ELECTROMAGNETICS

SECOND EDITION

Edward J. Rothwell
Michael J. Cloud

CRC Press
Taylor & Francis Group

ELECTROMAGNETICS

SECOND EDITION

ELECTROMAGNETICS

SECOND EDITION

Edward J. Rothwell
Michael J. Cloud

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-4200-6448-3 (Ebook-PDF)

Contents

Preface

This is the second edition of our book *Electromagnetics*. It is intended as a text for use by engineering graduate students in a first-year sequence where the basic concepts learned as undergraduates are solidified and a conceptual foundation is established for future work in research. It should also prove useful for practicing engineers who wish to improve their understanding of the principles of electromagnetics, or brush up on those fundamentals that have become cloudy with the passage of time.

Much of the material from the first edition remains intact. The most significant change is the addition of an extensive chapter on integral equations with application to electromagnetics. Many new problems of varying difficulty have also been added, and we have prepared a complete solution manual for all of the end-of-chapter exercises.

The assumed background of the reader is limited to standard undergraduate topics in physics and mathematics. Worthy of explicit mention are complex arithmetic, vector analysis, ordinary differential equations, and certain topics normally covered in a "signals and systems" course (e.g., convolution and the Fourier transform). Further analytical tools, such as contour integration, dyadic analysis, and separation of variables, are covered in a self-contained mathematical appendix.

The organization of the book is in seven chapters. In Chapter 1 we present essential background on the field concept, as well as information related specifically to the electromagnetic field and its sources. Chapter 2 is concerned with a presentation of Maxwell's theory of electromagnetism. Here attention is given to several useful forms of Maxwell's equations, the nature of the four field quantities and of the postulate in general, some fundamental theorems, and the wave nature of the time-varying field. The electrostatic and magnetostatic cases are treated in Chapter 3. In Chapter 4 we cover the representation of the field in the frequency domains: both temporal and spatial. Here the behavior of common engineering materials is also given some attention. The use of potential functions is discussed in Chapter 5, along with other field decompositions including the solenoidal–lamellar, transverse–longitudinal, and TE–TM types. In Chapter 6 we present the powerful integral solution to Maxwell's equations by the method of Stratton and Chu. Finally, in Chapter 7 we provide introductory coverage of integral equations and discuss how they may be used to solve a variety of problems in electromagnetics, including several classic problems in radiation and scattering. A main mathematical appendix near the end of the book contains brief but sufficient treatments of Fourier analysis, vector transport theorems, complex-plane integration, dyadic analysis, and boundary value problems. Several subsidiary appendices provide useful tables of identities, transforms, and so on.

We would like to express our deep gratitude to those persons who contributed to the development of the book. The reciprocity-based derivation of the Stratton–Chu formula was provided by Prof. Dennis Nyquist, as was the material on wave reflection from multiple layers. The groundwork for our discussion of the Kronig–Kramers relations was provided by Prof. Michael Havrilla, and material on the time-domain reflection coefficient was developed by Dr. Jungwook Suk. We owe thanks to Prof. Leo Kempel, Dr. David Infante, and Dr. Ahmet Kizilay for carefully reading large portions of the manuscript

for the first edition during its preparation, and to Dr. Christopher Coleman for helping to prepare the figures. We are indebted to Dr. John E. Ross for kindly permitting us to employ one of his computer programs for scattering from a sphere and another for numerical Fourier transformation. Helpful comments and suggestions on the figures were provided by Beth Lannon–Cloud. Technical assistance for the layout of the second edition was graciously provided by Shashi Kumar of International Typesetting and Composition. Finally, we would like to thank Michael Davidson, Jessica Vakili, and Nora Konopka of Taylor & Francis for their guidance and support throughout the publication process, and James Miller for designing the cover.

Authors

Edward J. Rothwell received the BS degree from Michigan Technological University, the MS and EE degrees from Stanford University, and the PhD from Michigan State University, all in electrical engineering. He has been a faculty member in the Department of Electrical and Computer Engineering at Michigan State University since 1985, and currently holds the rank of Professor. Before coming to Michigan State he worked at Raytheon and Lincoln Laboratory. Dr. Rothwell has published numerous articles in professional journals involving his research in electromagnetics and related subjects. He is a member of Phi Kappa Phi, Sigma Xi, URSI Commission B, and is a Fellow of the IEEE.

