force. The mass indicates the resistance to a change in velocity. In Newtonian mechanics, mass is constant, independent of particle motion.

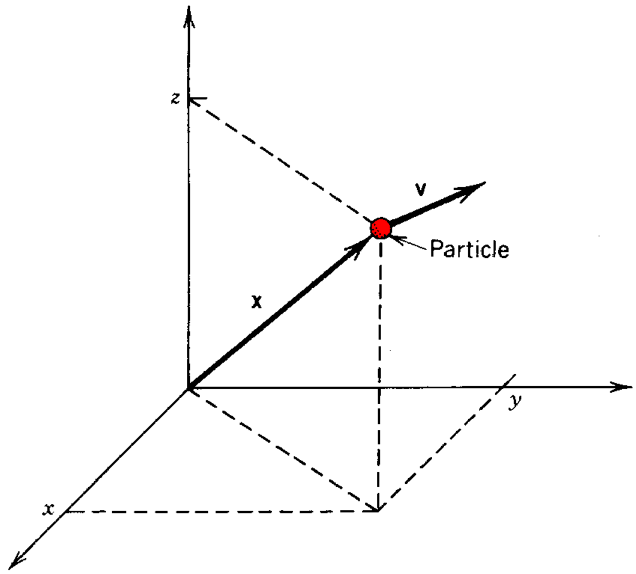


Figure 2.1. Position and velocity vectors of a particle in Cartesian coordinates.

The Newtonian mass (or *rest mass*) is denoted by a subscript: m_e for electrons, m_p for protons, and m_o for a general particle. A particle's behavior is described completely by its position in three-dimensional space and its velocity as a function of time. Three quantities are necessary to specify position; the position **x** is a vector. In the Cartesian coordinates (Figure 2.1), **x** can be written

$$\mathbf{x} = (x,y,z). \tag{2.1}$$

The particle velocity is

$$\mathbf{v} = (v_x,v_y,v_z) = (dx/dt,dy/dt,dz/dt) = d\mathbf{x}/dt, \tag{2.2}$$

Newton's first law states that a moving particle follows a straight-line path unless acted upon by a force. The tendency to resist changes in straight-line motion is called the momentum, **p**. Momentum is the product of a particle's mass and velocity,

$$\mathbf{p} = m_o\mathbf{v} = (p_x,p_y,p_z). \tag{2.3}$$

Newton's second law defines force F through the equation

$$d\mathbf{p}/dt = \mathbf{F}. \tag{2.4}$$

In Cartesian coordinates, Eq. (2.4) can be written

$$dp_x/dt = F_x, \quad dp_y/dt = F_y, \quad dp_z/dt = F_z. \tag{2.5}$$

Motions in the three directions are decoupled in Eq. (2.5). With specified force components, velocity components in the x, y, and z directions are determined by separate equations. It is important to note that this decoupling occurs only when the equations of motion are written in terms of Cartesian coordinates. The significance of straight-line motion is apparent in Newton's first law, and the laws of motion have the simplest form in coordinate systems based on straight lines. Caution must be exercised using coordinate systems based on curved lines. The analog of Eq. (2.5) for cylindrical coordinates (r, 0, z) will be derived in Chapter 3. In curvilinear coordinates, momentum components may change even with no force components along the coordinate axes.

2.3 KINETIC ENERGY

Kinetic energy is the energy associated with a particle's motion. The purpose of particle accelerators is to impart high kinetic energy. The kinetic energy of a particle, T, is changed by applying a force. Force applied to a static particle cannot modify T; the particle must be moved. The change in T (work) is related to the force by

$$\Delta T = \int \mathbf{F} \cdot \mathbf{dx}. \tag{2.6}$$

The integrated quantity is the vector dot product; **dx** is an incremental change in particle position.

In accelerators, applied force is predominantly in one direction. This corresponds to the symmetry axis of a linear accelerator or the main circular orbit in a betatron. With acceleration along the z axis, Eq. (2.6) can be rewritten

$$\Delta T = \int F_z dz = \int F_z \; (dz/dt) \; dt. \tag{2.7}$$

The chain rule of derivatives has been used in the last expression. The formula for T in Newtonian mechanics can be derived by (1) rewriting F, using Eq. (2.4), (2) taking T = 0 when v, = 0, and (3) assuming that the particle mass is not a function of velocity:

$$T = \int m_o v_z \; (dv_z/dt) \; dt = m_o v_z^2/2. \tag{2.8}$$

The differential relationship $d(m_o v_z^2/2)/dt = m_o v_z\, dv_z/dt$ leads to the last expression. The differences of relativistic mechanics arise from the fact that assumption 3 is not true at high energy.

When static forces act on a particle, the potential energy U can be defined. In this circumstance, the sum of kinetic and potential energies, T + U, is a constant called the total energy. If the force is axial, kinetic and potential energy are interchanged as the particle moves along the z axis, so that U = U(z). Setting the total time derivative of T + U equal to 0 and assuming $\partial U/\partial t = 0$ gives

$$m_o v_z (dv_z/dt) = -(\partial U/\partial z)(dz/dt). \tag{2.9}$$

The expression on the left-hand side equals $F_z v_z$. The static force and potential energy are related by

$$\mathbf{F_z} = -\partial U/\partial z, \quad \mathbf{F} = -\nabla U. \tag{2.10}$$

where the last expression is the general three-dimensional form written in terms of the vector gradient operator,

$$\nabla = \mathbf{u_x}\partial/\partial x + \mathbf{u_y}\partial/\partial y + \mathbf{u_z}\partial/\partial z. \tag{2.11}$$

The quantities $\mathbf{u}_x$, $\mathbf{u}_y$, and $\mathbf{u}_z$ are unit vectors along the Cartesian axes.

Potential energy is useful for treating electrostatic accelerators. Stationary particles at the source can be considered to have high U (potential for gaining energy). This is converted to kinetic energy as particles move through the acceleration column. If the potential function, U(x, y, z), is known, focusing and accelerating forces acting on particles can be calculated.

2.4 GALILEAN TRANSFORMATIONS

In describing physical processes, it is often useful to change the viewpoint to a frame of reference that moves with respect to an original frame. Two common frames of reference in accelerator theory are the stationary frame and the rest frame. The stationary frame is identified with the laboratory or accelerating structure. An observer in the rest frame moves at the average velocity of the beam particles; hence, the beam appears to be at rest. A coordinate transformation converts quantities measured in one frame to those that would be measured in another moving with velocity u. The transformation of the properties of a particle can be written symbolically as

$$(\mathbf{x},\mathbf{v},m,\mathbf{p},T) \Rightarrow (\mathbf{x}',\mathbf{v}',m',\mathbf{p}',T')$$

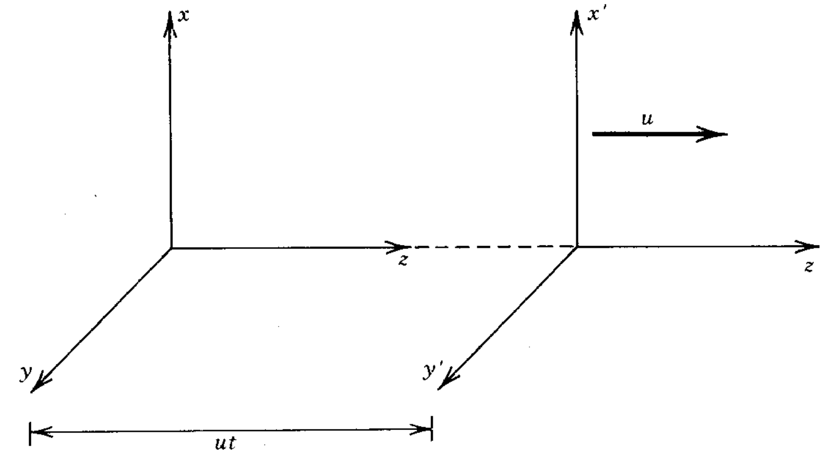


Figure 2.2. Galilean transformation between coordinate systems

where primed quantities are those measured in the moving frame. The operation that transforms quantities depends on **u**. If the transformation is from the stationary to the rest frame, **u** is the particle velocity **v**.

The transformations of Newtonian mechanics (Galilean transformations) are easily understood by inspecting Figure 2.2. Cartesian coordinate systems are defined so that the z axes are colinear with u and the coordinates are aligned at t = 0. This is consistent with the usual convention of taking the average beam velocity along the z axis. The position of a particle transforms as

$$x' = x, \quad y' = y, \quad z' = z - ut. \tag{2.12}$$

Newtonian mechanics assumes inherently that measurements of particle mass and time intervals in frames with constant relative motion are equal: m' = m and dt' = dt. This is not true in a relativistic description. Equations (2.12) combined with the assumption of invariant time intervals imply that $dx' = dx$ and $dx'/dt' = dx/dt$. The velocity transformations are

$$v_x' = v_x, \quad v_y' = v_y, \quad v_z' = v_x, - u. \tag{2.13}$$

Since m' = m, momenta obey similar equations. The last expression shows that velocities are additive. The axial velocity in a third frame moving at velocity w with respect to the x' frame is related to the original quantity by $v_z'' = v_z - u - w$.

Equations (2.13) can be used to determine the transformation for kinetic energy,

$$T' = T + ½m_o(-2uv_z + u^2). \tag{2.14}$$

Measured kinetic energy depends on the frame of reference. It can be either larger or smaller in a moving frame, depending on the orientation of the velocities. This dependence is an important factor in beam instabilities such as the two-stream instability.

2.5 POSTULATES OF RELATIVITV

The principles of special relativity proceed from two postulates:

1.The laws of mechanics and electromagnetism are identical in all inertial frames of reference.

2.Measurements of the velocity of light give the same value in all inertial frames.

Only the theory of special relativity need be used for the material of this book. General relativity incorporates the gravitational force, which is negligible in accelerator applications. The first postulate is straightforward; it states that observers in any *inertial frame* would derive the same laws of physics. An inertial frame is one that moves with constant velocity. A corollary is that it is impossible to determine an absolute velocity. Relative velocities can be measured, but there is no preferred frame of reference. The second postulate follows from the first. If the velocity of light were referenced to a universal stationary frame, tests could be devised to measure absolute velocity. Furthermore, since photons are the entities that carry the electromagnetic force, the laws of electromagnetism would depend on the absolute velocity of the frame in which they were derived. This means that the forms of the Maxwell equations and the results of electrodynamic experiments would differ in frames in relative motion. Relativistic mechanics, through postulate 2, leaves Maxwell's equations invariant under a coordinate transformation. Note that invariance does not mean that measurements of electric and magnetic fields will be the same in all frames. Rather, such measurements will always lead to the same governing equations.

The validity of the relativistic postulates is determined by their agreement with experimental measurements. A major implication is that no object can be induced to gain a measured velocity faster than that of light,

$$c \;=\; 2.998\times 10^{8}\;\; m/s. \tag{2.15}$$

This result is verified by observations in electron accelerators. After electrons gain a kinetic energy above a few million electron volts, subsequent acceleration causes no increase in electron velocity, even into the multi-GeV range. The constant velocity of relativistic particles is important in synchronous accelerators, where an accelerating electromagnetic wave must be

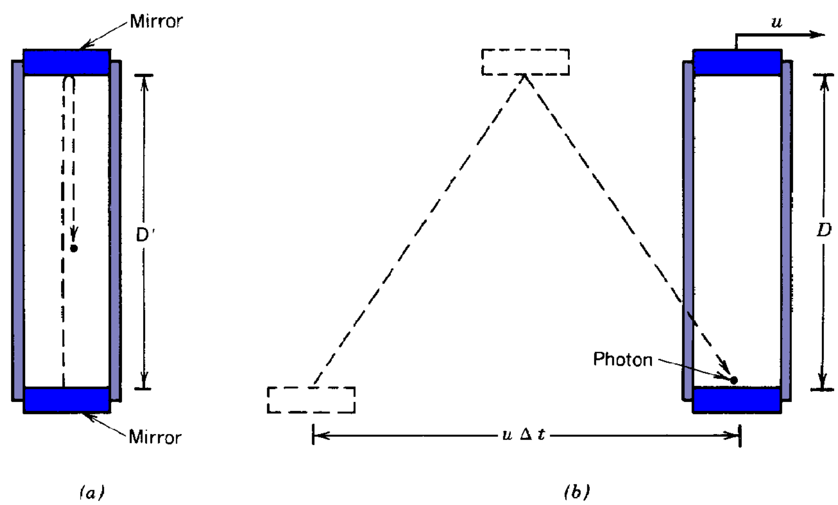


Figure 2.3 Effect of time dilation on the observed rates of a photon clock. (a) Clock rest frame. (b) Stationary frame.

matched to the motion of the particle.

2.6 TIME DILATION

In Newtonian mechanics, observers in relative motion measure the same time interval for an event (such as the decay of an unstable particle or the period of an atomic oscillation). This is not consistent with the relativistic postulates. The variation of observed time intervals (depending on the relative velocity) is called *time dilation*. The term *dilation* implies extending or spreading out.

The relationship between time intervals can be demonstrated by the clock shown in Figure 2.3, where double transits (back and forth) of a photon between mirrors with known spacing are measured. This test could actually be performed using a photon pulse in a mode-locked laser. In the rest frame (denoted by primed quantities), mirrors are separated by a distance D', and the photon has no motion along the z axis. The time interval in the clock rest frame is

$$\Delta t' = 2D'/c. \tag{2.16}$$

If the same event is viewed by an observer moving past the clock at a velocity - u, the photon appears to follow the triangular path shown in Figure 2.3b. According to postulate 2, the photon still travels with velocity c but follows a longer path in a double transit. The distance traveled in the laboratory frame is

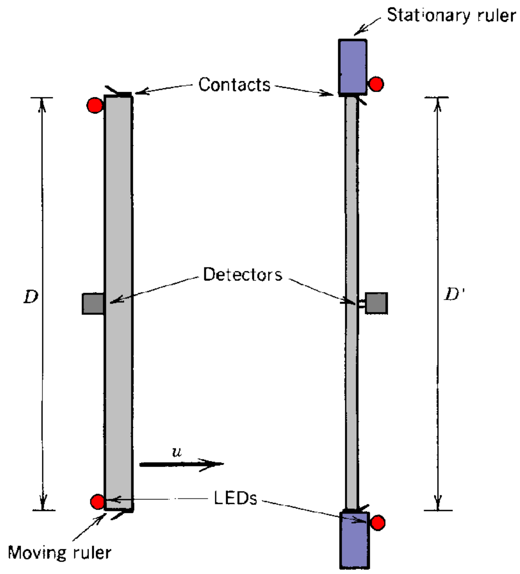


Figure 2.4. Experiment to demonstrate invariance of transverse lengths between frames in relative motion

$$c\Delta t = 2\left[D^2 + (u\Delta t/2)^2\right]^{½},$$

or

$$\Delta t = \frac{2D/c}{(1 - u^2/c^2)^{½}} . \tag{2.17}$$

In order to compare time intervals, the relationship between mirror spacing in the stationary and rest frames (D and D') must be known. A test to demonstrate that these are equal is illustrated in Figure 2.4. Two scales have identical length when at rest. Electrical contacts at the ends allow comparisons of length when the scales have relative motion. Observers are stationed at thecenters of the scales. Since the transit times of electrical signals from the ends to the middle are equal in all frames, the observers agree that the ends are aligned simultaneously. Measured length may depend on the magnitude of the relative velocity, but it cannot depend on the direction since there is no preferred frame or orientation in space. Let one of the scales move; the observer in the scale rest frame sees no change of length. Assume, for the sake of argument, that the stationary observer measures that the moving scale has shortened in the transverse direction, D < D'. The situation is symmetric, so that the roles of stationary and rest frames can be interchanged. This leads to conflicting conclusions. Both observers feel that their clock is the same length but the other is shorter. The only way to resolve the conflict is to take D = D'. The key to the argument is that the observers agree on simultaneity of the comparison events (alignment of the ends). This is not true in tests to compare axial length, as discussed in the next section. Taking D = D', the relationship between time intervals is

$$\Delta t \;=\; \frac{\Delta t'}{(1-u^2/c^2)^{1/2}} \;. \tag{2.18}$$

Two dimensionless parameters are associated with objects moving with a velocity u in a stationary frame:

$$\beta \;=\; u/c, \quad \gamma = (1-u^2/c^2)^{-1/2}. \tag{2.19}$$

These parameters are related by

$$\gamma \;=\; (1-\beta^2)^{-1/2}, \tag{2.20}$$

$$\beta \;=\; (1-1/\gamma^2)^{1/2}. \tag{2.21}$$

A time interval Δt measured in a frame moving at velocity u with respect to an object is related to an interval measured 'in the rest frame of the object, Δt', by

$$\Delta t \;=\; \gamma \Delta t'. \tag{2.22}$$

For example, consider an energetic π^+ pion (rest energy 140 MeV) produced by the interaction of a high-energy proton beam from an accelerator with a target. If the pion moves at velocity 2.968×10^8 m/s in the stationary frame, it has a β value of 0.990 and a corresponding γ value of 8.9. The pion is unstable, with a decay time of 2.5×10^{-8} s at rest. Time dilation affects the decay time measured when the particle is in motion. Newtonian mechanics predicts that the average distance traveled from the target is only 7.5 in, while relativistic mechanics (in agreement with observation) predicts a decay length of 61 in for the high-energy particles.

2.7 LORENTZ CONTRACTION

Another familiar result from relativistic mechanics is that a measurement of the length of a moving object along the direction of its motion depends on its velocity. This phenomenon is

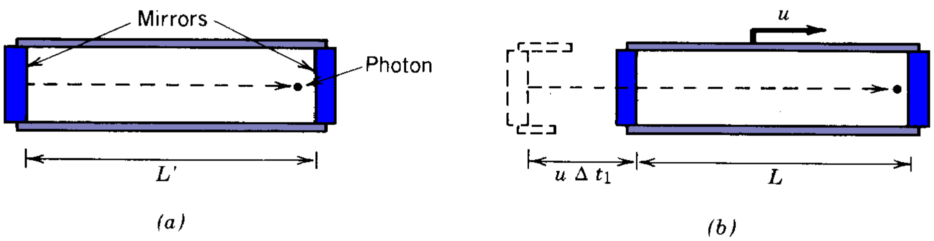


Figure 2.5 Lorentz contraction of a photon clock. (a) Clock rest frame.
(bl) Stationary frame

known as Lorentz contraction. The effect can be demonstrated by considering the clock of Section 2.6 oriented as shown in Figure 2.5.

The detector on the clock measures the double transit time of light between the mirrors. Pulses are generated when a photon leaves and returns to the left-hand mirror. Measurement of the single transit time would require communicating the arrival time of the photon at the right-hand mirror to the timer at the left-hand mirror. Since the maximum speed with which this information can be conveyed is the speed of light, this is equivalent to a measurement of the double transit time. In the clock rest frame, the time interval is $\Delta t' = 2L'/c$.

To a stationary observer, the clock moves at velocity u. During the transit in which the photon leaves the timer, the right-hand mirror moves away. The photon travels a longer distance in the stationary frame before being reflected. Let Δt_1, be the time for the photon to travel from the left to right mirrors. During this time, the right-hand mirror moves a distance u At,. Thus,

$$c\Delta t_1 = (L + u\Delta t_1),$$

where L is the distance between mirrors measured in the stationary frame. Similarly, on the reverse transit, the left-hand mirror moves toward the photon. The time to complete this leg is

$$\Delta t_2 = (L - u\Delta t_2)/c.$$

The total time for the event in the stationary frame is

$$\Delta t = \Delta t_1 + \Delta t_2 = \frac{L}{c-u} + \frac{L}{c+u} \ ,$$

or

$$\Delta t = \frac{2L/c}{1-u^2/c^2} \ .$$

Time intervals cannot depend on the orientation of the clock, so that Eq. (2.22) holds. The above equations imply that

$$L = L'/\gamma. \tag{2.23}$$

Thus, a moving object appears to have a shorter length than the same object at rest.

The acceleration of electrons to multi-GeV energies in a linear accelerator provides an interesting example of a Lorentz contraction effect. Linear accelerators can maintain longitudinal accelerating gradients of, at most, a few megavolts per meter. Lengths on the kilometer scale are required to produce high-energy electrons. To a relativistic electron, the accelerator appears to be rushing by close to the speed of light. The accelerator therefore has a contracted apparent length of a few meters. The short length means that focusing lenses are often unnecessary in electron linear accelerators with low-current beams.

2.8 LORENTZ TRANSFORMATIONS

Charged particle orbits are characterized by position and velocity at a certain time, (**x**, **v**, t). In Newtonian mechanics, these quantities differ if measured in a frame moving with a relative velocity with respect to the frame of the first measurement. The relationship between quantities was summarized in the Galilean transformations.

The Lorentz transformations are the relativistic equivalents of the Galilean transformations. In the same manner as Section 2.4, the relative velocity of frames is taken in the z direction and the z and z' axes are colinear. Time is measured from the instant that the two coordinate systems are aligned (z = z' = 0 at t = t' = 0). The equations relating position and time measured in one frame (unprimed quantities) to those measured in another frame moving with velocity u (primed quantities) are

$$x' = x, \tag{2.24}$$

$$y' = y, \tag{2.25}$$

$$z' = \frac{z-ut}{(1-u^2/c^2)^{1/2}} = \gamma(z-ut), \tag{2.26}$$

$$t' = \frac{t-uz/c^2}{(1-u^2/c^2)^{1/2}} = \gamma\left(t-\frac{uz}{c^2}\right). \tag{2.27}$$

Particle Dynamics

The primed frame is not necessarily the rest frame of a particle. One major difference between the Galilean and Lorentz transformations is the presence of the γ factor. Furthermore, measurements of time intervals are different in frames in relative motion. Observers in both frames may agree to set their clocks at t = t' = 0 (when z = z' = 0), but they will disagree on the subsequent passage of time [Eq. (2.27)]. This also implies that events at different locations in z that appear to be simultaneous in one frame of reference may not be simultaneous in another.

Equations (2.24)-(2.27) may be used to derive transformation laws for particle velocities. The differentials of primed quantities are

$$dx' = dx, \quad dy' = dy, \quad dz' = \gamma(dz - u\,dt),$$

$$dt' = \gamma dt \;(1 - uv_z/c^2).$$

In the special case where a particle has only a longitudinal velocity equal to u, the particle is at rest in the primed frame. For this condition, time dilation and Lorentz contraction proceed directly from the above equations.

Velocity in the primed frame is dx'/dt'. Substituting from above,

$$v_x' = \frac{v_x}{\gamma\;(1 - uv_z/c^2)} \;. \tag{2.28}$$

When a particle has no longitudinal motion in the primed frame (i.e., the primed frame is the rest frame and v_z = u), the transformation of transverse velocity is

$$v_x' = \gamma\; v_x. \tag{2.29}$$

This result follows directly from time dilation. Transverse distances are the same in both frames, but time intervals are longer in the stationary frame.

The transformation of axial particle velocities can be found by substitution for dz' and dt',

$$\frac{dz'}{dt'} = \frac{\gamma dt\;(dz/dt - u)}{\gamma dt\;(1 - uv_z/c^2)} \;,$$

or

$$v_z' = \frac{v_z - u}{1 - uv_z/c^2} \;. \tag{2.30}$$

This can be inverted to give

$$v_z = \frac{v_z^{/}+u}{1+uv_z^{/}/c^2} \ . \tag{2.31}$$

Equation (2.31) is the relativistic velocity addition law. If a particle has a velocity v_z' in the primed frame, then Eq. (2.31) gives observed particle velocity in a frame moving at *-u*. For v_z' approaching c, inspection of Eq. (2.31) shows that v_z also approaches c. The implication is that there is no frame of reference in which a particle is observed to move faster than the velocity of light. A corollary is that no matter how much a particle's kinetic energy is increased, it will never be observed to move faster than c. This property has important implications in particle acceleration. For example, departures from the Newtoniain velocity addition law set a limit on the maximum energy available from cyclotrons. In high-power, multi-MeV electron extractors, saturation of electron velocity is an important factor in determining current propagation limits.

2.9 RELATIVISTIC FORMULAS

The motion of high-energy particles must be described by relativistic laws of motion. Force is related to momentum by the same equation used in Newtonian mechanics

$$d\mathbf{p}/dt = \mathbf{F}. \tag{2.32}$$

This equation is consistent with the Lorentz transformations if the momentum is defined as

$$\mathbf{p} = \gamma m_o \mathbf{v}. \tag{2.33}$$

The difference from the Newtonian expression is the γ factor. It is determined by the total particle velocity v observed in the stationary frame, $\gamma = (1-v^2/c^2)^{-½}$. One interpretation of Eq. (2.33) is that a particle's effective mass increases as it approaches the speed of light. The relativistic mass is related to the rest mass by

$$m = \gamma m_o. \tag{2.34}$$

The relativistic mass grows without limit as v_z approaches c. Thus, the momentum increases although there is a negligible increase in velocity.

In order to maintain Eq. (2.6), relating changes of energy to movement under the influence of a force, particle energy must be defined as

$$E = \gamma m_o c^2. \tag{2.35}$$

The energy is not zero for a stationary particle, but approaches $m_o c^2$, which is called the rest energy. The kinetic energy (the portion of energy associated with motion) is given by

$$T = E - m_o c^2 = m_o c^2 (\gamma - 1). \tag{2.36}$$

Two useful relationships proceed directly from Eqs. (2.20). (2.33), and (2.35):

$$E = \sqrt{c^2 p^2 + m_o^2 c^4}, \tag{2.37}$$

where $p^2 = \mathbf{p} \cdot \mathbf{p}$, and

$$\boldsymbol{v} = c^2 \boldsymbol{p}/E. \tag{2.38}$$

The significance of the rest energy and the region of validity of Newtonian mechanics is clarified by expanding Eq. (2.35) in limit that v/c « 1.

$$E = \frac{m_o c^2}{\sqrt{1 - v^2/c^2}} = m_o c^2 \ (1 + v^2/c^2 + ...). \tag{2.39}$$

The Newtonian expression for T [Eq. (2.8)] is recovered in the second term. The first term is a constant component of the total energy, which does not affect Newtonian dynamics. Relativistic expressions must be used when $T \geq m_o c2$. The rest energy plays an important role in relativistic mechanics.

Rest energy is usually given in units of electron volts. Electrons are relativistic when T is in the MeV range, while ions (with a much larger mass) have rest energies in the GeV range. Figure 2.6 plots β for particles of interest for accelerator applications as a function of kinetic energy. The Newtonian result is also shown. The graph shows saturation of velocity at high energy and the energy range where departures from Newtonian results are significant.

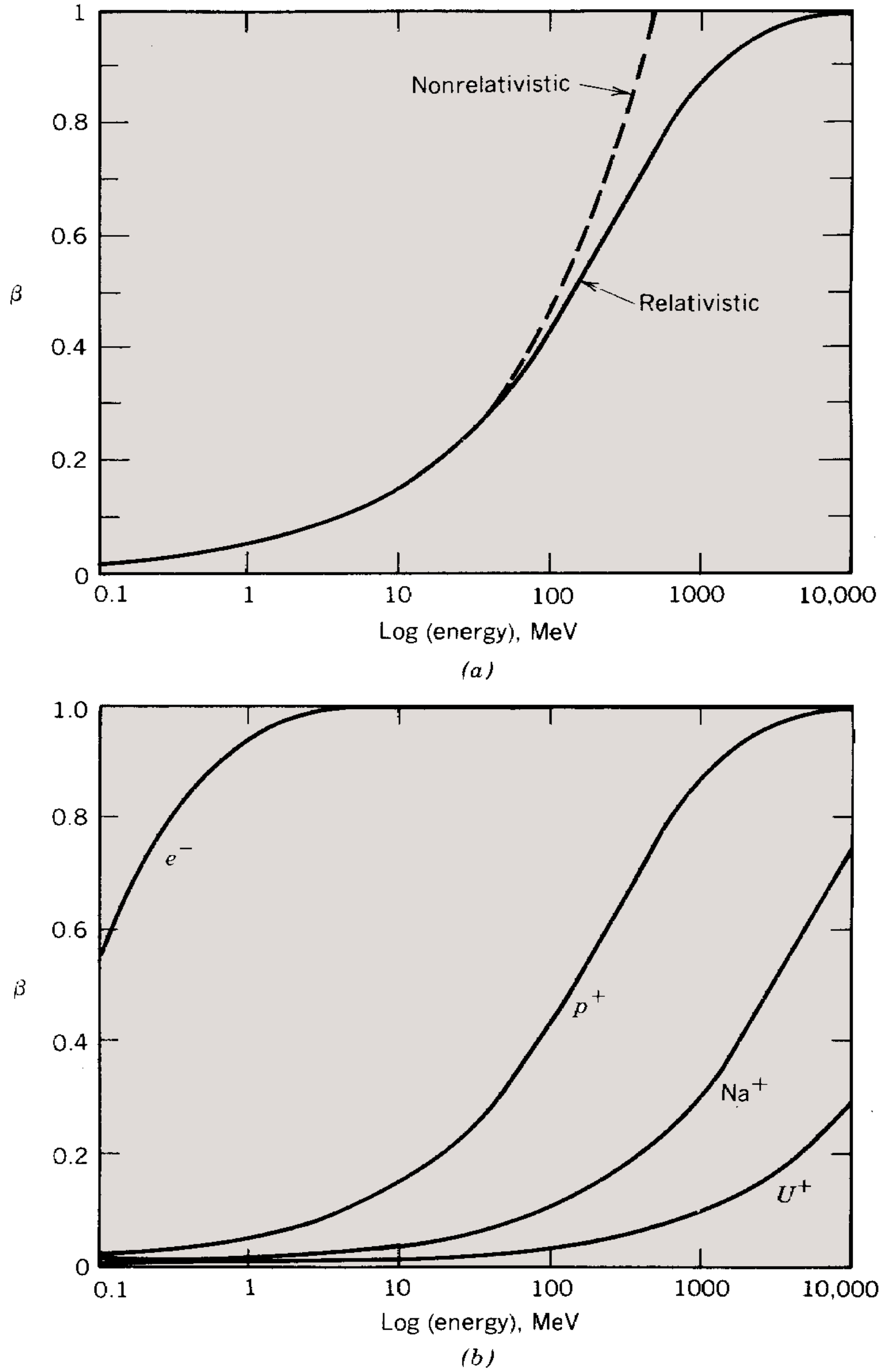


Figure 2.6. Particle velocity normalized to the speed of light as a function of kinetic energy. (a) Protons: solid line, relativistic predicted, dashed line, Newtonian predicition. (b) Relativistic predictions for various particles.

2.10 NONRELATIVISTIC APPROXIMATION FOR TRANSVERSE MOTION

A relativistically correct description of particle motion is usually more difficult to formulate and solve than one involving Newtonian equations. In the study of the transverse motions of charged particle beams, it is often possible to express the problem in the form of Newtonian equations with the rest mass replaced by the relativistic mass. This approximation is valid when the beam is well directed so that transverse velocity components are small compared to the axial velocity of beam particles. Consider the effect of focusing forces applied in the x direction to confine particles along the z axis. Particles make small angles with this axis, so that v_x is always small compared to v_z. With $\mathbf{F} = \mathbf{u}_x F_x$, Eq. (2.32) can be written in the form

$$\gamma m_o v_x \left(\frac{d\gamma/dt}{\gamma} + \frac{dv_x/dt}{v_x} \right) = F_x. \tag{2.40}$$

Equation (2.39) can be rewritten as

$$E = m_o c^2 \; [1 + (v_z^2 + v_x^2)/2c^2 + 3(v_z^2 + v_x^2)^2/8c^4 + \ldots].$$

When $v_x \ll v_z$, relative changes in γ resulting from the transverse motion are small. In Eq. (2.40), the first term in parenthesis is much less than the second, so that the equation of motion is approximately

$$\gamma m_o \frac{dv_x}{dt} = F_x. \tag{2.41}$$

This has the form of a Newtonian expression with m_o replaced by γm_o.

3

Electric and Magnetic Forces

Electromagnetic forces determine all essential features of charged particle acceleration and transport. This chapter reviews basic properties of electromagnetic forces. Advanced topics, such as particle motion with time-varying forces, are introduced throughout the book as they are needed.

It is convenient to divide forces between charged particles into electric and magnetic components. The relativistic theory of electrodynamics shows that these are manifestations of a single force. The division into electric and magnetic interactions depends on the frame of reference in which particles are observed.

Section 3.1 introduces electromagnetic forces by considering the mutual interactions between pairs of stationary charges and current elements. Coulomb's law and the law of Biot and Savart describe the forces. Stationary charges interact through the electric force. Charges in motion constitute currents. When currents are present, magnetic forces also act.

Although electrodynamics is described completely by the summation of forces between individual particles, it is advantageous to adopt the concept of fields. Fields (Section 3.2) are mathematical constructs. They summarize the forces that could act on a test charge in a region with a specified distribution of other charges. Fields characterize the electrodynamic properties of the charge distribution. The Maxwell equations (Section 3.3) are direct relations between electric and magnetic fields. The equations determine how fields arise from distributed charge and current and specify how field components are related to each other.

Electric and magnetic fields are often visualized as vector lines since they obey equations similar to those that describe the flow of a fluid. The field magnitude (or strength) determines the density of tines. In this interpretation, the Maxwell equations are fluidlike equations that describe the creation and flow of field lines. Although it is unnecessary to assume the physical existence of field lines, the concept is a powerful aid to intuit complex problems.

The Lorentz law (Section 3.2) describes electromagnetic forces on a particle as a function of fields and properties of the test particle (charge, position and velocity). The Lorentz force is the basis for all orbit calculations in this book. Two useful subsidiary functions of field quantities, the electrostatic and vector potentials, are discussed in Section 3.4. The electrostatic potential (a function of position) has a clear physical interpretation. If a particle moves in a static electric field, the change in kinetic energy is equal to its charge multiplied by the change in electrostatic potential. Motion between regions of different potential is the basis of electrostatic acceleration. The interpretation of the vector potential is not as straightforward. The vector potential will become more familiar through applications in subsequent chapters.

Section 3.6 describes an important electromagnetic force calculation, motion of a charged particle in a uniform magnetic field. Expressions for the relativistic equations of motion in cylindrical coordinates are derived in Section 3.5 to apply in this calculation.

3.1 FORCES BETWEEN CHARGES AND CURRENTS

The simplest example of electromagnetic forces, the mutual force between two stationary point charges, is illustrated in Figure 3.1a. The force is directed along the line joining the two particles, $\mathbf{r}$. In terms of $\mathbf{u}_r$ (a vector of unit length aligned along r), the force on particle 2 from particle 1 is

$$\boldsymbol{F}(1\Rightarrow 2) \;=\; \frac{1}{4\pi\varepsilon_o}\,\frac{q_1 q_2 \boldsymbol{u_r}}{r^2}\;\;(newtons). \tag{3.1}$$

The value of ε_o is

$$\varepsilon_o \;=\; 8.85\times 10^{-12}\;\;(A{-}s/V{-}m).$$

In Cartesian coordinates, $\mathbf{r} = (x_2-x_1)\mathbf{u}_x + (y_2-y_1)\mathbf{u}_y + (z_2-z_1)\mathbf{u}_z$. Thus, $r^2=(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2$. The force on particle 1 from particle 2 is equal and opposite to that of Eq. (3.1). Particles with the same polarity of charge repel one another. This fact affects high-current beams. The electrostatic repulsion of beam particles causes beam expansion in the absence of strong focusing.

Currents are charges in motion. Current is defined as the amount of charge in a certain cross section (such as a wire) passing a location in a unit of time. The mks unit of current is the ampere (coulombs per second). Particle beams may have charge and current. Sometimes, charge effects

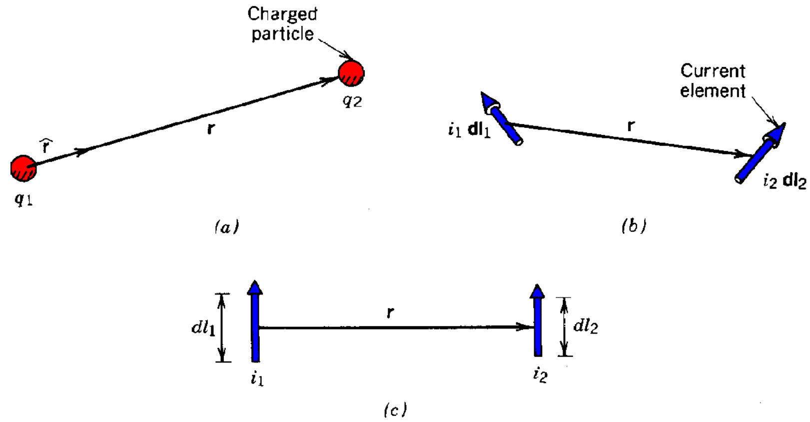


Figure 3.1 Forces between charged particles. (*a*) Electrostatic force. (*b*) Magnetostatic force. (*c*) Magnetostatic force between parallel current elements.

can be neutralized by adding particles of opposite-charge sign, leaving only the effects of current. This is true in a metal wire. Electrons move through a stationary distribution of positive metal ions. The force between currents is described by the law of Biot and Savart. If $i_1\mathbf{dl}_1$ and $i_2\mathbf{dl}_2$ are current elements (e.g., small sections of wires) oriented as in Figure 3.1b, the force on element 2 from element 1 is

$$d\boldsymbol{F} = \frac{\mu_o}{4\pi}\,\frac{i_2\boldsymbol{dl}_2\times(i_1\boldsymbol{dl}_1\times\boldsymbol{u}_r)}{r^2}\ . \qquad (3.2)$$

where $\mathbf{u}_r$ is a unit vector that points from 1 to 2 and

$$\mu_o = 4\pi\times10^{-7} = 1.26\times10^{-6}\ (V\text{-}s/A\text{-}m).$$

Equation (3.2) is more complex than (3.1); the direction of the force is determined by vector cross products. Resolution of the cross products for the special case of parallel current elements is shown in Figure 3.1c. Equation (3.2) becomes

$$d\boldsymbol{F}(1\Rightarrow2) = -\frac{\mu_o}{4\pi}\,\frac{i_1 i_2 dl_1 dl_2}{r^2}\ \boldsymbol{u}_r.$$

Currents in the same direction attract one another. This effect is important in high-current relativistic electron beams. Since all electrons travel in the same direction, they constitute parallel current elements, and the magnetic force is attractive. If the electric charge is neutralized by ions, the magnetic force dominates and relativistic electron beams can be self-confined.

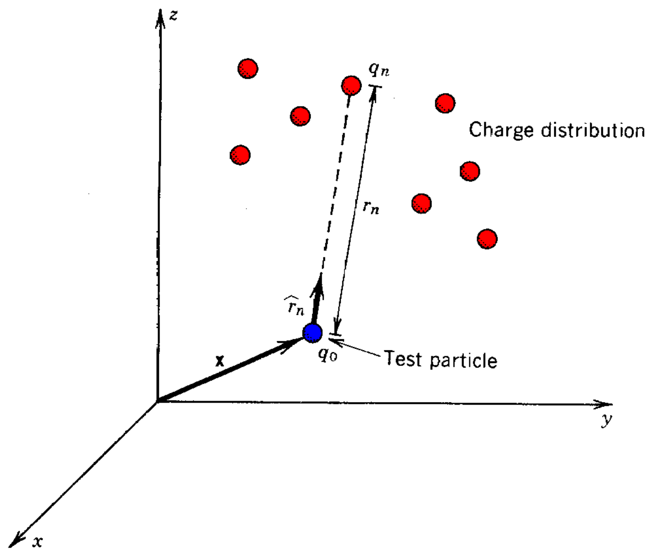


Figure 3.2 Electrical force exerted on a test particle by a distribution of charged particles.

3.2 THE FIELD DESCRIPTION AND THE LORENTZ FORCE

It is often necessary to calculate electromagnetic forces acting on a particle as it moves through space. Electric forces result from a specified distribution of charge. Consider, for instance, a low-current beam in an electrostatic accelerator. Charges on the surfaces of the metal electrodes provide acceleration and focusing. The electric force on beam particles at any position is given in terms of the specified charges by

$$\boldsymbol{F} = \sum_n \frac{1}{4\pi\varepsilon_o} \frac{q_o q_n \boldsymbol{u_{rn}}}{r_n^2} ,$$

where q_o is the charge of a beam particle and the sum is taken over all the charges on the electrodes (Fig. 3.2).