Michael J. Cloud received the BS, MS, and PhD degrees from Michigan State University, all in electrical engineering. He has been a faculty member in the Department of Electrical and Computer Engineering at Lawrence Technological University since 1987, and currently holds the rank of Associate Professor. Dr. Cloud has coauthored six other books: Computational Functional Analysis (Horwood Ltd, 2007), Applications of Functional Analysis and Operator Theory (Elsevier, 2005), Approximating Perfection (Princeton University Press, 2004), The Calculus of Variations and Functional Analysis (World Scientific, 2003), Tensor Analysis (World Scientific, 2003), and Inequalities with Applications to Engineering (Springer-Verlag, 1998; Morikita Shuppan, 2004). He is a senior member of the IEEE.

1

Introductory concepts

1.1 Notation, conventions, and symbology

Any book that covers a broad range of topics will likely harbor some problems with notation and symbology. This results from having the same symbol used in different areas to represent different quantities, and also from having too many quantities to represent. Rather than invent new symbols, we choose to stay close to the standards and warn the reader about any symbol used to represent more than one distinct quantity.

The basic nature of a physical quantity is indicated by typeface or by the use of a diacritical mark. Scalars are shown in ordinary typeface: q, Φ, for example. Vectors are shown in boldface: $\mathbf{E}, \mathbf{\Pi}$. Dyadics are shown in boldface with an overbar: $\bar{\boldsymbol{\epsilon}}, \bar{\mathbf{A}}$. Frequency dependent quantities are indicated by a tilde, whereas time dependent quantities are written without additional indication; thus we write $\tilde{\mathbf{E}}(\mathbf{r}, \omega)$ and $\mathbf{E}(\mathbf{r}, t)$. (Some quantities, such as impedance, are used in the frequency domain to interrelate Fourier spectra; although these quantities are frequency dependent they are seldom written in the time domain, and hence we do not attach tildes to their symbols.) We often combine diacritical marks: for example, $\bar{\tilde{\boldsymbol{\epsilon}}}$ denotes a frequency domain dyadic. We distinguish carefully between phasor and frequency domain quantities. The variable ω is used for the frequency variable of the Fourier spectrum, while $\check{\omega}$ is used to indicate the constant frequency of a time harmonic signal. We thus further separate the notion of a phasor field from a frequency domain field by using a *check* to indicate a phasor field: $\check{\mathbf{E}}(\mathbf{r})$. However, there is often a simple relationship between the two, such as $\check{\mathbf{E}} = \tilde{\mathbf{E}}(\check{\omega})$.

We designate the field and source point position vectors by $\mathbf{r}$ and $\mathbf{r}'$, respectively, and the corresponding relative displacement or distance vector by $\mathbf{R}$:

$$\mathbf{R} = \mathbf{r} - \mathbf{r}'.$$

A hat designates a vector as a unit vector (e.g., $\hat{\mathbf{x}}$). The sets of coordinate variables in rectangular, cylindrical, and spherical coordinates are denoted by

$$(x, y, z), \qquad (\rho, \phi, z), \qquad (r, \theta, \phi),$$

respectively. (In the spherical system ϕ is the azimuthal angle and θ is the polar angle.) We freely use the "del" operator notation ∇ for gradient, curl, divergence, Laplacian, and so on.

The SI (MKS) system of units is employed throughout the book.

1.2 The field concept of electromagnetics

Introductory treatments of electromagnetics often stress the role of the field in force transmission: the individual fields **E** and **B** are defined via the mechanical force on a small test charge. This is certainly acceptable, but does not tell the whole story. We might, for example, be left with the impression that the EM field always arises from an interaction between charged objects. Often coupled with this is the notion that the field concept is meant merely as an aid to the calculation of force, a kind of notational convenience not placed on the same physical footing as force itself. In fact, fields are more than useful — they are fundamental. Before discussing electromagnetic fields in more detail, let us attempt to gain a better perspective on the field concept and its role in modern physical theory. Fields play a central role in any attempt to describe physical reality. They are as real as the physical substances we ascribe to everyday experience. In the words of Einstein [71],

> "It seems impossible to give an obvious qualitative criterion for distinguishing between matter and field or charge and field."

We must therefore put fields and particles of matter on the same footing: both carry energy and momentum, and both interact with the observable world.