In principle, particle orbits can be determined by performing the above calculation at each point of each orbit. A more organized approach proceeds from recognizing that (1) the potential force on a test particle at any position is a function of the distribution of external charges and (2) the net force is proportional to the charge of the test particle. The function $\mathbf{F}(x)/q_o$ characterizes the action of the electrode charges. It can be used in subsequent calculations to determine the orbit of any test particle. The function is called the *electric field* and is defined by

$$\boldsymbol{E}(\boldsymbol{x}) = \sum_n \frac{1}{4\pi\varepsilon_o} \frac{q_n \boldsymbol{u_{rn}}}{r_n^2} . \tag{3.3}$$

The sum is taken over all specified charges. It may include freely moving charges in conductors, bound charges in dielectric materials, and free charges in space (such as other beam particles). If the specified charges move, the electric field may also be a function of time-, in this case, the equations that determine fields are more complex than Eq. (3.3).

The electric field is usually taken as a smoothly varying function of position because of the l/r^2 factor in the sum of Eq. (3.3). The smooth approximation is satisfied if there is a large number of specified charges, and if the test charge is far from the electrodes compared to the distance between specified charges. As an example, small electrostatic deflection plates with an applied voltage of 100 V may have more than 10" electrons on the surfaces. The average distance between electrons on the conductor surface is typically less than 1 µm.

When **E** is known, the force on a test particle with charge q_o as a function of position is

$$\boldsymbol{F}(\boldsymbol{x}) = q_o\ \boldsymbol{E}(\boldsymbol{x}). \tag{3.4}$$

This relationship can be inverted for measurements of electric fields. A common nonperturbing technique is to direct a charged particle beam through a region and infer electric field by the acceleration or deflection of the beam.

A summation over current elements similar to Eq. (3.3) can be performed using the law of Biot and Savart to determine forces that can act on a differential test element of current. This function is called the magnetic field **B**. (Note that in some texts, the term magnetic field is reserved for the quantity **H**, and **B** is called the magnetic induction.) In terms of the field, the magnetic force on i**dl** is

$$d\boldsymbol{F} = i\boldsymbol{dl} \times \boldsymbol{B}. \tag{3.5}$$

Equation (3.5) involves the vector cross product. The force is perpendicular to both the current element and magnetic field vector.

An expression for the total electric and magnetic forces on a single particle is required to treat beam dynamics. The differential current element, i**dl**, must be related to the motion of a single charge. The correspondence is illustrated in Figure 3.3. The test particle has charge q and velocity **v**. It moves a distance dl in a time dt = |dl|/|v|. The current (across an arbitrary cross section) represented by this motion is q/(|dl|/|v|). A moving charged particle acts like a current element with

$$i\boldsymbol{dl} = \frac{q\boldsymbol{dl}}{|dl|/|v|} = q\boldsymbol{v}.$$

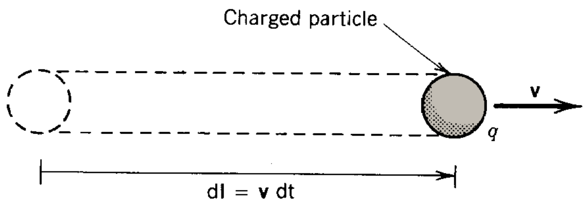


Figure 3.3 Representation of a moving charged particle as a current element.

The magnetic force on a charged particle is

$$\boldsymbol{F} = q\boldsymbol{v} \times \boldsymbol{B}. \tag{3.6}$$

Equations (3.4) and (3.6) can be combined into a single expression (the Lorentz force law)

$$\boldsymbol{F}(\boldsymbol{x},t) = q\ (\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}). \tag{3.7}$$

Although we derived Equation (3.7) for static fields, it holds for time-dependent fields as well. The Lorentz force law contains all the information on the electromagnetic force necessary to treat charged particle acceleration. With given fields, charged particle orbits are calculated by combining the Lorentz force expression with appropriate equations of motion. In summary, the field description has the following advantages.

1. Fields provide an organized method to treat particle orbits in the presence of large numbers of other charges. The effects of external charges are summarized in a single, continuous function.
2. Fields are themselves described by equations (Maxwell equations). The field concept extends beyond the individual particle description. Chapter 4 will show that field lines obey geometric relationships. This makes it easier to visualize complex force distributions and to predict charged particle orbits.
3. Identification of boundary conditions on field quantities sometimes makes it possible to circumvent difficult calculations of charge distributions in dielectrics and on conducting boundaries.
4. It is easier to treat time-dependent electromagnetic forces through direct solution for field quantities.

The following example demonstrates the correspondence between fields and charged particle distributions. The parallel plate capacitor geometry is shown in Figure 3.4. Two infinite parallel metal plates are separated by a distance d. A battery charges the plates by transferring electrons from one plate to the other. The excess positive charge and negative electron charge spread uniformerly on the inside surfaces. If this were not true, there would be electric fields inside the metal. The problem is equivalent to calculating the electric fields from two thin sheets of charge,

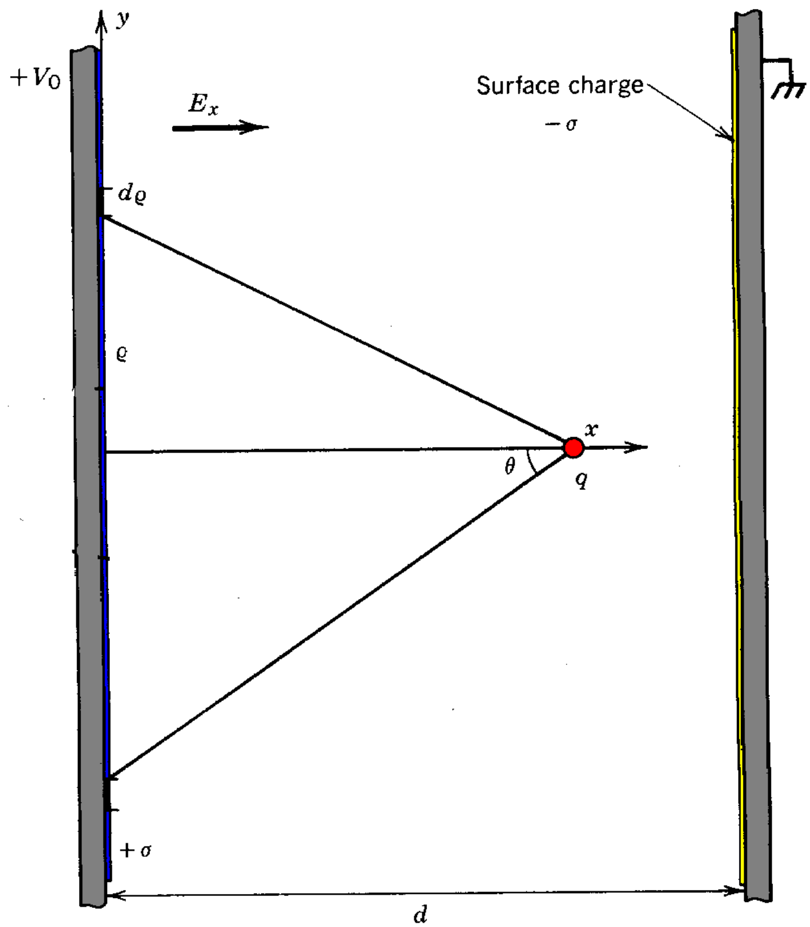


Figure 3.4 Calculation of force on a test particle between uniformly charged plates.

as shown in Figure 3.4. The surface charge densities, $\pm\,\sigma$ (in coulombs per square meter), are equal in magnitude and opposite in sign.

A test particle is located between the plates a distance *x* from the positive electrode. Figure 3.4 defines a convenient coordinate system. The force from charge in the differential annulus illustrated is repulsive. There is force only in the x direction; by symmetry transverse forces cancel. The annulus has charge ($2\pi\rho\ d\rho\ \sigma$) and is a distance $(\rho^2 + x^2)^{1/2}$ from the test charge. The total force [from Eq. (3.1)] is multiplied by $\cos\theta$ to give the x component.

$$dF_x = \frac{2\pi\rho\ d\rho\ \sigma q_o\ \cos\theta}{4\pi\varepsilon_o\ (\rho^2+x^2)}\ ,$$

where $\cos\theta = x/(\rho^2 + x^2)^{1/2}$. Integrating the above expression over ρ from 0 to ∞ gives the net force

$$F^{+} = \int_{0}^{\infty} \frac{\rho \; d\rho \;\; \sigma q_o \; x}{2\varepsilon_o \;\; (\rho^2 + x^2)^{3/2}} = \frac{q_o \sigma}{2\varepsilon_o} \; . \tag{3.8}$$

A similar result is obtained for the force from the negative-charge layer. It is attractive and adds to the positive force. The electric field is found by adding the forces and dividing by the charge of the test particle

$$E_x(x) = (F^{+} - F^{-})/q = \sigma/\varepsilon_o . \tag{3.9}$$

The electric field between parallel plates is perpendicular to the plates and has uniform magnitude at all positions. Approximations to the parallel plate geometry are used in electrostatic deflectors; particles receive the same impulse independent of their position between the plates.

3.3 THE MAXWELL EQUATIONS

The Maxwell equations describe how electric and magnetic fields arise from currents and charges. They are continuous differential equations and are most conveniently written if charges and currents are described by continuous functions rather than by discrete quantities. The source functions are the *charge density*, $\rho(x, y, z, t)$ and *current density* $\mathbf{j}(x, y, z, t)$.

The charge density has units of coulombs per cubic meters (in MKS units). Charges are carried by discrete particles, but a continuous density is a good approximation if there are large numbers of charged particles in a volume element that is small compared to the minimum scale length of interest. Discrete charges can be included in the Maxwell equation formulation by taking a charge density of the form $\rho = q\delta[\mathbf{x} - \mathbf{x}_o(t)]$. The delta function has the following properties:

$$\begin{aligned} &\delta(\boldsymbol{x} - \boldsymbol{x_o}) = 0, \quad if \; \boldsymbol{x} \neq \boldsymbol{x_o}, \\ &\int dx \int dy \int dz \;\; \delta(\boldsymbol{x} - \boldsymbol{x_o}) = 1. \end{aligned} \tag{3.10}$$

The integral is taken over all space.

The current density is a vector quantity with units amperes per square meter. It is defined as the differential flux of charge, or the charge crossing a small surface element per second divided by the area of the surface. Current density can be visualized by considering how it is measured (Fig. 3.5). A small current probe of known area is adjusted in orientation at a point in space until

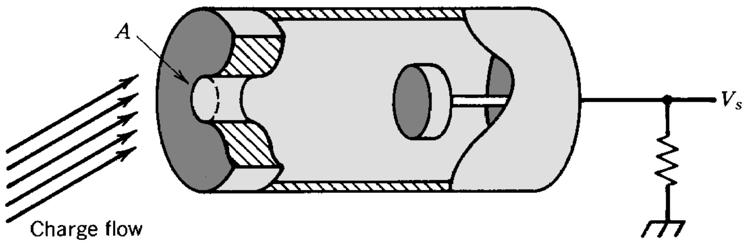


Figure 3.5 Probe to measure current density.

the current reading is maximized. The orientation of the probe gives the direction, and the current divided by the area gives the magnitude of the current density.

The general form of the Maxwell equations in MKS units is

$$\nabla\times\boldsymbol{E} = -\partial\boldsymbol{B}/\partial t, \tag{3.11}$$

$$\nabla\times\boldsymbol{B} = (1/c^2)\ \partial\boldsymbol{E}/\partial t + \mu_o\boldsymbol{j}, \tag{3.12}$$

$$\nabla\cdot\boldsymbol{E} = \rho/\varepsilon_o, \tag{3.13}$$

$$\nabla\cdot\boldsymbol{B} = 0. \tag{3.14}$$

Although these equations will not be derived, there will be many opportunities in succeeding chapters to discuss their physical implications. Developing an intuition and ability to visualize field distributions is essential for understanding accelerators. Characteristics of the Maxwell equations in the static limit and the concept of field lines will be treated in the next chapter.

No distinction has been made in Eqs. (3.1l)-(3.14) between various classes of charges that may constitute the charge density and current density. The Maxwell equations are sometimes written in terms of vector quantities **D** and **H**. These are subsidiary quantities in which the contributions from charges and currents in linear dielectric or magnetic materials have been extracted. They will be discussed in Chapter 5.

3.4 ELECTROSTATIC AND VECTOR POTENTIALS

The electrostatic potential is a scalar function of the electric field. In other words, it is specified by a single value at every point in space. The physical meaning of the potential can be demonstrated by considering the motion of a charged particle between two parallel plates (Fig. 3.6). We want to find the change in energy of a particle that enters that space between the plates with kinetic energy

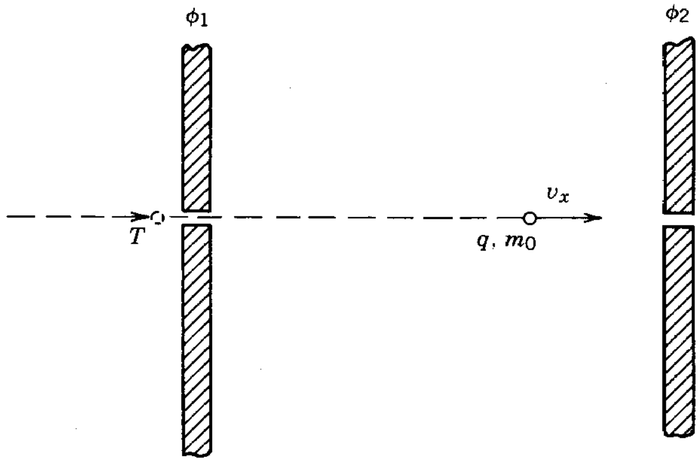


Figure 3.6 Acceleration of a particle between charged parallel plates.

T. Section 3.2 has shown that the electric field E_x, is uniform. The equation of motion is therefore

$$dp_x/dt = F_x = qE_x.$$

The derivative can be rewritten using the chain rule to give

$$(dp_x/dx)(dx/dt) = v_x \; dp_x/dx = qE_x.$$

The relativistic energy *E* of a particle is related to momentum by Eq. (2.37). Taking the derivative in x of both sides of Eq. (2.37) gives

$$c^2 p_x \; dp_x/dx = E \; dE/dx.$$

This can be rearranged to give

$$dE/dx = [c^2 p_x/E] \; dp_x/dx = v_x \; dp_x/dx. \tag{3.15}$$

The final form on the right-hand side results from substituting Eq. (2.38) for the term in brackets. The expression derived in Eq. (3.15) confirms the result quoted in Section 2.9. The right-hand side is dp_x/dt which is equal to the force F_x. Therefore, the relativistic form of the energy [Eq. (2.35)] is consistent with Eq. (2.6). The integral of Eq. (3.15) between the plates is

$$\Delta E = q \int dx E_x. \tag{3.16}$$

The *electrostatic potential* φ is defined by

$$\varphi = -\int \boldsymbol{E}\cdot\boldsymbol{dx}. \tag{3.17}$$

The change in potential along a path in a region of electric fields is equal to the integral of electric field tangent to the path times differential elements of pathlength. Thus, by analogy with the example of the parallel plates [Eq. (3.1 6)] ΔE = -qΔφ. If electric fields are static, the total energy of a particle can be written

$$E = m_o c^2 + T_o - q(\varphi - \varphi_o), \tag{3.18}$$

where T_o is the particle kinetic energy at the point where $\varphi = \varphi_o$.

The potential in Eq. (3.18) is not defined absolutely; a constant can be added without changing the electric field distribution. In treating electrostatic acceleration, we will adopt the convention that the zero point of potential is defined at the particle source (the location where particles have zero kinetic energy). The potential defined in this way is called the absolute potential (with respect to the source). In terms of the absolute potential, the total energy can be written

$$E = m_o c^2 - q\varphi$$

or

$$\gamma = 1 - q\varphi/m_o c^2. \tag{3.19}$$

Finally, the static electric field can be rewritten in the differential form,

$$\begin{aligned} \boldsymbol{E} &= -\nabla\varphi = (\partial\varphi/\partial x)\ \boldsymbol{u_x} + (\partial\varphi/\partial y)\ \boldsymbol{u_y} + (\partial\varphi/\partial z)\ \boldsymbol{u_z} \\ &= E_x\ \boldsymbol{u_x} + E_y\ \boldsymbol{u_y} + E_z\ \boldsymbol{u_z}. \end{aligned} \tag{3.20}$$

If the potential is known as a function of position, the three components of electric field can be found by taking spatial derivatives (the gradient operation). The defining equation for electrostatic fields [Eq. (3.3)] can be combined with Eq. (3.20) to give an expression to calculate potential directly from a specified distribution of charges

$$\varphi(\boldsymbol{x}) = \sum_n \frac{q_n/4\pi\varepsilon_o}{|\boldsymbol{x}-\boldsymbol{x_n}|}\ . \tag{3.21}$$

The denominator is the magnitude of the distance from the test charge to the *n*th charge. The integral form of this equation in terms of charge density is

$$\varphi(\boldsymbol{x}) = \frac{1}{4\pi\varepsilon_o} \iiint d^3x' \; \frac{\rho(\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|} \;. \tag{3.22}$$

Although Eq. 3.22 can be used directly to find the potential, we will usually use differential equations derived from the Maxwell equations combined with boundary conditions for such calculations (Chapter 4). The *vector potential* **A** is another subsidiary quantity that can be valuable for computing magnetic fields. It is a vector function related to the magnetic field through the vector curl operation

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}. \tag{3.23}$$

This relationship is general, and holds for time-dependent fields. We will use **A** only for static calculations. In this case, the vector potential can be written as a summation over source current density

$$\boldsymbol{A}(\boldsymbol{x}) = \frac{\mu_o}{4\pi} \iiint d^3x' \; \frac{\boldsymbol{j}(\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|} \;. \tag{3.24}$$

Compared to the electrostatic potential, the vector potential does not have a straightforward physical interpretation. Nonetheless, it is a useful computational device and it is helpful for the solution of particle orbits in systems with geometry symmetry. In cylindrical systems it is proportional to the number of magnetic field lines encompassed within particle orbits (Section 7.4).

3.5 INDUCTIVE VOLTAGE AND DISPLACEMENT CURRENT

The static concepts already introduced must be supplemented by two major properties of time-dependent fields for a complete and consistent theory of electrodynamics. The first is the fact that time-varying magnetic fields lead to electric fields. This is the process of magnetic induction. The relationship between inductively generated electric fields and changing magnetic flux is stated in Faraday's law. This effect is the basis of betatrons and linear induction accelerators. The second phenomenon, first codified by Maxwell, is that a time-varying electric field leads to a virtual current in space, the displacement current. We can verify that displacement currents "exist" by measuring the magnetic fields they generate. A current monitor such as a Rogowski loop enclosing an empty space with changing electric fields gives a current reading. The combination of inductive fields with the displacement current leads to predictions of electromagnetic oscillations.

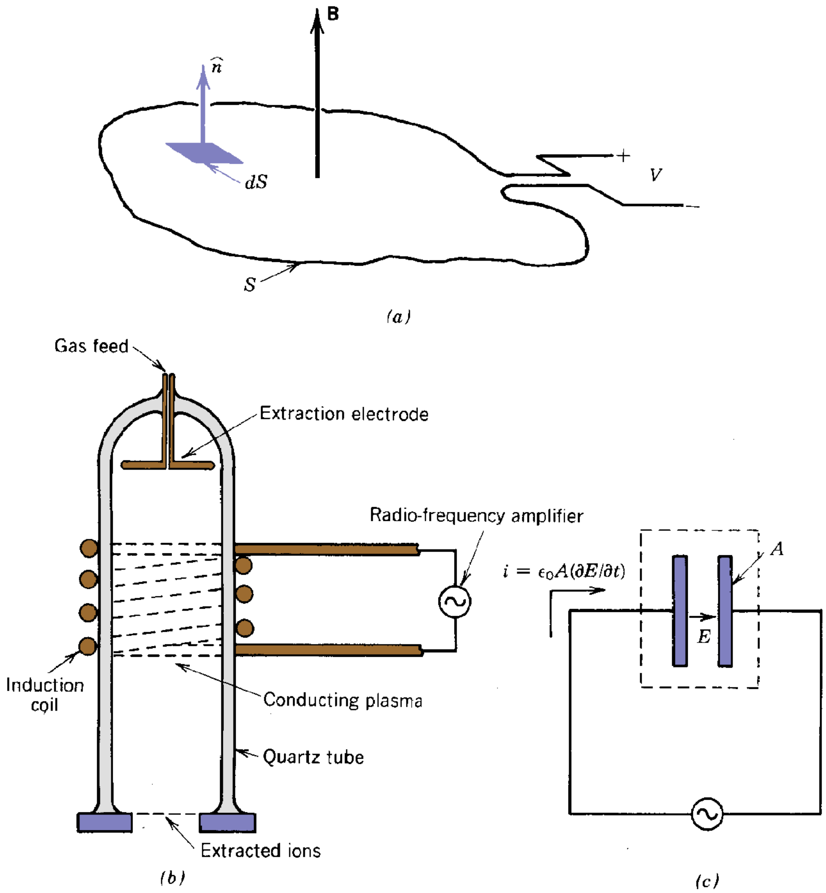


Figure 3.7 Inductive voltage and displacement currents. (*a*) Faraday's law. (*b*) Inductively coupled plasma source. (*c*) Alternating-current circuit with a parallel-plate capacitor.

Propagating and stationary electromagnetic waves are the bases for RF (radio-frequency) linear accelerators.

Faraday's law is illustrated in Figure 3.7a. A wire loop defines a surface *S*. The magnetic flux ψ passing through the loop is given by

$$\psi = \iint \boldsymbol{B}\cdot\boldsymbol{n} \; dS, \tag{3.25}$$

where **n** is a unit vector normal to *S* and *dS* is a differential element of surface area. Faraday's law states that a voltage is induced around the loop when the magnetic flux changes according to

$$V = -d\psi/dt. \tag{3.26}$$

The time derivative of ψ is the total derivative. Changes in ψ can arise from a time-varying field at constant loop position, motion of the loop to regions of different field magnitude in a static field, or a combination of the two.

The term *induction* comes from induce, to produce an effect without a direct action. This is illustrated by the example of Figure 3.7b. an inductively coupled plasma source. (A plasma is a conducting medium of hot, ionized gas.) Such a device is often used as an ion source for accelerators. In this case, the plasma acts as the loop. Currents driven in the plasma by changing magnetic flux ionize and heat the gas through resistive effects. The magnetic flux is generated by windings outside the plasma driven by a high-frequency ac power supply. The power supply couples energy to the plasma through the intermediary of the magnetic fields. The advantage of inductive coupling is that currents can be generated without immersed electrodes that may introduce contaminants.

The sign convention of Faraday's law implies that the induced plasma currents flow in the direction opposite to those of the driving loop. Inductive voltages always drive reverse currents in conducting bodies immersed in the magnetic field; therefore, oscillating magnetic fields are reduced or canceled inside conductors. Materials with this property are called diamagnetic. Inductive effects appear in the Maxwell equations on the right-hand side of Eq. (3.11). Application of the Stokes theorem (Section 4.1) shows that Eqs. (3.11) and (3.26) are equivalent.

The concept of displacement current can be understood by reference to Figure 3.7c. An electric circuit consists of an ac power supply connected to parallel plates. According to Eq. 3.9, the power supply produces an electric field E. between the plates by moving an amount of charge

$$Q = \varepsilon_o E_x A.$$

where A is the area of the plates. Taking the time derivative, the current through the power supply is related to the change in electric field by

$$i = \varepsilon_o \; A \; (\partial E_x/\partial t). \tag{3.27}$$

The partial derivative of Eq. (3.27) signifies that the variation results from the time variation of E_x with the plates at constant position. Suppose we considered the plate assembly as a black box without knowledge that charge was stored inside. In order to guarantee continuity of current around the circuit, we could postulate a virtual current density between the plates given bv

$$j_d = \varepsilon_o \; (\partial E_x/\partial t). \tag{3.28}$$

This quantity, the displacement current density, is more than just an abstraction to account for a change in space charge inside the box. The experimentally observed fact is that there are magnetic fields around the plate assembly that identical to those that would be produced by a real wore connecting the plates and carrying the current specified by Eq. (3.27) (see Section 4.6). There is thus a parallelism of time-dependent effects in electromagnetism. Time-varving magnetic fields produce electric fields, and changing electric fields produce magnetic fields. The coupling and interchange of electric and magnetic field energy is the basis of electromagnetic oscillations. Displacement currents or, equivalently, the generation of magnetic fields by time-varying electric fields, enter the Maxwell equations on the right side of Eq. (3.12). Noting that

$$c = 1/\sqrt{\varepsilon_o \mu_o}\ , \tag{3.29}$$

we see that the displacement current is added to any real current to determine the net magnetic field.

3.6 RELATIVISTIC PARTICLE MOTION IN CYLINDRICAL COORDINATES

Beams with cylindrical symmetry are encountered frequently in particle accelerators. For example, electron beams used in applications such as electron microscopes or cathode ray tubes have cylindrical cross sections. Section 3.7 will introduce an important application of the Lorentz force, circular motion in a uniform magnetic field. In order to facilitate this calculation and to derive useful formulas for subsequent chapters, the relativistic equations of motion for particles in cylindrical coordinates are derived in this section.

Cylindrical coordinates, denoted by (r, 0, z), are based on curved coordinate lines. We recognize immediately that equations of the form $dp_r/dt = F_r$ are incorrect. This form implies that particles subjected to no radial force move in a circular orbit (r = constant, $dp_r/dt = 0$). This is not consistent with Newton's first law. A simple method to derive the proper equations is to express d**p**/dt = **F** in Cartesian coordinates and make a coordinate transformation by direct substitution.

Reference to Figure 3.8 shows that the following equations relate Cartesian coordinates to cylindrical coordinates sharing a common origin and a common z axis, and with the line (r, 0, 0) lying on the x axis:

$$x = r\ \cos\theta,\quad y = r\ \sin\theta,\quad z = z, \tag{3.30}$$

and

$$r = \sqrt{x^2+y^2},\quad \theta = \tan^{-1}(y/x). \tag{3.31}$$

Motion along the z axis is described by the same equations in both frames, $dp_z/dt = F_z$. We will thus concentrate on equations in the (r, 0) plane. The Cartesian equation of motion in the x direction is

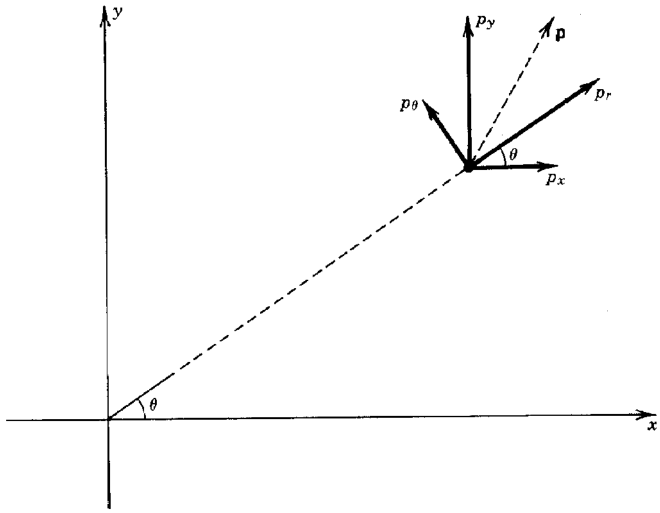


Figure 3.8 Relationship between cylindrical and Cartesian coordinates.

$$dp_x/dt = F_x. \tag{3.32}$$

Figure 3.8 shows that

$$p_x = p_r\cos\theta - p_\theta\sin\theta, \quad F_x = F_r\cos\theta - F_\theta\sin\theta.$$

Substituting in Eq. (3.32),

$$(dp_r/dt)\cos\theta - p_r\sin\theta(d\theta/dt) - (dp_\theta/dt)\sin\theta - p_\theta\cos\theta(d\theta/dt) = F_r\cos\theta - F_\theta\sin\theta.$$

The equation must hold at all positions, or at any value of θ. Thus, terms involving cosθ and sinθ must be separately equal. This yields the cylindrical equations of motion

$$dp_r/dt = F_r + [p_\theta\ d\theta/dt], \tag{3.33}$$

$$dp_\theta/dt = F_\theta - [p_r\ d\theta/dt]. \tag{3.34}$$

The quantities in brackets are correction terms for cylindrical coordinates. Equations (3.33) and (3.34) have the form of the Cartesian equations if the bracketed terms are considered as virtual forces. The extra term in the radial equation is called the centrifugal force, and can be rewritten

$$Centrigfugal\ force = \gamma m_o v_\theta^2/r, \tag{3.35}$$

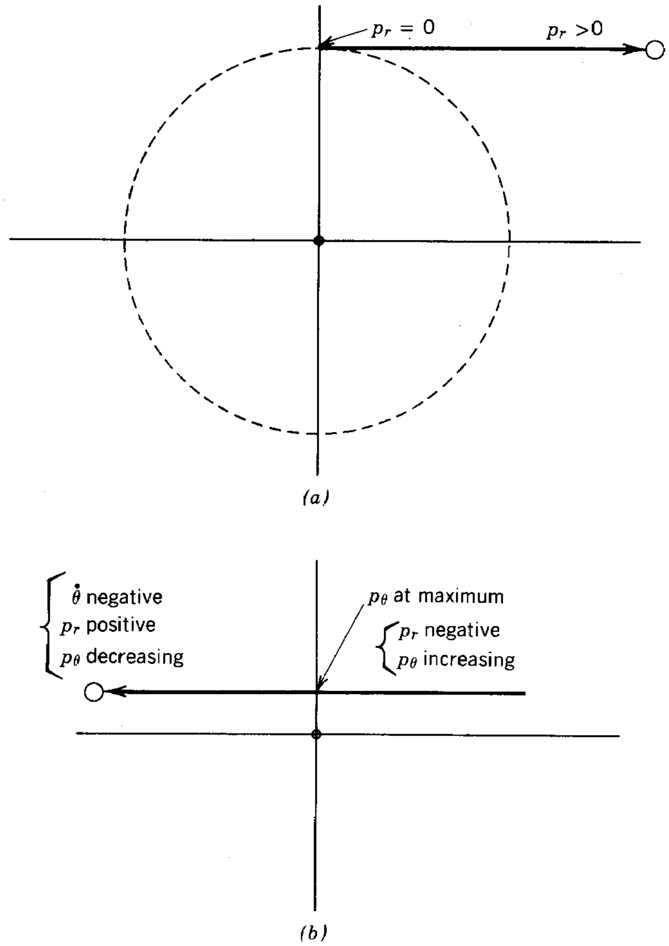


Figure 3.9 Virtual forces in a cylindrical coordinate system. (*a*) Centrifugal force: particle moves in straight line; p_r increases continually. (b) Coriolis force.

noting that $v_\theta = rd\theta/dt$. The bracketed term in the azimuthal equation is the Coriolis force, and can be written

$$\text{Coriolis force} = -\gamma m_o v_r v_\theta / r. \tag{3.36}$$

Figure 3.9 illustrates the physical interpretation of the virtual forces. In the first example, a particle moves on a force-free, straight-line orbit. Viewed in the cylindrical coordinate system, the particle (with no initial v_r) appears to accelerate radially, propelled by the centrifugal force. At large radius, when v_θ approaches 0, the acceleration appears to stop, and the particle moves outward at constant velocity. The Coriolis force is demonstrated in the second example. A particle from large radius moves in a straight line past the origin with nonzero impact parameter.

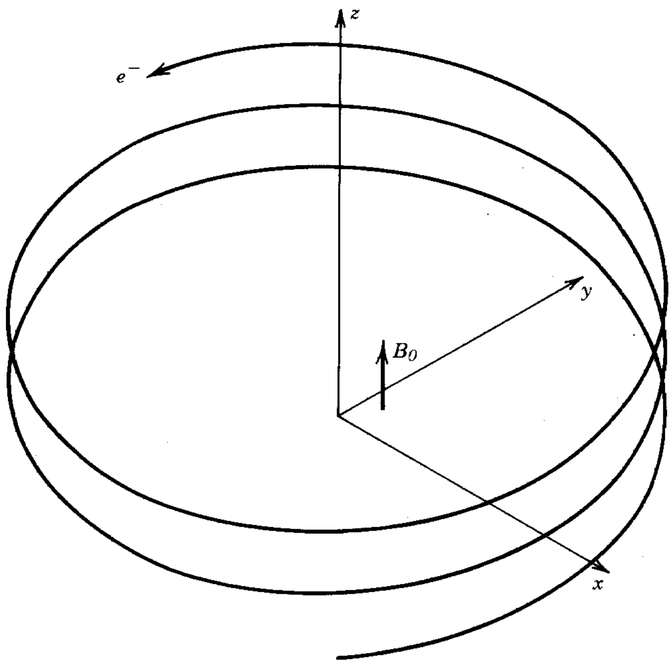


Figure 3.10 Helical orbit of a charged particle in a uniform magnetic field.

The azimuthal velocity, which was initially zero, increases as the particle moves inward with negative u, and decreases as the particle moves out. The observer in the cylindrical coordinate system notes a negative and then positive azimuthal acceleration.

Cylindrical coordinates appear extensively in accelerator theory. Care must be exercised to identify properly the orientation of the coordinates. For example, the z axis is sometimes aligned with the beam axis, white in other cases, the z axis may be along a symmetry axes of the accelerator. In this book, to avoid excessive notation, (r, 0, z) will be used for all cylindrical coordinate systems. Illustrations will clarify the geometry of each case as it is introduced.

3.7 MOTION OF CHARGED PARTICLES IN A UNIFORM MAGNETIC FIELD

Motion of a charged particle in a uniform magnetic field directed along the z axis, $\mathbf{B} = B_o\mathbf{u}_z$, is illustrated in Figure 3.10. Only the magnetic component of the Lorentz force is included. The equation of motion is

$$d\boldsymbol{p}/dt = d(\gamma m_o \boldsymbol{v})/dt = q\ \boldsymbol{v} \times \boldsymbol{B}. \tag{3.37}$$

By the nature of the cross product, the magnetic force is always perpendicular to the velocity of the particle. There is no force along a differential element of pathlength, **dx**. Thus, $\int \mathbf{F}\cdot\mathbf{dx} = 0$. According to Eq. (2.6), magnetic fields perform no work and do not change the kinetic energy of the particle. In Eq. (3.37), γ is constant and can be removed from the time derivative.

Because the force is perpendicular to B, there is no force along the z axis. Particles move in this direction with constant velocity. There is a force in the x-y plane. It is of constant magnitude (since the total particle velocity cannot change), and it is perpendicular to the particle motion. The projection of particle motion in the x-y plane is therefore a circle. The general three-dimensional particle orbit is a helix.

If we choose a cylindrical coordinate system with origin at the center of the circular orbit, then $dp_r/dt = 0$, and there is no azimuthal force. The azimuthal equation of motion [Eq. (3.34)] is satisfied trivially with these conditions. The radial equation [Eq. (3.33)] is satisfied when the magnetic force balances the centrifugal force, or

$$q v_\theta B_o = \gamma m_o v_\theta^2 / r.$$

The particle orbit radius is thus

$$r_g = \gamma m_o v_\theta / |q| B_o. \tag{3.38}$$

This quantity is called the *gyroradius*. It is large for high-momentum particles; the gyroradius is reduced by applying stronger magnetic field. The point about which the particle revolves is called the *gyrocenter*. Another important quantity is the angular frequency of revolution of the particle, the *gyrofrequency*. This is given by $\omega_g = v_\theta/r$, or

$$\omega_g = |q| B_o / \gamma m_o. \tag{3.39}$$

The particle orbits in Cartesian coordinates are harmonic,

$$x(t) - x_o = r_g \cos(\omega_g t),$$

$$y(t) - y_o = r_g \sin(\omega_g t),$$

where x_o and y_o are the coordinates of the gyrocenter. The gyroradius and gyrofrequency arise in all calculations involving particle motion in magnetic fields. Magnetic confinement of particles in circular orbits forms the basis for recirculating high-energy accelerators, such as the cyclotron, synchrotron, microtron, and betatron.

4

Steady-State Electric and Magnetic Fields

A knowledge of electric and magnetic field distributions is required to determine the orbits of charged particles in beams. In this chapter, methods are reviewed for the calculation of fields produced by static charge and current distributions on external conductors. Static field calculations appear extensively in accelerator theory. Applications include electric fields in beam extractors and electrostatic accelerators, magnetic fields in bending magnets and spectrometers, and focusing forces of most lenses used for beam transport.

Slowly varying fields can be approximated by static field calculations. A criterion for the static approximation is that the time for light to cross a characteristic dimension of the system in question is short compared to the time scale for field variations. This is equivalent to the condition that connected conducting surfaces in the system are at the same potential. Inductive accelerators (such as the betatron) appear to violate this rule, since the accelerating fields (which may rise over many milliseconds) depend on time-varying magnetic flux. The contradiction is removed by noting that the velocity of light may be reduced by a factor of 100 in the inductive media used in these accelerators. Inductive accelerators are treated in Chapters 10 and 11. The study of rapidly varying vacuum electromagnetic fields in geometries appropriate to particle acceleration is deferred to Chapters 14 and 15.

The static form of the Maxwell equations in regions without charges or currents is reviewed in Section 4.1. In this case, the electrostatic potential is determined by a second-order differential equation, the Laplace equation. Magnetic fields can be determined from the same equation by defining a new quantity, the magnetic potential. Examples of numerical (Section 4.2) and analog

(Section 4.3) methods for solving the Laplace equation are discussed. The numerical technique of successive overrelaxation is emphasized since it provides insight into the physical content of the Laplace equation. Static electric field calculations with field sources are treated in Section 4.4. The classification of charge is emphasized; a clear understanding of this classification is essential to avoid confusion when studying space charge and plasma effects in beams. The final sections treat the calculation of magnetic fields from specific current distributions through direct solution of the Maxwell equations (Section 4.5) and through the intermediary of the vector potential (Section 4.6).

4.1 STATIC FIELD EQUATIONS WITH NO SOURCES

When there are no charges or currents present. the Maxwell equations have the form

$$\nabla\cdot\boldsymbol{E} = 0, \tag{4.1}$$

$$\nabla\times\boldsymbol{E} = 0, \tag{4.2}$$

$$\nabla\cdot\boldsymbol{B} = 0, \tag{4.3}$$

$$\nabla\times\boldsymbol{B} = 0. \tag{4.4}$$

These equations resolve into two decoupled and parallel sets for electric fields [Eqs. (4.1) and (4.2)] and magnetic fields [Eqs. (4.3) and (4.4)]. Equations (4.1)-(4.4) hold in regions such as that shown in Figure 4.1. The charges or currents that produce the fields are external to the volume of interest. In electrostatic calculations, the most common calculation involves charge distributed on the surfaces of conductors at the boundaries of a vacuum region.

Equations (4.1)-(4.4) have straightforward physical interpretations. Similar conclusions hold for both sets, so we will concentrate on electric fields. The form for the divergence equation [Eq. (4.1)] in Cartesian coordinates is

$$\partial E_x/\partial x + \partial E_y/\partial y + \partial E_z/\partial z = 0. \tag{4.5}$$

An example is illustrated in Figure 4.2. The electric field is a function of x and y. The meaning of the divergence equation can be demonstrated by calculating the integral of the normal electric field over the surface of a volume with cross-sectional area A and thickness Δx. The integral over the left-hand side is $AE_x(x)$. If the electric field is visualized in terms of vector field lines, the integral is the flux of lines into the volume through the left-hand face. The electric field line flux out of the volume through the right-hand face is $AE_x(x + \Delta x)$.