1.2.1 Historical perspective

Early nineteenth century physical thought was dominated by the *action at a distance* concept, formulated by Newton more than 100 years earlier in his immensely successful theory of gravitation. In this view the influence of individual bodies extends across space, instantaneously affects other bodies, and remains completely unaffected by the presence of an intervening medium. Such an idea was revolutionary; until then *action by contact*, in which objects are thought to affect each other through physical contact or by contact with the intervening medium, seemed the obvious and only means for mechanical interaction. Priestly's experiments in 1766 and Coulomb's torsion-bar experiments in 1785 seemed to indicate that the force between two electrically charged objects behaves in strict analogy with gravitation: both forces obey inverse square laws and act along a line joining the objects. Oersted, Ampere, Biot, and Savart soon showed that the magnetic force on segments of current-carrying wires also obeys an inverse square law.

The experiments of Faraday in the 1830s placed doubt on whether action at a distance really describes electric and magnetic phenomena. When a material (such as a dielectric) is placed between two charged objects, the force of interaction decreases; thus, the intervening medium does play a role in conveying the force from one object to the other. To explain this, Faraday visualized "lines of force" extending from one charged object to another. The manner in which these lines were thought to interact with materials they intercepted along their path was crucial in understanding the forces on the objects. This also held for magnetic effects. Of particular importance was the number of lines passing through a certain area (the *flux*), which was thought to determine the amplitude of the effect observed in Faraday's experiments on electromagnetic induction.

Faraday's ideas presented a new world view: electromagnetic phenomena occur in the region surrounding charged bodies, and can be described in terms of the laws governing the "field" of his lines of force. Analogies were made to the stresses and strains in material objects, and it appeared that Faraday's force lines created equivalent electromagnetic

stresses and strains in media surrounding charged objects. His law of induction was formulated not in terms of positions of bodies, but in terms of lines of magnetic force. Inspired by Faraday's ideas, Gauss restated Coulomb's law in terms of flux lines, and Maxwell extended the idea to time changing fields through his concept of displacement current.

In the 1860s Maxwell created what Einstein called "the most important invention since Newton's time" — a set of equations describing an entirely field-based theory of electromagnetism. These equations do not model the forces acting between bodies, as do Newton's law of gravitation and Coulomb's law, but rather describe only the dynamic, time-evolving structure of the electromagnetic field. Thus bodies are not seen to interact with each other, but rather with the (very real) electromagnetic field they create, an interaction described by a supplementary equation (the Lorentz force law). To better understand the interactions in terms of mechanical concepts, Maxwell also assigned properties of stress and energy to the field.

Using constructs that we now call the electric and magnetic fields and potentials, Maxwell synthesized all known electromagnetic laws and presented them as a system of differential and algebraic equations. By the end of the nineteenth century, Hertz had devised equations involving only the electric and magnetic fields, and had derived the laws of circuit theory (Ohm's law and Kirchhoff's laws) from the field expressions. His experiments with high-frequency fields verified Maxwell's predictions of the existence of electromagnetic waves propagating at finite velocity, and helped solidify the link between electromagnetism and optics. But one problem remained: if the electromagnetic fields propagated by stresses and strains on a medium, how could they propagate through a vacuum? A substance called the *luminiferous aether*, long thought to support the transverse waves of light, was put to the task of carrying the vibrations of the electromagnetic field as well. However, the pivotal experiments of Michelson and Morely showed that the aether was fictitious, and the physical existence of the field was firmly established.

The essence of the field concept can be conveyed through a simple thought experiment. Consider two stationary charged particles in free space. Since the charges are stationary, we know that (1) another force is present to balance the Coulomb force between the charges, and (2) the momentum and kinetic energy of the system are zero. Now suppose one charge is quickly moved and returned to rest at its original position. Action at a distance would require the second charge to react immediately (Newton's third law), but by Hertz's experiments it does not. There appears to be no change in energy of the system: both particles are again at rest in their original positions. However, after a time (given by the distance between the charges divided by the speed of light) we find that the second charge does experience a change in electrical force and begins to move away from its state of equilibrium. But by doing so it has gained net kinetic energy and momentum, and the energy and momentum of the system seem larger than at the start. This can only be reconciled through field theory. If we regard the field as a physical entity, then the nonzero work required to initiate the motion of the first charge and return it to its initial state can be seen as increasing the energy of the field. A disturbance propagates at finite speed and, upon reaching the second charge, transfers energy into kinetic energy of the charge. Upon its acceleration this charge also sends out a wave of field disturbance, carrying energy with it, eventually reaching the first charge and creating a second reaction. At any given time, the net energy and momentum of the system, composed of both the bodies and the field, remain constant. We thus come to regard the electromagnetic field as a true physical entity: an entity capable of carrying energy and momentum.