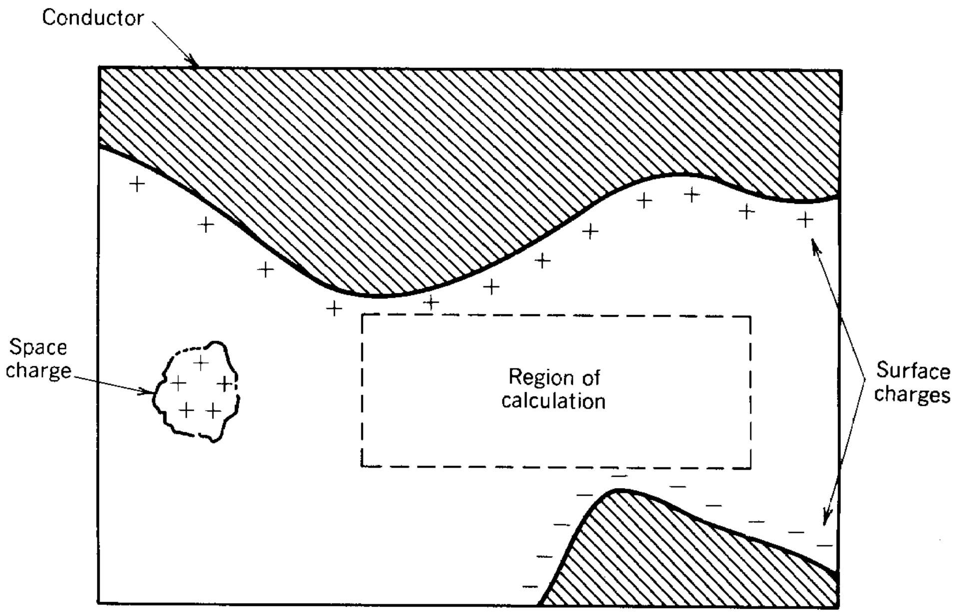


Figure 4.1 Calculation of static electric fields in a region with no source charge.

When the electric field is a smooth function of x, variations about a point can be approximated by a Taylor expansion. The right-hand integral is $A[E_x(x) + \Delta x\ \partial E_x/\partial x]$. The condition that $\partial E_x/\partial x = 0$ leads to a number of parallel conclusions.

1. The integrals of normal electric field over both faces of the volume are equal.
2. All field lines that enter the volume must exit.
3. The net flux of electric field lines into the volume is zero.
4. No field lines originate inside the volume.

Equation (4.5) is the three-dimensional equivalent of these statements.

The *divergence operator* applied to a vector quantity gives the effluence of the quantity away from a point in space. The divergence theorem can be written

$$\iint \boldsymbol{E}\cdot\boldsymbol{n}\ da = \iiint (\nabla\cdot\boldsymbol{E})\ dV. \tag{4.6}$$

Equation (4.6) states that the integral of the divergence of a vector quantity over all points of a volume is equal to the surface integral of the normal component of the vector over the surface of the volume. With no enclosed charges, field lines must flow through a volume as shown in Figure 4.3. The same holds true for magnetic fields. The main difference between electric and magnetic fields is that magnetic field lilies have zero divergence under all conditions, even in regions with currents. This means that magnetic field lines never emanate from a source point. They either extend indefinitely or are self-connected.

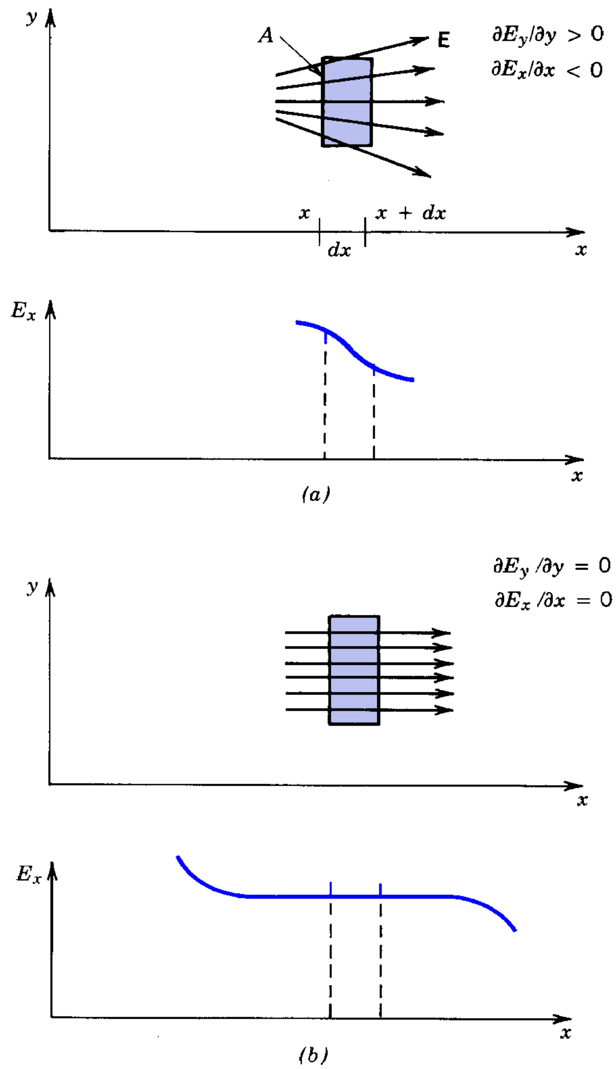


Figure 4.2 Relationship between gradients of electric field components and geometry of electric

The curl equations determine another geometric property of field lines. This property proceeds from the Stokes theorem, which states that

$$\int \mathbf{E}\cdot d\mathbf{l} = \iint (\nabla\times\mathbf{E})\cdot\mathbf{n}\ da. \tag{4.7}$$

The quantities in Eq. (4.7) are defined in Figure 4.4; *S* is a two-dimensional surface in space and **dl** is a length element oriented along the circumference. The integral on the left-hand side is taken

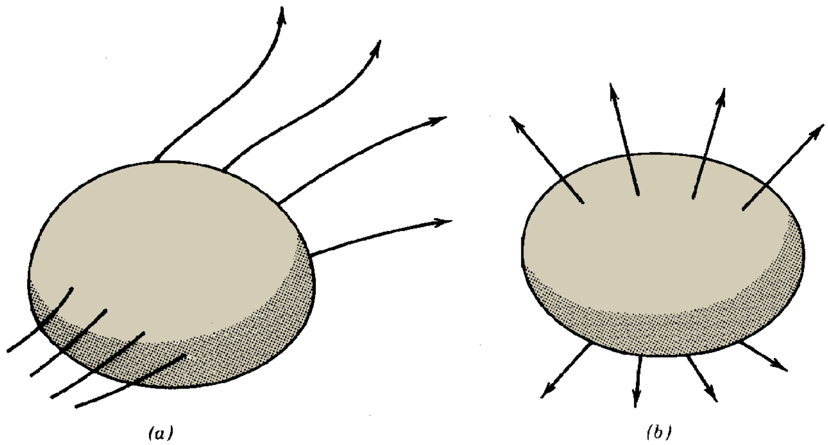


Figure 4.3 Geometry of vector field lines near a point with (a) zero divergence and (b) nonzero divergence.

around the periphery. The right-hand side is the surface integral of the component of the vector v x E normal to the surface. If the curl is nonzero at a point in space, then field lines form closed loops around the point. Figure 4.5 'illustrates points in vector fields with zero and nonzero curl.. The study of magnetic fields around current-carrying wires (Section 4.5) will illustrate a vector function with a nonzero curl.

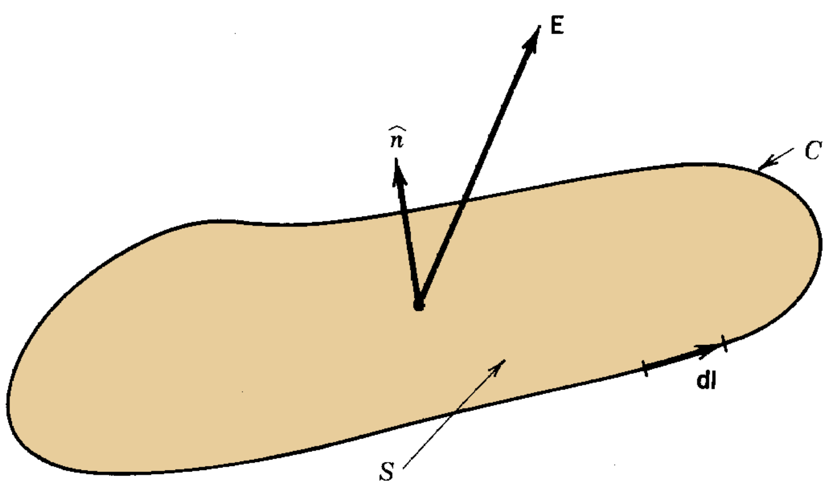


Figure 4.4 Definition of quantities used in Stokes theorem.

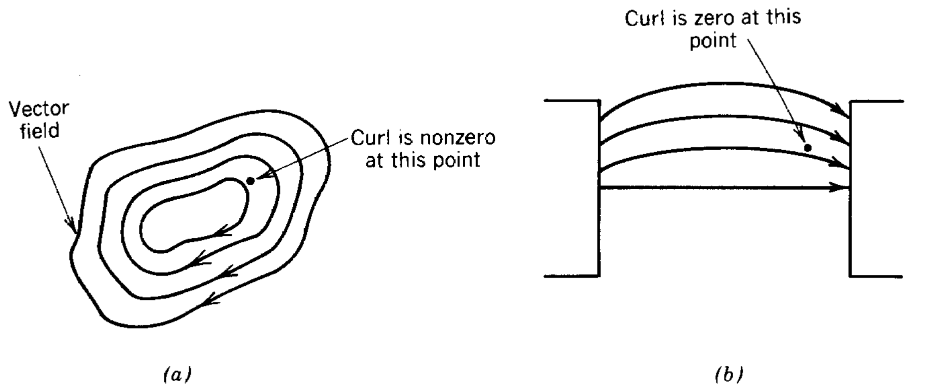


Figure 4.5 Geometry of vector field lines near a point with (*a*) nonzero curl and (*b*) zero curl.

For reference, the curl operator is written in Cartesian coordinates as

$$\nabla\times\boldsymbol{E} = \begin{bmatrix} \boldsymbol{u}_x & \boldsymbol{u}_y & \boldsymbol{u}_z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ E_x & E_y & E_z \end{bmatrix}. \tag{4.8}$$

The usual rule for evaluating a determinant is used. The expansion of the above expression is

$$\nabla\times = \boldsymbol{u}_x\left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) + \boldsymbol{u}_y\left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right) + \boldsymbol{u}_z\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right). \tag{4.9}$$

The electrostatic potential function φ can be defined when electric fields are static. The electric field is the gradient of this function,

$$\boldsymbol{E} = -\nabla\varphi. \tag{4.10}$$

Substituting for **E** in Eq. (4.1) gives

$$\nabla\cdot(\nabla\varphi) = 0,$$

or

$$\nabla^2\varphi = \partial^2\varphi/\partial x^2 + \partial^2\varphi/\partial y^2 + \partial^2\varphi/\partial z^2 = 0. \tag{4.11}$$

The operator symbolized by ∇^2 in Eq. (4.11) is called the Laplacian operator. Equation (4.11) is the *Laplace equation*. It determines the variation of φ (and hence **E**) in regions with no charge.

The curl equation is automatically satisfied through the vector identity $\nabla\times(\nabla\varphi)= 0$.

The main reason for using the Laplace equation rather than solving for electric fields directly is that boundary conditions can be satisfied more easily. The difficulty in solving the Maxwell equations directly lies in determining boundary conditions for vector fields on surrounding conducting surfaces. The electrostatic potential is a scalar function; we can show that the potential is a constant on a connected metal surface. Metals contain free electrons; an electric field parallel to the surface of a metal drives large currents. Electrons in the metal adjust their positions to produce a parallel component of field equal and opposite to the applied field. Thus, at a metal surface **E**(parallel) = 0 and **E**(normal) is unspecified. Equation (4.10) implies that electric field lines are always normal to surfaces of constant φ. This comes about because the gradient of a function (which indicates the direction in which a function has maximum rate of variation) must always be perpendicular to surfaces on which the function is constant (Fig. 4.6). Since a metal surface is everywhere perpendicular to the electric field, it must be an equipotential surface with the boundary condition φ = constant.

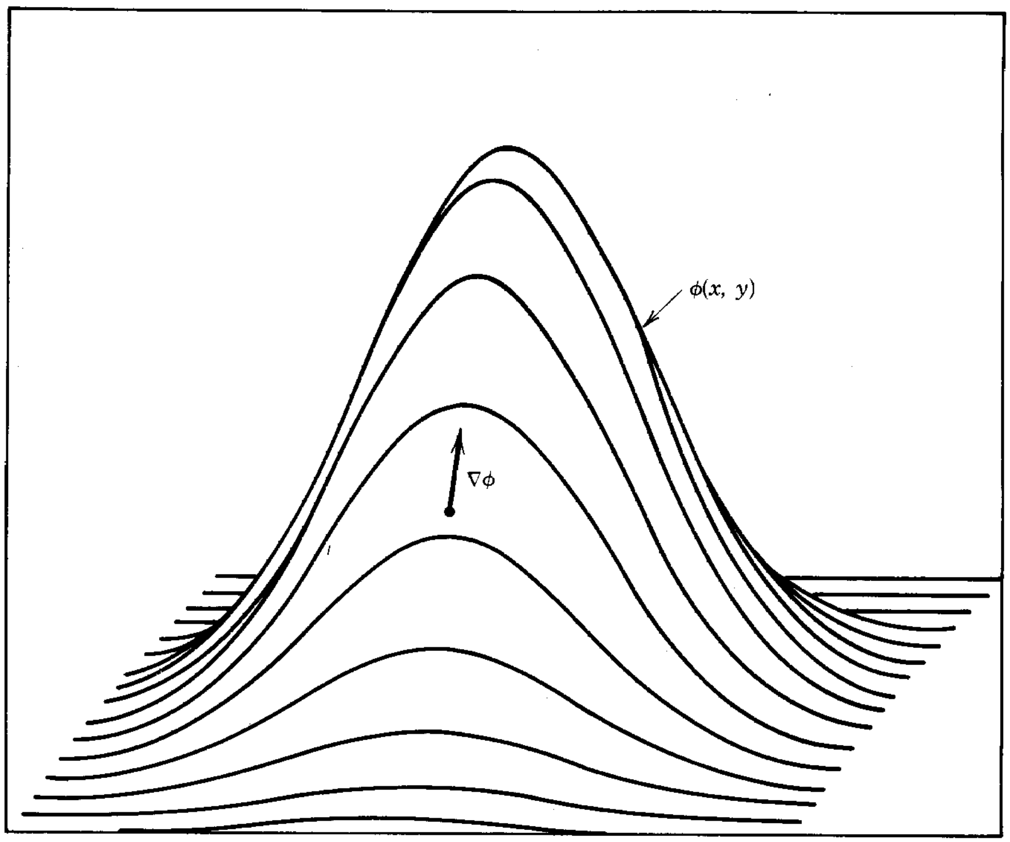


Figure 4.6 Three-dimensional plot of a scalar function $\phi(x, y)$, illustrating the orientation of the gradient vector $\nabla\phi$.

In summary, electric field lines have the following properties in source-free regions:

(a) Field lines are continuous. All lines that enter a volume eventually exit.
(b)Field lines do not kink, curl, or cross themselves.
(c)Field lines do not cross each other, since this would result in a point of infinite flux.
(d)Field lines are normal to surfaces of constant electrostatic potential.
(e) Electric fields are perpendicular to metal surfaces.

Fairly accurate electric field sketches can be made utilizing the laminar flow nature of electric field lines and the above properties. Even with the availability of digital computers, it is valuable to generate initial sketches of field patterns. This saves time and gives insight into the nature of fields. An example of an electrostatic field pattern generated by the method of squares is shown in

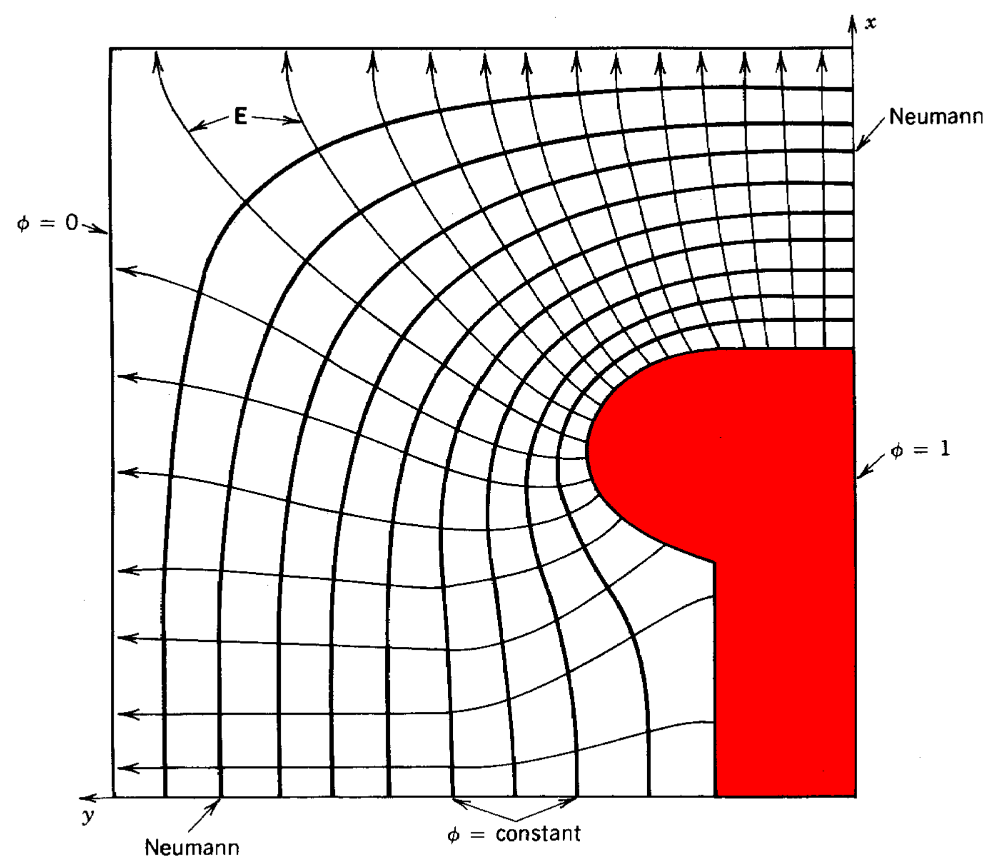


Figure 4.7 Equipotential lines and electric field lines around a high-voltage electrode in a grounded case. Electrode and case have infinite extent in z direction (out of page). Solution carried out only in one quadrant because of symmetry about x and y axes.

Figure 4.7. In this method, a number of equipotential lines between metal surfaces are sketched. Electric field lines normal to the equipotential lines and electrodes are added. Since the density of field lines is proportional to the distance between equipotentials, a valid final solution results when the elements between equipotential and field lines approach as close as possible to squares. The process is iterative and requires only some drawing ability and an eraser.

It is also possible to define formally a magnetic potential U_m such that

$$\nabla^2 U_m = 0. \tag{4.12}$$

The function U_m should not be confused with the vector potential. Methods used for electric field problems in source-free regions can also be applied to determine magnetic fields. We will defer use of Eq. (4.12) to Chapter 5. An understanding of magnetic materials is necessary to determine boundary conditions for U_m.

4.2 NUMERICAL SOLUTIONS TO THE LAPLACE EQUATION

The Laplace equation determines electrostatic potential as a function of position. Resulting electric fields can then be used to calculate particle orbits. Electrostatic problems may involve complex geometries with surfaces at many different potentials. In this case, numerical methods of analysis are essential.

Digital computers handle discrete quantities, so the Laplace equation must be converted from a continuous differential equation to a finite difference formulation. As shown in Figure 4.8, the quantity $\Phi(i, j, k)$ is defined at discrete points in space. These points constitute a three-dimensional mesh. For simplicity, the mesh spacing Δ between points in the three Cartesian directions is assumed uniform. The quantity Φ has the property that it equals $\varphi(x, y, z)$ at the mesh points. If φ is a smoothly varying function, then a linear interpolation of Φ gives a good approximation for φ at any point in space. In summary, Φ is a mathematical construct used to estimate the physical quantity, φ.

The Laplace equation for φ implies an algebraic difference equation for Φ. The spatial position of a mesh point is denoted by (i, j, k), with $x = i\Delta$, $y = j\Delta$, and $z = k\Delta$. The x derivative of φ to the right of the point (x, y, z) is approximated by

$$\partial\varphi(x+\Delta/2)/\partial x = [\Phi(i+1,j,k)-\Phi(i,j,k)]/\Delta. \tag{4.13}$$

A similar expression holds for the derivative at $x - \Delta/2$. The second derivative is the difference of derivatives divided by Δ, or

$$\frac{\partial}{\partial x}\left(\frac{\partial\varphi(x)}{\partial x}\right) \simeq \frac{1}{\Delta}\left(\frac{\partial\varphi(x+\Delta/2)}{\partial x} - \frac{\partial\varphi(x-\Delta/2)}{\partial x}\right)$$

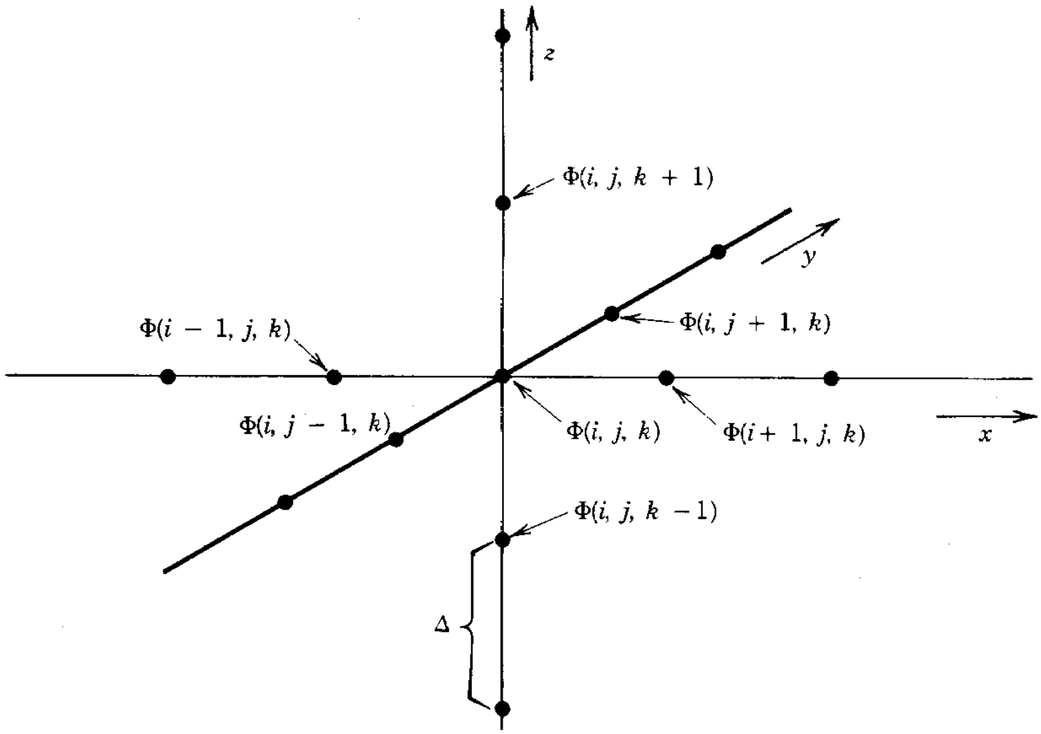


Figure 4.8 Finite difference approximation for electrostatic potential.

Combining expressions,

$$\frac{\partial^2 \varphi}{\partial x^2} = \frac{\Phi(i+1,j,k) - 2\Phi(i,j,k) + \Phi(i-1,j,k)]}{\Delta^2} . \tag{4.14}$$

Similar expressions can be found for the $\partial^2\varphi/\partial y^2$ and $\partial^2\varphi/\partial z^2$ terms. Setting $\nabla^2\varphi 1 = 0$ implies

$$\Phi(i,j,k) = 1/6\ [\Phi(i+1,j,k) + \Phi(i-1,j,k) + \Phi(i,j+1,k) + \Phi(i,j-1,k) + \Phi(i,j,k+1) + \Phi(i,j,k-1)]. \tag{4.15}$$

In summary, (1) Φ(i, j, k) is a discrete function defined as mesh points, (2) the interpolation of Φ(i, j, k) approximates φ(x, y, z), and (3) if φ(x, y, z) satisfies the Laplace equation, then Φ(i, j, k) is determined by Eq. (4.15).

According to Eq. (4.15), individual values of Φ(i, j, k) are the average of their six neighboring points. Solving the Laplace equation is an averaging process; the solution gives the smoothest flow of field lines. The net length of all field lines is minimized consistent with the boundary conditions. Therefore, the solution represents the state with minimum field energy (Section 5.6).

There are many numerical methods to solve the finite difference form for the Laplace equation. We will concentrate on the *method of successive ouerrelaxation*. Although it is not the fastest method of solution, it has the closest relationship to the physical content of the Laplace equation. To illustrate the method, the problem will be formulated on a two-dimensional, square mesh. Successive overrelaxation is an iterative approach. A trial solution is corrected until it is close to a valid solution. Correction consists of sweeping through all values of an intermediate solution to calculate *residuals*, defined by

$$R(i,j) = \tfrac{1}{4}[\Phi(i+1,j) + \Phi(i-1,j) + \Phi(i,j+1) + \Phi(i,j-1)] - \Phi(i,j) \tag{4.16}$$

If *R(i, j)* is zero at all points, then Φ(i, j) is the desired solution. An intermediate result can be improved by adding a correction factor proportional to R(i, j),

$$\Phi(i,j)_{n+1} = \Phi(i,j)_n + \omega R(i,j)_n. \tag{4.17}$$

The value $\omega = 1$ is the obvious choice, but in practice values of ω between 1 and 2 produce a faster convergence (hence the term overrelaxation). The succession of approximations resembles a time-dependent solution for a system with damping, relaxing to its lowest energy state. The elastic sheet analog (described in Section 4.3) is a good example of this interpretation. Figure 4.9 shows intermediate solutions for a one-dimensional mesh with 20 points and with $\omega = 1.00$. Information on the boundary with elevated potential propagates through the mesh.

The method of successive overrelaxation is quite slow for large numbers of points. The number of calculations on an $n \times n$ mesh is proportional to n^2. Furthermore, the number of iterations necessary to propagate errors out of the mesh is proportional to n. The calculation time increases as n^3 . A BASIC algorithm to relax internal points in a 40 x 48 point array is listed in Table 4.1. Corrections are made continuously during the sweep. Sweeps are first carried out along the *x* direction and then along the *y* direction to allow propagation of errors in both directions. The electrostatic field distribution in Figure 4.10 was calculated by a relaxation program.

Advanced methods for solving the Laplace equation generally use more efficient algorithms based on Fourier transforms. Most available codes to solve electrostatic problems utilize a more complex mesh. The mesh may have a rectangular or even triangular divisions to allow a close match to curved boundary surfaces.

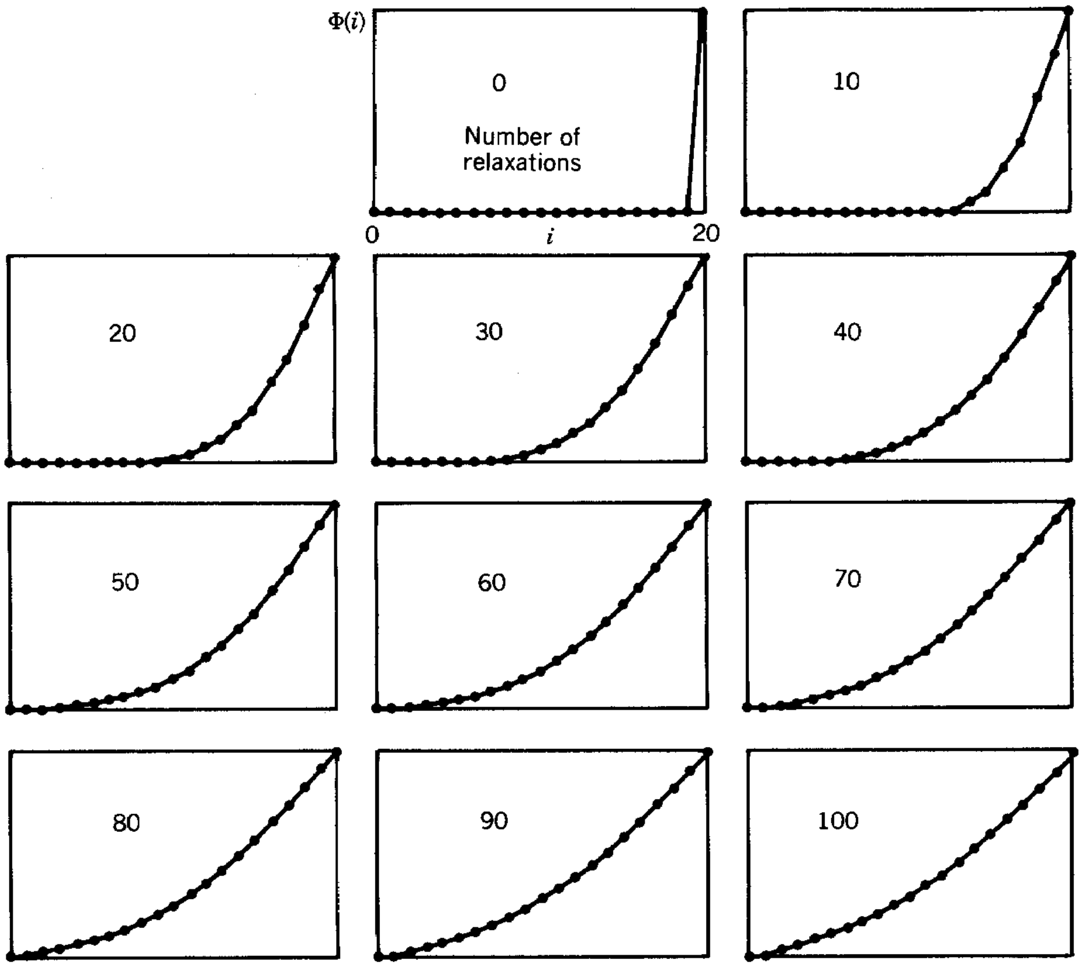


Figure 4.9 Relaxation method for solving the Laplace equation. Successive one-dimensional solutions for electrostatic potential between charged plates as a function of the number of mesh relaxations. Initial conditions: $\Phi(1) - \Phi(19) = 0$; $\Phi(20) = 1$.

Boundary conditions present special problems and must be handled differently from internal points representing the vacuum region. Boundary points may include those on the actual boundary of the calculational mesh, or points on internal electrodes maintained at a constant potential. The latter points are handled easily. They are marked by a flag to indicate locations of nonvariable potential. The relaxation calculation is not performed at such points. Locations on the mesh boundary have no neighbors outside the mesh, so that Eq. (4.16) can not be applied. If these points have constant potential, there is no problem since the residual need not be computed. Constant-potential points constitute a Dirichlet boundary condition.

The other commonly encountered boundary specification is the Neumann condition in which the normal derivative of the potential at the boundary is specified. In most cases where the Neumann condition is used, the derivative is zero, so that there is no component of the electric field normal

TABLE 4.1 Subroutines: Successive Over-Relaxation Program

```
1000 REM ---RELAXATION SUBROUTINES FOR POINTS
1010 REM ---INTERNAL TO BOUNDARY

1200 REM ---SWEEP ALONG I DIRECTION
1202 FOR J=1 TO JMAX
1204 FOR I=1 TO IMAX
1206 IF IB(I,J)<>0 THEN GOTO 1212:REM ---SPECIFIED POTENTIAL
1208 RES=(P(I+1,J)+P(I-1,J)+P(I,J+1)+P(I,J-1))/4-P(I,J)
1210 P(I,J)=P(I,J)+OMEGA*RES
1212 NEXT I
1214 NEXT J
1216 RETURN

1300 REM ---SWEEP ALONG J DIRECTION
1302 FOR I=1 TO IMAX
1304 FOR J=1 TO JMAX
1306 IF IB(I,J)<>0 THEN GOTO 1312:REM ---SPECIFIED POTENTIAL
1308 RES=(P(I+1,J)+P(I-1,J)+P(I,J+1)+P(I,J-1))/4-P(I,J)
1310 P(I,J)=P(I,J)+OMEGA*RES
1312 NEXT J
1314 NEXT I
1316 RETURN
```

to the boundary. This condition applies to boundaries with special symmetry, such as the axis in a cylindrical calculation or a symmetry plane of a periodic system. Residues can be calculated at Neumann boundaries since the potential outside the mesh is equal to the potential at the first point inside the mesh. For example, on the boundary $i = 0$, the condition $\Phi(-1, j) = \Phi(+1, j)$ holds. The residual is

$$R(0,j) = \tfrac{1}{4}\,[\Phi(0,j+1) + 2\Phi(1,j) + \Phi(0,j-1)] - \Phi(0,j). \qquad (4.18)$$

Two-dimensional systems with cylindrical symmetry are often encountered in accelerator applications. Potential is a function of (r, z), with no azimuthal dependence. The Laplace equation for a cylindrical system is

$$\frac{1}{r}\,\frac{\partial}{\partial r}\left(r\frac{\partial\varphi}{\partial r}\right) + \frac{\partial^2\varphi}{\partial z^2} = 0. \qquad (4.19)$$

The finite difference form for the Laplace equation for this case is

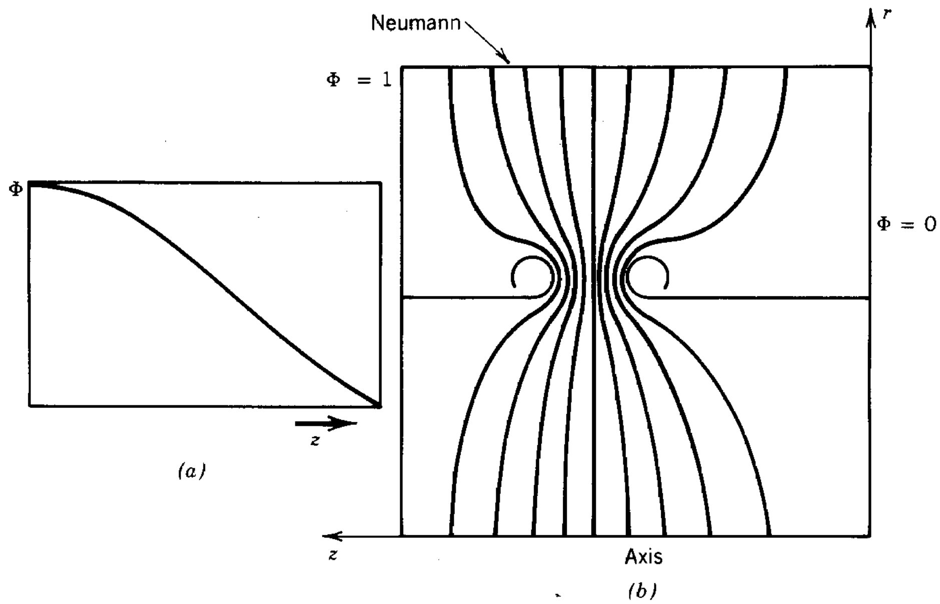


Figure 4.10 Two-dimensional solution for equipotentials between cylindrical electrodes of a drift tube linear accelerator calculated by the method of successive overrelaxation ($\omega = 1.75$). (*a*) Variation of potential on axis, $\Phi(0, z)$. (*b*) Boundary conditions, equipotential lines.

$$\Phi(i,j) = \frac{1}{4}\left[\frac{(i+½)\Phi(i+1,j)}{i} + \frac{(i-½)\Phi(i-1,j)}{i} + \Phi(i,j+1) + \Phi(i,j-1)\right] . \tag{4.20}$$

where r = iΔ and z = jΔ.

Figure 4.10 shows results for a relaxation calculation of an electrostatic immersion lens. It consists of two cylinders at different potentials separated by a gap. Points of constant potential and Neumann boundary conditions are indicated. Also shown is the finite difference approximation for the potential variation along the axis, 0(0, z). This data can be used to determine the focal properties of the lens (Chapter 6).

4.3 ANALOG METHODS TO SOLVE THE LAPLACE EQUATION

Analog methods were used extensively to solve electrostatic field problems before the advent of digital computers. We will consider two analog techniques that clarify the nature of the Laplace equation. The approach relies on finding a physical system that obeys the Laplace equation but that allows easy measurements of a characteristic quantity (the analog of the potential).

One system, the tensioned elastic sheet, is suitable for two-dimensional problems (symmetry along the z axis). As shown in Figure 4.11, a latex sheet is stretched with uniform tension on a frame. If the sheet is displaced vertically a distance *H(x, y)*, there will be vertical restoring forces. In equilibrium, there is vertical force balance at each point. The equation of force balance can be determined from the finite difference approximation defined in Figure 4.11. In terms of the surface tension, the forces to the left and right of the point (iΔ, jΔ) are

$$F[(i-½)\Delta] = T\ [H(i\Delta,j\Delta)-H([i-1]\Delta,j\Delta)]/\Delta,$$

$$F[(i+½)\Delta] = T\ [H([i+1]\Delta,j\Delta)-H(i\Delta,j\Delta)]/\Delta.$$

Similar expressions can be determined for the *y* direction. The height of the point (iΔ, jΔ) is constant in time; therefore,

$$F[(i-½)\Delta] = -F[(i+½)\Delta],$$

and

$$F[(j+½)\Delta] = -F[(j-½)\Delta].$$

Substituting for the forces shows that the height of a point on a square mesh is the average of its four nearest neighbors. Thus, inverting the arguments of Section 4.2, *H(x, y)* is described by the two-dimensional Laplace equation

$$\partial^2 H(x,y)/\partial x^2 + \partial^2 H(x,y)/\partial y^2 = 0.$$

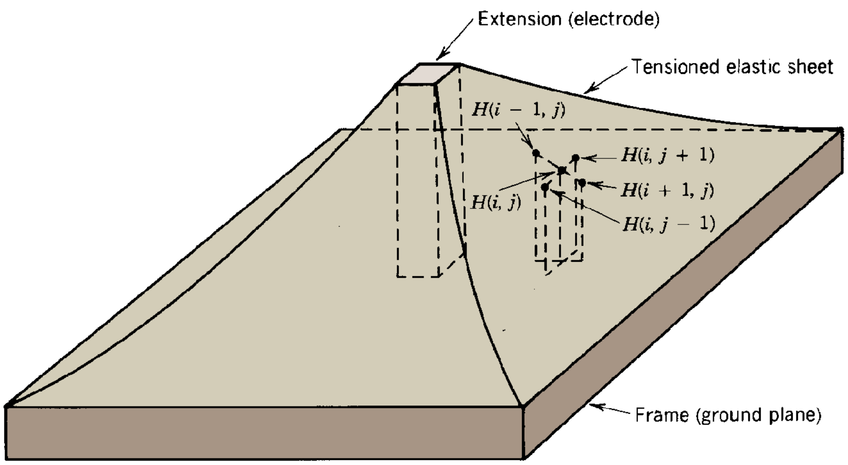


Figure 4.11 Tensioned elastic sheet analog for electrostatic potential.

Height is the analog of potential. To make an elastic potential solution, parts are cut to the shape of the electrodes. They are fastened to the frame to displace the elastic sheet up or down a distance proportional to the electrode potential. These pieces determine equipotential surfaces. The frame is theground plane.

An interesting feature of the elastic sheet analog is that it can also be used to determine orbits of charged particles in applied electrostatic fields. Neglecting rotation, the total energy of a ball bearing on the elastic sheet is $E = T + mgh(x, y)$, where g is the gravitational constant. The transverse forces acting on a ball bearing on the elastic sheet are $F_x = \partial H/\partial x$ and $F_y = \partial H/\partial y$. Thus, ball bearings on the elastic sheet follow the same orbits as charged particles in the analogous electrostatic potential, although over a considerably longer time scale.

Figure 4.12 is a photograph of a model that demonstrates the potentials in a planar electron extraction gap with a coarse grid anode made of parallel wires. The source of the facet lens effect associated with extraction grids (Section 6.5) is apparent.