1.2.2 Formalization of field theory

Before we can invoke physical laws, we must find a way to describe the *state* of the system we intend to study. We generally begin by identifying a set of *state variables* that can depict the physical nature of the system. In a mechanical theory such as Newton's law of gravitation, the state of a system of point masses is expressed in terms of the instantaneous positions and momenta of the individual particles. Hence $6N$ state variables are needed to describe the state of a system of N particles, each particle having three position coordinates and three momentum components. The time evolution of the system state is determined by a supplementary force function (e.g., gravitational attraction), the initial state (initial conditions), and Newton's second law $\mathbf{F} = d\mathbf{P}/dt$.

Descriptions using finite sets of state variables are appropriate for action-at-a-distance interpretations of physical laws such as Newton's law of gravitation or the interaction of charged particles. If Coulomb's law were taken as the force law in a mechanical description of electromagnetics, the state of a system of particles could be described completely in terms of their positions, momenta, and charges. Of course, charged particle interaction is not this simple. An attempt to augment Coulomb's force law with Ampere's force law would not account for kinetic energy loss via radiation. Hence we abandon* the mechanical viewpoint in favor of the field viewpoint, selecting a different set of state variables. The essence of field theory is to regard electromagnetic phenomena as affecting all of space. We shall find that we can describe the field in terms of the four vector quantities **E**, **D**, **B**, and **H**. Because these fields exist by definition at each point in space and each time t, a finite set of state variables cannot describe the system.

Here then is an important distinction between field theories and mechanical theories: the state of a field at any instant can only be described by an infinite number of state variables. Mathematically we describe fields in terms of functions of continuous variables; however, we must be careful not to confuse all quantities described as "fields" with those fields innate to a scientific field theory. For instance, we may refer to a temperature "field" in the sense that we can describe temperature as a function of space and time. However, we do *not* mean by this that temperature obeys a set of physical laws analogous to those obeyed by the electromagnetic field.

What special character, then, can we ascribe to the electromagnetic field that has meaning beyond that given by its mathematical implications? In this book, **E**, **D**, **B**, and **H** are integral parts of a *field-theory description* of electromagnetics. In any field theory we need two types of fields: a *mediating field* generated by a source, and a field describing the source itself. In free-space electromagnetics the mediating field consists of **E** and **B**, while the source field is the distribution of charge or current. An important consideration is that the source field must be independent of the mediating field that it "sources." Additionally, fields are generally regarded as unobservable: they can only be measured indirectly through interactions with observable quantities. We need a link to mechanics to observe **E** and **B**: we might measure the change in kinetic energy of a particle as it interacts with the field through the Lorentz force. The Lorentz force becomes the force function in the mechanical interaction that uniquely determines the (observable) mechanical state of the particle.

A field is associated with a set of *field equations* and a set of *constitutive relations*. The field equations describe, through partial derivative operations, both the spatial distribution and temporal evolution of the field. The constitutive relations describe the effect

*Attempts have been made to formulate electromagnetic theory purely in action-at-a-distance terms, but this viewpoint has not been generally adopted [77].

of the supporting medium on the fields and are dependent upon the physical state of the medium. The state may include macroscopic effects, such as mechanical stress and thermodynamic temperature, as well as the microscopic, quantum-mechanical properties of matter.

The value of the field at any position and time in a bounded region V is then determined uniquely by specifying the sources within V, the initial state of the fields within V, and the value of the field or finitely many of its derivatives on the surface bounding V. If the boundary surface also defines a surface of discontinuity between adjacent regions of differing physical characteristics, or across discontinuous sources, then *jump conditions* may be used to relate the fields on either side of the surface.

The variety of forms of field equations is restricted by many physical principles including reference-frame invariance, conservation, causality, symmetry, and simplicity. Causality prevents the field at time $t = 0$ from being influenced by events occurring at subsequent times $t > 0$. Of course, we prefer that a field equation be mathematically robust and well-posed to permit solutions that are unique and stable.