A second analog technique, the electrolytic tank, permits accurate measurements of potential distributions. The method is based on the flow of current in a liquid medium of constant-volume resistivity, ρ (measured in units of ohm-meters). A dilute solution of copper sulfate in water is a common medium. A model of the electrode structure is constructed to scale from copper sheet and immersed in the solution. Alternating current voltages with magnitude proportional to those in the actual system are applied to the electrodes.

According to the definition of volume resistivity, the current density is proportional to the electric field

$$\boldsymbol{E} = \rho\, \boldsymbol{j}$$

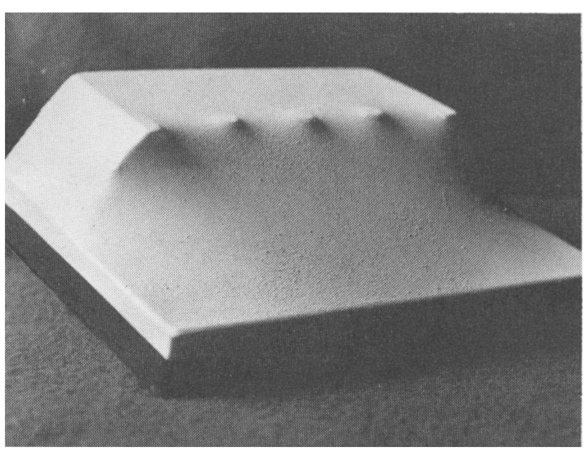


Figure 4.12 Elastic sheet analog for electrostatic potential near an extraction grid. Elevated section represents a high-voltage electrode surrounded by a grounded enclosure. Note the distortion of the potential near the grid wires that results in focusing of extracted particles. (Photograph and model by the author. Latex courtesy of the Hygenic Corporation.)

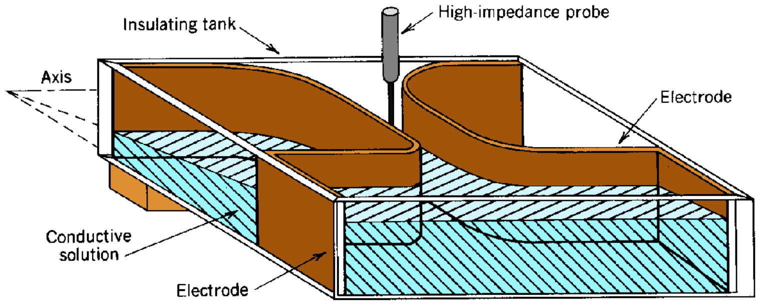


Figure 4.13 Electrolytic tank analog to solve the Laplace equation in a cylindrically symmetric system.

or

$$\boldsymbol{j} = -\nabla\varphi/\rho. \tag{4.21}$$

The steady-state condition that charge at any point in the liquid is a constant implies that all current that flows into a volume element must flow out. This condition can be written

$$\nabla\cdot\boldsymbol{j} = 0. \tag{4.22}$$

Combining Eq. (4.21) with (4.22), we find that potential in the electrolytic solution obeys the Laplace equation.

In contrast to the potential in the real system, the potential in the electrolytic analog is maintained by a real current flow. Thus, energy is available for electrical measurements. A high-impedance probe can be inserted into the solution without seriously perturbing the fields. Although the electrolytic method could be applied to three-dimensional problems, in practice it is usually limited to two-dimensional simulations because of limitations on insertion of a probe. A typical setup is shown in Figure 4.13. Following the arguments given above, it is easy to show that a tipped tank can be used to solve for potentials in cylindrically symmetric systems.

4.4 ELECTROSTATIC QUADRUPOLE FIELD

Although numerical calculations are often necessary to determine electric and magnetic fields in accelerators, analytic calculations have advantages when they are tractable. Analytic solutions show general features and scaling relationships. The field expressions can be substituted into equations of motion to yield particle orbit expressions in closed form. Electrostatic solutions for a wide variety of electrode geometries have been derived. In this section. we will examine the quadrupole field, a field configuration used in all high-energy transport systems. We will concentrate on the electrostatic quadrupole; the magnetic equivalent will be discussed in Chapter

5.

The most effective procedure to determine electrodes to generate quadrupole fields is to work in reverse, starting with the desired electric field distribution and calculating the associated potential function. The equipotential lines determine a set of electrode surfaces and potentials that generate the field. We assume the following two-dimensional fields:

$$E_x = +kx = E_o x/a, \tag{4.23}$$

$$E_y = -ky = -E_o y/a. \tag{4.24}$$

It is straightforward to verify that both the divergence and curl of **E** are zero. The fields of Eqs. (4.23) and (4.24) represent a valid solution to the Maxwell equations in a vacuum region. The electric fields are zero at the axis and increase (or decrease) linearly with distance from the axis. The potential is related to the electric field by

$$\partial\varphi/\partial x = -E_o x/a, \quad \partial\varphi/\partial y = +E_o y/a,$$

Integrating the partial differential equations

$$\varphi = -E_o x^2/2a + f(y) + C, \quad \varphi = +E_o y^2/2a + g(x) + C'.$$

Taking $\varphi(0, 0) = 0$, both expressions are satisfied if

$$\varphi(x,y) = (E_o/2a)\ (y^2 - x^2). \tag{4.25}$$

This can be rewritten in a more convenient, dimensionless form:

$$\frac{\varphi(x,y)}{E_o a/2} = \left(\frac{y}{a}\right)^2 - \left(\frac{x}{a}\right)^2 . \tag{4.26}$$

Equipotential surfaces are hyperbolas in all four quadrants. There is an infinite set of electrodes that will generate the fields of Eqs. (4.23) and (4.24). The usual choice is symmetric electrodes on the equipotential lines $\varphi_o = \pm E_o a/2$. Electrodes, field lines, and equipotential surfaces are plotted in Figure 4.14. The quantity a is the minimum distance from the axis to the electrode, and E_o is the electric field on the electrode surface at the position closest to the origin. The equipotentials in Figure 4.14 extend to infinity. In practice, focusing fields are needed only near the axis. These fields are not greatly affected by terminating the electrodes at distances a few times a from the axis.

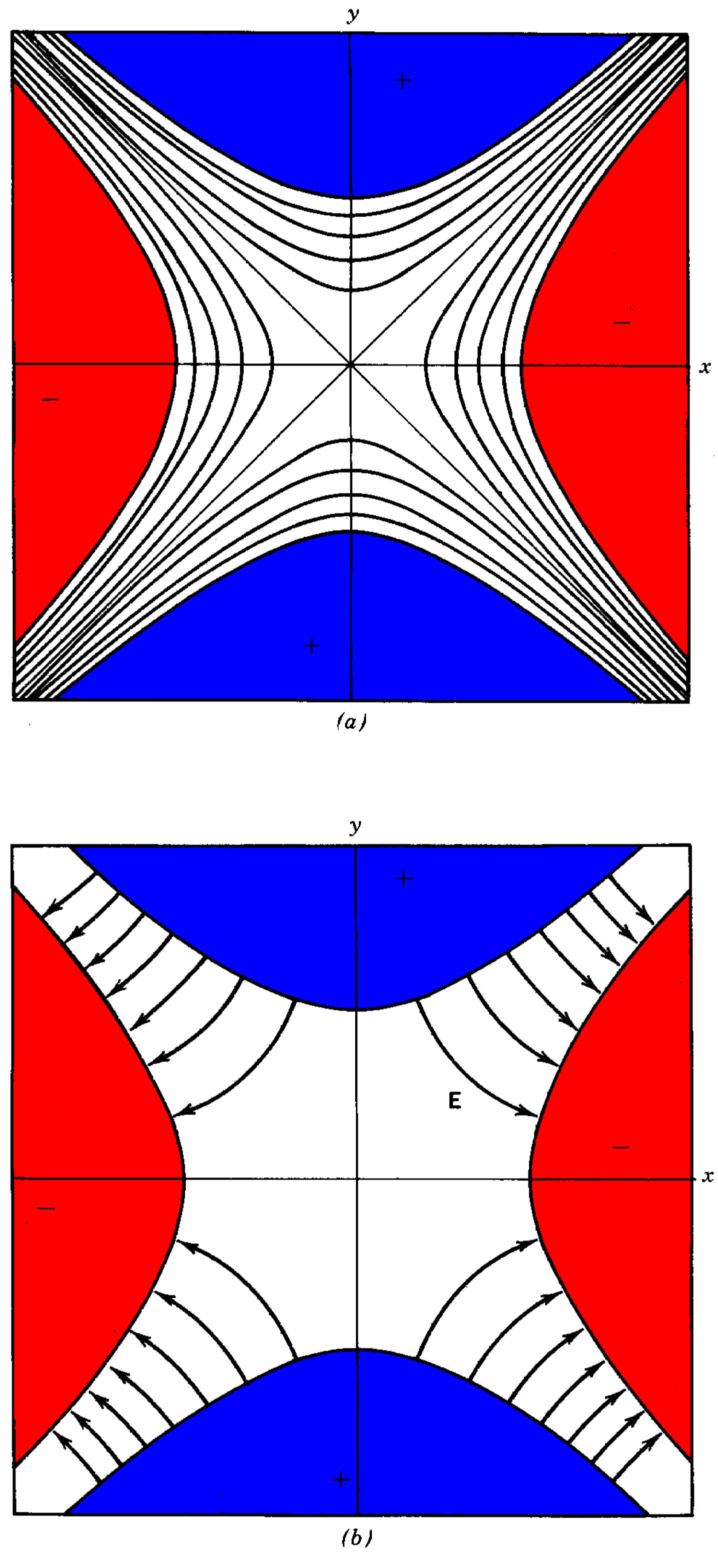


Figure 4.14 Electrostatic quadrupole field. (*a*) Equipotential lines. (*b*) Field lines.

4.5 STATIC ELECTRIC FIELDS WITH SPACE CHARGE

Space charge is charge density present in the region in which an electric field is to be calculated. Clearly, space charge is not included in the Laplace equation, which describes potential arising from charges on external electrodes. In accelerator applications, space charge is identified with the charge of the beam; it must be included in calculations of fields internal to the beam. Although we will not deal with beam self-fields in this book, it is useful to perform at least one space charge calculation. It gives insight into the organization of various types of charge to derive electrostatic solutions. Furthermore, we will derive a useful formula to estimate when beam charge can be neglected.

Charge density can be conveniently divided into three groups: (1) applied, (2) dielectric, and (3) space charge. Equation 3.13 can be rewritten

$$\nabla\cdot\boldsymbol{E} = (\rho_1 + \rho_2 + \rho_3)/\varepsilon_o. \tag{4.27}$$

The quantity ρ_1 is the charge induced on the surfaces of conducting electrodes by the application of voltages. The second charge density represents charges in *dielectric materials*. Electrons in dielectric materials cannot move freely. They are bound to a positive charge and can be displaced only a small distance. The dielectric charge density can influence fields in and near the material. Electrostatic calculations with the inclusion of ρ_2 are discussed in Chapter 5. The final charge density, ρ_3, represents space charge, or free charge in the region of the calculation. This usually includes the charge density of the beam. Other particles may contribute to ρ_3, such as low-energy electrons in a neutralized ion beam.

Electric fields have the property of superposition. Given fields corresponding to two or more charge distributions, then the total electric field is the vector sum of the individual fields if the charge distributions do not perturb one another. For instance, we could calculate electric fields individually for each of the charge components, $\mathbf{E}_1$, $\mathbf{E}_2$, and $\mathbf{E}_3$. The total field is

$$\boldsymbol{E} = \boldsymbol{E}_1(applied) + \boldsymbol{E}_2(dielectric) + \boldsymbol{E}_3(spacecharge). \tag{4.28}$$

Only the third component occurs in the example of Figure 4.15. The cylinder with uniform charge density is a commonly encountered approximation for beam space charge. The charge density is constant, ρ_o, from $r = 0$ to $r = r_b$. There is no variation in the axial (z) or azimuthal (θ) directions so that $\partial/\partial z = \partial/\partial\theta = 0$. The divergence equation (3.13) implies that there is only a radial component of electric field. Because all field lines radiate straight outward (or inward for $\rho_o < 0$), there can be no curl, and Eq. (3.11) is automatically satisfied.

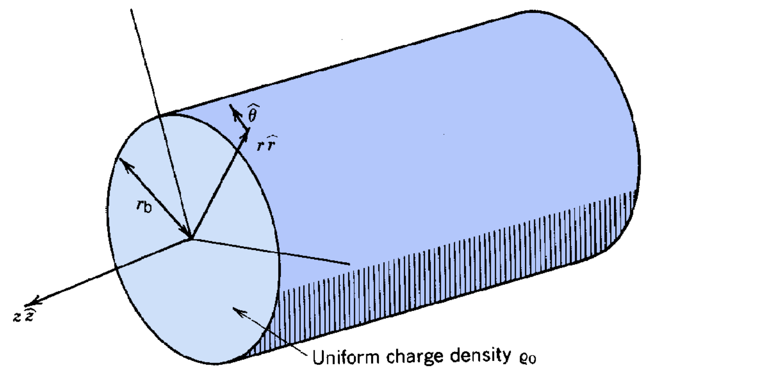


Figure 4.15 Quantities for calculating electric fields in and around a cylinder with uniform charge density (ρ_0).

Inside the charge cylinder, the electric field is determined by

$$\frac{1}{r}\frac{d(rE_r)}{dr} = \frac{\rho_o}{\varepsilon_o} . \tag{4.29}$$

Electric field lines are generated by the charge inside a volume. The size of the radial volume element goes to zero near the origin. Since no field lines can emerge from the axis, the condition $E_r(r = 0) = 0$ must hold. The solution of Eq. (4.29) is

$$E_r(r < r_b) = \frac{\rho_o r}{2\varepsilon_o} . \tag{4.30}$$

Outside the cylinder, the field is the solution of Eq. (4.29) with the right-hand side equal to zero. The electric field must be a continuous function of radius in the absence of a charge layer. (A charge layer is a finite quantity of charge in a layer'of zero thickness; this is approximately the condition on the surface of an electrode.) Thus, $E_r(r = r_b^+) = E_r(r = r_b^-)$, so that

$$E_r(r > r_b) = \frac{\rho_o r_b^2}{2\varepsilon_o r} . \tag{4.31}$$

The solution is plotted in Figure 4,16. The electric field increases linearly away from the axis in the charge region. It decreases as $1/r$ for $r > r_b$ because the field lines are distributed over a larger area.

The problem of the charge cylinder can also be solved through the electrostatic potential. The Poisson equation results when the gradient of, the potential is substituted in Eq. (3.13):

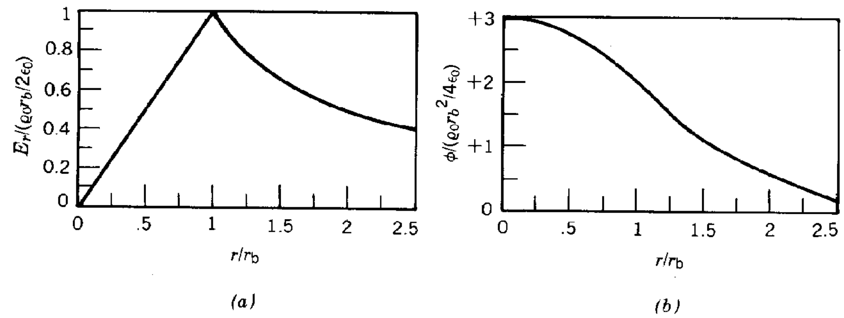


Figure 4.16 Solutions for the electrostatic fields of a cylinder with uniform charge density. (*a*) Normalized radial electric field $E_r/(\rho_0 r_b/2\varepsilon_0)$. (*b*) Normalized electrostatic potential, $\phi/(\rho_0 r_b^2/4\varepsilon_0)$.

$$\nabla^2\varphi = -\frac{\rho(x)}{\varepsilon_o}, \tag{4.32}$$

or

$$\frac{1}{r}\frac{d}{dr}\left(r\,\frac{d\varphi}{dr}\right) = -\frac{\rho_o}{\varepsilon_o}. \tag{4.33}$$

The solution to the Poisson equation for the charge cylinder is

$$\varphi(r < r_b) = -\frac{\rho_o r^2}{4\varepsilon_o}, \tag{4.34}$$

$$\varphi(r > r_b) = -\frac{\rho_o r_b^2}{4\varepsilon_o}\left(2\,\ln\frac{r}{r_b} + 1\right). \tag{4.35}$$

The potential is also plotted in Figure 4.16.

The Poisson equation can be solved by numerical methods developed in Section 4.2. If the finite difference approximation to $\nabla^2\varphi$ [Eq. (4.14)] is substituted in the Poisson equation in Cartesian coordinates (4.32) and both sides are multiplied by Δ^2, the following equation results:

$$\begin{aligned} &-6\Phi(i,j,k) + \Phi(i+1,j,k) + \Phi(i-1,j,k) + \Phi(i,j+1,k) \\ &+ \Phi(i,j-1,k) + \Phi(i,j,k+1) + \Phi(i,j,k-1) = -\rho(x,y,z)\Delta^3/\Delta\varepsilon_o. \end{aligned} \tag{4.36}$$

The factor $\rho\Delta^3$ is approximately the total charge in a volume Δ^3 surrounding the mesh point *(i, j,k)* when (1) the charge density is a smooth function of position and (2) the distance Δ is small compared to the scale length for variations in ρ. Equation (4.36) can be converted to a finite difference equation by defining $Q(i, j, k) = \rho(x, y, z)\Delta^3$. Equation (4.36) becomes

$$\Phi(i,j,k) = 1/6\ [\Phi(i+1,j,k) + \Phi(i-1,j,k) + \Phi(i,j+1,k) + \Phi(i,j-1,k) + \Phi(i,j,k+1) + \Phi(i,j,k-1)] + Q(i,j,k)/6\varepsilon_o. \tag{4.37}$$

Equation (4.37) states that the potential at a point is the average of 'its nearest neighbors elevated (or lowered) by a term proportional to the space charge surrounding the point.

The method of successive relaxation can easily be modified to treat problems with space charge. In this case, the residual [Eq. (4.16)] for a two-dimensional problem is

$$R(i,j,) = 1/4\ [\Phi(i+1,j) + \Phi(i-1,j) + \Phi(i,j+1) + \Phi(i,j-1)] - \Phi(i,j) + Q(i,j)/4\varepsilon_o\ . \tag{4.38}$$

4.6 MAGNETIC FIELDS IN SIMPLE GEOMETRIES

This section illustrates some methods to find static magnetic fields by direct use of the Maxwell equations [(4.3) and (4.4)]. The fields are produced by current-carrying wires. Two simple, but often encountered, geometries are included: the field outside a long straight wire and the field inside of solenoidal winding of infinite extent.

The wire (Fig. 4.17) has current I in the z direction. There are no radial magnetic field lines since $\nabla\cdot\mathbf{B} = 0$. There is no component B_z since the fields must be perpendicular to the current. Thus, magnetic field lines are azimuthal. By symmetry, the field lines are circles. The magnitude of the azimuthal field (or density of lines) can be determined by rewriting the static form of Eq. (3.12) in integral form according to the Stokes law [Eq. (4.7)],

$$\int \boldsymbol{B}\cdot\boldsymbol{dl} = \mu_o \iint j_z dA = \mu_o I. \tag{4.39}$$

Using the fact that field lines are circles, we find that

$$B_\theta = \mu_o I/2\pi r. \tag{4.40}$$

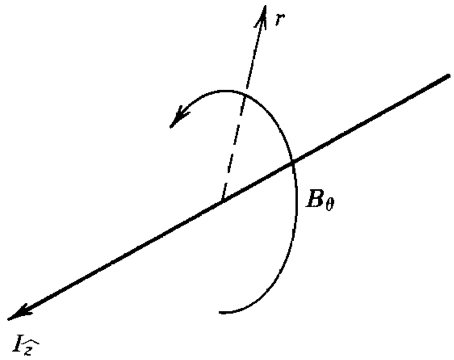


Figure 4.17 Magnetic field lines near a current-carrying wire.