Many of these ideas are well illustrated by a consideration of electrostatics. We can describe the electrostatic field through a mediating scalar field $\Phi(x, y, z)$ known as the electrostatic potential. The spatial distribution of the field is governed by Poisson's equation

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = -\frac{\rho}{\epsilon_0},$$

where $\rho = \rho(x, y, z)$ is the source charge density. No temporal derivatives appear, and the spatial derivatives determine the spatial behavior of the field. The function ρ represents the spatially-averaged distribution of charge that acts as the source term for the field Φ. Note that ρ incorporates no information about Φ. To uniquely specify the field at any point, we must still specify its behavior over a boundary surface. We could, for instance, specify Φ on five of the six faces of a cube and the normal derivative $\partial\Phi/\partial n$ on the remaining face. Finally, we cannot directly observe the static potential field, but we can observe its interaction with a particle. We relate the static potential field theory to the realm of mechanics via the electrostatic force $\mathbf{F} = q\mathbf{E}$ acting on a particle of charge q.

In future chapters we shall present a classical field theory for macroscopic electromagnetics. In that case the mediating field quantities are $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{H}$, and the source field is the current density $\mathbf{J}$.

1.3 The sources of the electromagnetic field

Electric charge is an intriguing natural entity. Human awareness of charge and its effects dates back to at least 600 BC, when the Greek philosopher Thales of Miletus observed that rubbing a piece of amber could enable the amber to attract bits of straw. Although *charging by friction* is probably still the most common and familiar manifestation of electric charge, systematic experimentation has revealed much more about the behavior of charge and its role in the physical universe. There are two kinds of charge, to which Benjamin Franklin assigned the respective names *positive* and *negative*. Franklin observed that charges of opposite kind attract and charges of the same kind repel. He also found that an increase in one kind of charge is accompanied by an increase in the other,

and so first described the principle of *charge conservation*. Twentieth century physics has added dramatically to the understanding of charge:

1. Electric charge is a fundamental property of matter, as is mass or dimension.
2. Charge is *quantized*: there exists a smallest quantity (*quantum*) of charge that can be associated with matter. No smaller amount has been observed, and larger amounts always occur in integral multiples of this quantity.
3. The charge quantum is associated with the smallest subatomic particles, and these particles interact through electrical forces. In fact, matter is organized and arranged through electrical interactions; for example, our perception of physical contact is merely the macroscopic manifestation of countless charges in our fingertips pushing against charges in the things we touch.
4. Electric charge is an *invariant*: the value of charge on a particle does not depend on the speed of the particle. In contrast, the mass of a particle increases with speed.
5. Charge acts as the source of an electromagnetic field; the field is an entity that can carry energy and momentum away from the charge via propagating waves.

We begin our investigation of the properties of the electromagnetic field with a detailed examination of its source.

1.3.1 Macroscopic electromagnetics

We are interested primarily in those electromagnetic effects that can be predicted by classical techniques using continuous sources (charge and current densities). Although macroscopic electromagnetics is limited in scope, it is useful in many situations encountered by engineers. These include, for example, the determination of currents and voltages in lumped circuits, torques exerted by electrical machines, and fields radiated by antennas. Macroscopic predictions can fall short in cases where quantum effects are important: e.g., with devices such as tunnel diodes. Even so, quantum mechanics can often be coupled with classical electromagnetics to determine the macroscopic electromagnetic properties of important materials.

Electric charge is not of a continuous nature. The quantization of atomic charge — $\pm e$ for electrons and protons, $\pm e/3$ and $\pm 2e/3$ for quarks — is one of the most precisely established principles in physics (verified to 1 part in 10^{21}). The value of e itself is known to great accuracy:

$$e = 1.60217733 \times 10^{-19} \text{ Coulombs (C)}.$$

However, the discrete nature of charge is not easily incorporated into everyday engineering concerns. The strange world of the individual charge — characterized by particle spin, molecular moments, and thermal vibrations — is well described only by quantum theory. There is little hope that we can learn to describe electrical machines using such concepts. Must we therefore retreat to the macroscopic idea and ignore the discretization of charge completely? A viable alternative is to use atomic theories of matter to estimate the useful scope of macroscopic electromagnetics.

Remember, we are completely free to postulate a theory of nature whose scope may be limited. Like continuum mechanics, which treats distributions of matter as if they were continuous, macroscopic electromagnetics is regarded as valid because it is verified by experiment over a certain range of conditions. This applicability range generally corresponds to dimensions on a laboratory scale, implying a very wide range of validity for engineers.