The solenoidal coil is illustrated in Figure 4.18. It consists of a helical winding of insulated wire on a cylindrical mandrel. The wire carries current I. The quantity (N/L) is the number of turns per unit length. *Solenoid* derives from the Greek word for pipe; magnetic field lines are channeled through the windings. In a finite length winding, the field lines return around the outside. We will consider the case of an infinitely long structure with no axial variations. Furthermore, we assume there are many windings over a length comparable to the coil radius, or $(N/L)r_c \gg 1$. In this limit, we can replace the individual windings with a uniform azimuthal current sheet. The sheet has a current per unit length J $(A/m) = (N/L)I$.

The current that produces the field is azimuthal. By the law of Biot and Savart, there can be no component of azimuthal magnetic field. By symmetry, there can be no axial variation of field. The conditions of zero divergence and curl of the magnetic field inside the winding are written

$$\frac{1}{r}\frac{\partial(rB_r)}{\partial r} + \frac{\partial B_z}{\partial z} = 0, \qquad \frac{\partial B_r}{\partial z} - \frac{\partial B_z}{\partial r} = 0. \tag{4.41}$$

Setting $\partial/\partial z$ equal to zero in Eqs. (4.41); we find that B_r is zero and that B_z has equal magnitude at all radii. The magnitude of the axial field can be determined by applying Eq. (4.39) to the loop illustrated in Figure 4.18. The field outside a long solenoid is negligible since return magnetic flux is spread over a large area. There are no contributions to the loop integral from the radial segments because fields are axial. The only component of the integral comes from the part of the path inside the solenoid, so that

$$B_o = \mu_o J = \mu_o (N/L) I. \tag{4.42}$$

Many magnetic confinement systems for intense electron beams or for high-temperature plasmas are based on a solenoidal coil bent in a circle and connected, as shown in Figure 4.19. The geometry is that of a doughnut or *torus* with circular cross section. The axial fields that circulate

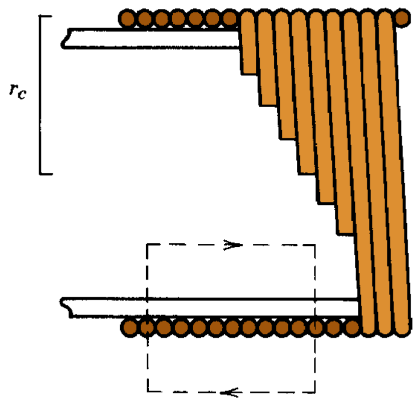


Figure 4.18 Solenoidal magnet coil.

around the torus are called *toroidal field lines*. Field lines are continuous and self-connected. All field lines are contained within the winding. The toroidal field magnitude inside the winding is not uniform. Modification of the loop construction of Figure 4.19 shows that the field varies as the inverse of the major radius. Toroidal field variation is small when the minor radius (the radius of the solenoidal windings) is much less than the major radius.

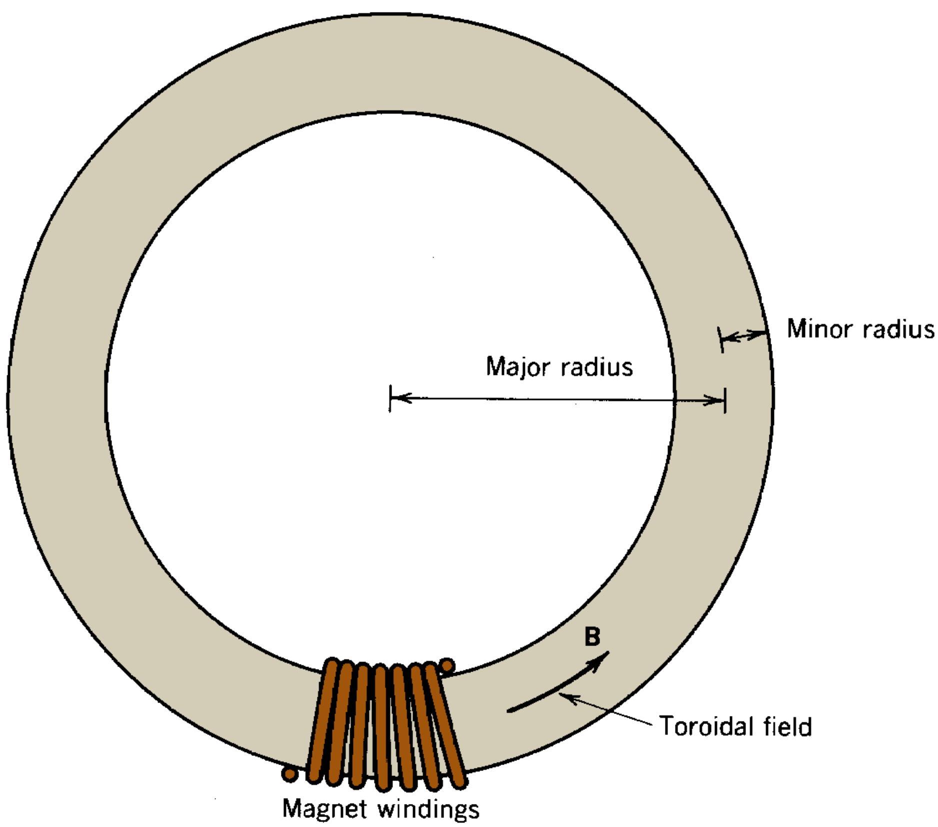


Figure 4.19 Toroidal magnet coil.

4.7 MAGNETIC POTENTIALS

The magnetic potential and the vector potential aid in the calculation of magnetic fields. In this section, we will consider how these functions are related and investigate the physical meaning of the vector potential in a two-dimensional geometry. The vector potential will be used to derive the magnetic field for a circular current loop. Assemblies of loop currents are used to generate magnetic fields in many particle beam transport devices.

In certain geometries, magnetic field lines and the vector potential are closely related. Figure 4.20 illustrates lines of constant vector potential in an axially uniform system in which fields are generated by currents in the *z* direction. Equation (3.24) implies that the vector potential has only an axial component, A_z. Equation (3.23) implies that

$$B_x = \partial A_z/\partial y, \quad B_y = -\partial A_z/\partial x. \tag{4.43}$$

Figure 4.20 shows a surface of constant A_z in the geometry considered. This line is defined by

$$dA_z = 0 = (\partial A_z/\partial x)\ dx + (\partial A_z/\partial y)\ dy. \tag{4.44}$$

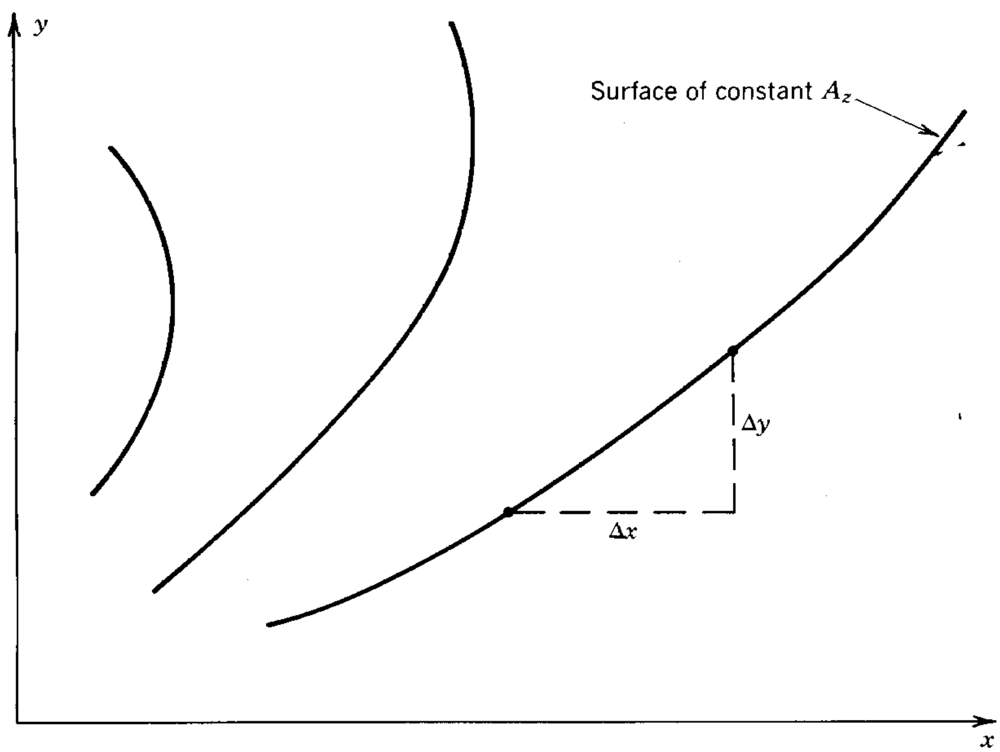


Figure 4.20 Lines of constant vector potential A_z. Two-dimensional system, symmetric along the *z* direction (out of page), axial currents only (j_z).

Substituting Eqs. (4.43) into Eq. (4.44), an alternate equation for a constant A_z line is

$$dy/dx = B_y/B_x. \tag{4.45}$$

Equation (4.45) is also the equation for a magnetic field line. To summarize, when magnetic fields are generated by axial currents uniform in z, magnetic field lines are defined by lines of constant A_z.

A similar construction shows that magnetic field lines are normal to surfaces of constant magnetic potential. In the geometry of Figure 4.20,

$$B_x = \partial U_m/\partial x, \quad B_y = \partial U_m/\partial y. \tag{4.46}$$

by the definition of U_m. The equation for a line of constant U_m is

$$dU_m = (\partial U_m/\partial x)\ dx + (\partial U_m/\partial y)\ dy = B_x dx + B_y dy.$$

Lines of constant magnetic potential are described by the equation

$$dy/dx = -B_y/B_x. \tag{4.47}$$

Analytic geometry shows that the line described by Eq. (4.47) is perpendicular to that of Eq. (4.45).

The correspondence of field lines and lines of constant A_z can be used to find magnetic fields of arrays of currents. As an example, consider the geometry illustrated in Figure 4.21. Two infinite length wires carrying opposed currents $\pm I$ are separated by a distance $2d$. It is not difficult to show that the vector potential for a single wire is

$$A_z = \pm \tfrac{1}{2}\mu_o I\ \ln(x'^2 + y'^2)/2\pi.$$

where the origin of the coordinate system (x', y') is centered on the wire. The total vector potential is the sum of contributions from both wires. In terms of the coordinate system (x, y) defined in Figure 4.21, the total vector potential is

$$A_z = \frac{\mu_o I}{4\pi}\ \ln\left(\frac{(x-d)^2+y^2}{(x+d)^2+y^2}\right).$$

Lines of constant A_z (corresponding to magnetic field tines) are plotted in Figure 4.21.

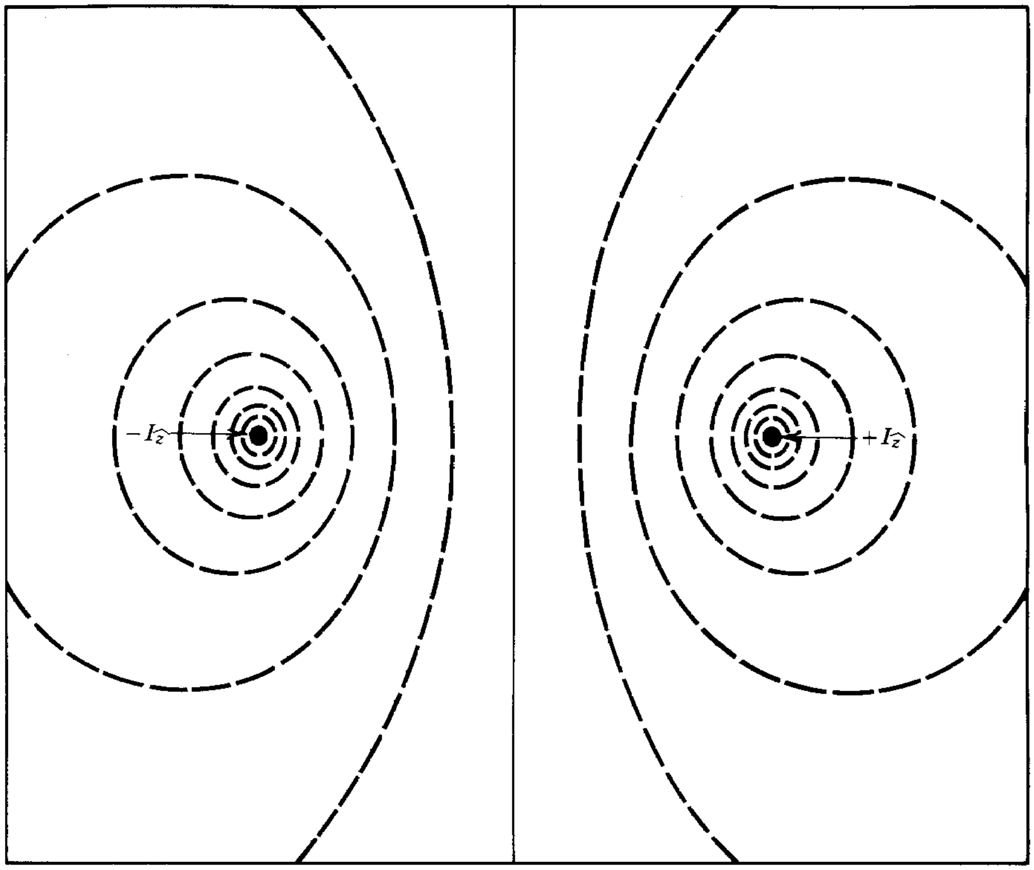


Figure 4.21 Lines of constant vector potential A_z from two wires of infinite extent in z (out of page) carrying equal and opposite currents.

There are many instances in accelerator applications in which magnetic fields are produced by azimuthal currents in cylindrically symmetric systems. For instance, the field of a solenoidal lens (Section 6.7) is generated by axicentered current loops of various radii. There is only one nonzero component of the vector potential, A_θ. It can be shown that magnetic field lines follow surfaces of constant $2\pi rA_\theta$. The function $2\pi rA_\theta$ is called the *stream function*. The contribution from many loops can be summed to find a net stream function.

The vector potential of a current loop of radius a (Fig. 4.22) can be found by application of Eq. (3.24). In terms of cylindrical coordinates centered at the loop axis, the current density is

$$j_\theta = I\ \delta(z') \ \delta(r'-a). \tag{4.48}$$

Care must be exercised in evaluating the integrals, since Eq. (3.24) holds only for a Cartesian coordinate system. The result is

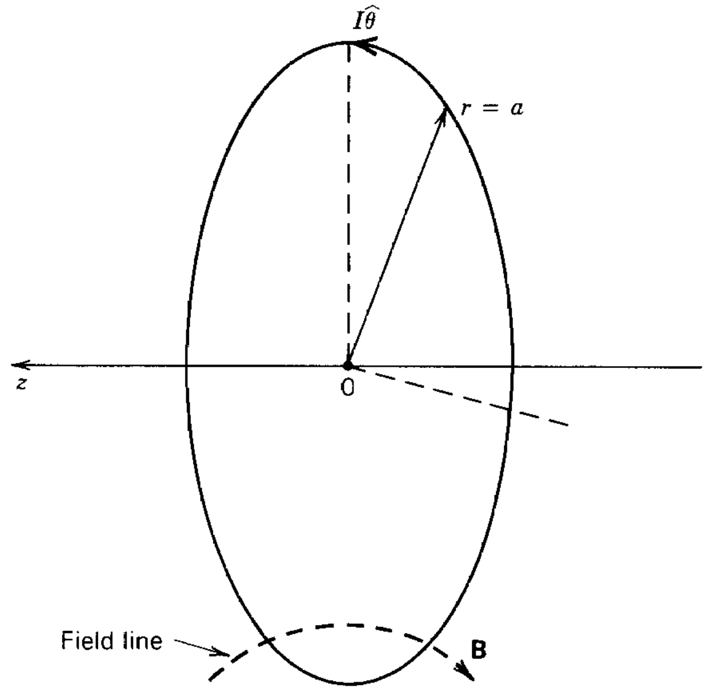


Figure 4.22 Quantities for calculating magnetic fields from a circular current loop.

$$A_\theta = \left(\frac{\mu_o I a}{4\pi} \right) \int_0^{2\pi} \frac{\cos\theta' d\theta'}{(a^2 + r^2 + z^2 - 2ar \cos\theta')^{1/2}} . \tag{4.49}$$

Defining the quantity

$$M = 4ar/(a^2 + r^2 + z^2 + 2ar),$$

Eq. (4.49) can be written in terms of the complete elliptic integrals E(M) and K(M) as

$$A_\theta = \frac{\mu_o I a}{\pi (a^2 + r^2 + z^2 + 2ar)^{1/2}} \frac{(2-M)\ K(M) - 2\ E(M)}{M} . \tag{4.50}$$

Although the expressions in Eq. (4.50) are relatively complex, the vector potential can be calculated quickly on a computer. Evaluating the elliptic integrals directly is usually ineffective and time consuming. A better approach is to utilize empirical series tabulated in many mathematical handbook s. These series give an accurate approximation in terms of power series and elementary transcendental functions. For example, the elliptic integrals are given to an accuracy of 4 x 10^{-5} by [adapted from M. Abramowitz and I. A. Stegun, Eds., **Handbook of Mathematical Functions** (Dover, New York, 1970), p. 591].

$$K(M) = 1.38629 + 0.111972(1-M) + 0.0725296(1-M)^2$$
$$+ [0.50000 + 0.121348(1-M) + 0.0288729(1-M)^2)]\ \ln(1/(1-M)l \tag{4.51}$$

$$E(M) = 1+0.463015(1-M) + 0.107781(1-M)^2$$
$$+ [0.245273(1-M) + 0.0412496(1-M)^2]\ \ln[1/(1-M)]. \tag{4.52}$$

(4.52)

The vector potential can be calculated for multiple coils bv transforming coordinates and then summing A_θ. The transformations are $z \rightarrow (z - z_{cn})$ and $a \rightarrow r_{cn}$, where z_{cn} and r_{cn} are the coordinates of the nth coil. Given the net vector potential, the magnetic fields are

$$B_z = \frac{1}{r}\ \frac{\partial (rA_\theta)}{\partial r}, \quad B_r = -\frac{\partial A_\theta}{\partial z} . \tag{4.53}$$

A quantity of particular interest for paraxial orbit calculations (Section 7.5) is the longitudinal field magnitude on the axis $B_z(0, z)$. The vector potential for a single coil [Eq. (4.49)] can be expanded for r « a as

$$A_\theta(r,z) = \left(\frac{\mu_o I a}{4\pi}\right) \int_0^{2\pi} \left(\frac{\cos\theta'\ d\theta'}{\sqrt{a^2+z^2}} + \frac{ar\cos^2\theta'\ d\theta'}{\sqrt{a^2+z^2}^{\,3}}\right) . \tag{4.54}$$

The integral of the first term is zero, while the second term gives

$$A_\theta = \frac{\mu_o I a^2 r}{4\ \sqrt{a^2+z^2}^{\,3}} . \tag{4.55}$$

Applying Eq. (4.53), the axial field is

$$B_z(0,z) = \frac{\mu_o I a^2}{2\sqrt{a^2+z^2}^{\,3}} . \tag{4.56}$$

We can use Eq. (4.56) to derive the geometry of the Helmholtz coil configuration. Assume that two loops with equal current are separated by an axial distance *d*. A Taylor expansion of the axial field near the axis about the midpoint of the coils gives

$$B_z(0,z) = (B_1+B_2) + (\partial B_1/\partial z + \partial B_2/\partial z)\ z$$

$$+ (\partial^2 B_1/\partial z^2 + \partial^2 B_2/\partial z^2)\ z^2 + \ldots$$

The subscript *1* refers to the contribution from the coil at $z = -d/2$, while *2* is associated with the coil at $z = +d/2$. The derivatives can be determined from Eq. (4.56). The zero-order components from both coils add. The first derivatives cancel at all values of the coil spacing. At a spacing of $d = a$, the second derivatives also cancel. Thus, field variations near the symmetry point are only on the order of $(z/a)^3$. Two coils with $d = a$ are called Helmholtz coils. They are used when a weak but accurate axial field is required over a region that is small compared to the dimension of the coil. The field magnitude for Helmholtz coils is

$$B_z = \mu_o I \ /\ (1.25)^{3/2}\ a \qquad (4.57)$$