Macroscopic effects as averaged microscopic effects. Macroscopic electromagnetics can hold in a world of discrete charges because applications usually occur over physical scales that include vast numbers of charges. Common devices, generally much larger than individual particles, "average" the rapidly varying fields that exist in the spaces between charges, and this allows us to view a source as a continuous "smear" of charge. To determine the range of scales over which the macroscopic viewpoint is valid, we must compare averaged values of microscopic fields to the macroscopic fields we measure in the lab. But if the effects of the individual charges are describable only in terms of quantum notions, this task will be daunting at best. A simple compromise, which produces useful results, is to extend the macroscopic theory right down to the microscopic level and regard discrete charges as "point" entities that produce electromagnetic fields according to Maxwell's equations. Then, in terms of scales much larger than the classical radius of an electron ($\approx 10^{-14}$ m), the expected rapid fluctuations of the fields in the spaces between charges is predicted. Finally, we ask: over what spatial scale must we average the effects of the fields and the sources in order to obtain agreement with the macroscopic equations?

In the spatial averaging approach a convenient weighting function $f(\mathbf{r})$ is chosen, and is normalized so that $\int f(\mathbf{r})\, dV = 1$. An example is the Gaussian distribution

$$f(\mathbf{r}) = (\pi a^2)^{-3/2} e^{-r^2/a^2},$$

where a is the approximate radial extent of averaging. The spatial average of a microscopic quantity $F(\mathbf{r}, t)$ is given by

$$\langle F(\mathbf{r}, t) \rangle = \int F(\mathbf{r} - \mathbf{r}', t) f(\mathbf{r}')\, dV'. \tag{1.1}$$

The scale of validity of the macroscopic model can be found by determining the averaging radius a that produces good agreement between the averaged microscopic fields and the macroscopic fields.

The macroscopic volume charge density. At this point we do not distinguish between the "free" charge that is unattached to a molecular structure and the charge found near the surface of a conductor. Nor do we consider the dipole nature of polarizable materials or the microscopic motion associated with molecular magnetic moment or the magnetic moment of free charge. For the consideration of free-space electromagnetics, we assume charge exhibits either three degrees of freedom (*volume charge*), two degrees of freedom (*surface charge*), or one degree of freedom (*line charge*).

In typical matter, the microscopic fields vary spatially over dimensions of 10^{-10} m or less, and temporally over periods (determined by atomic motion) of 10^{-13} s or less. At the surface of a material such as a good conductor where charge often concentrates, averaging with a radius on the order of 10^{-10} m may be required to resolve the rapid variation in the distribution of individual charged particles. However, within a solid or liquid material, or within a free-charge distribution characteristic of a dense gas or an electron beam, a radius of 10^{-8} m proves useful, containing typically 10^6 particles. A diffuse gas, on the other hand, may have a particle density so low that the averaging radius takes on laboratory dimensions, and in such a case the microscopic theory must be employed even at macroscopic dimensions.

Once the averaging radius has been determined, the value of the charge density may be found via (1.1). The volume density of charge for an assortment of point sources can

be written in terms of the three-dimensional Dirac delta as

$$\rho^o(\mathbf{r},t) = \sum_i q_i \delta(\mathbf{r} - \mathbf{r}_i(t)),$$

where $\mathbf{r}_i(t)$ is the position of the charge q_i at time t. Substitution into (1.1) gives

$$\rho(\mathbf{r},t) = \langle \rho^o(\mathbf{r},t) \rangle = \sum_i q_i f(\mathbf{r} - \mathbf{r}_i(t)) \tag{1.2}$$

as the averaged charge density appropriate for use in a macroscopic field theory. Because the oscillations of the atomic particles are statistically uncorrelated over the distances used in spatial averaging, the time variations of microscopic fields are not present in the macroscopic fields and temporal averaging is unnecessary. In (1.2) the time dependence of the spatially-averaged charge density is due entirely to bulk motion of the charge aggregate (macroscopic charge motion).

With the definition of macroscopic charge density given by (1.2), we can determine the total charge $Q(t)$ in any macroscopic volume region V using

$$Q(t) = \int_V \rho(\mathbf{r},t)\, dV. \tag{1.3}$$

We have

$$Q(t) = \sum_i q_i \int_V f(\mathbf{r} - \mathbf{r}_i(t))\, dV = \sum_{\mathbf{r}_i(t) \in V} q_i.$$

Here we ignore the small discrepancy produced by charges lying within distance a of the boundary of V. It is common to employ a box B having volume ΔV:

$$f(\mathbf{r}) = \begin{cases} 1/\Delta V, & \mathbf{r} \in B, \\ 0, & \mathbf{r} \notin B. \end{cases}$$

In this case

$$\rho(\mathbf{r},t) = \frac{1}{\Delta V} \sum_{\mathbf{r} - \mathbf{r}_i(t) \in B} q_i.$$

The size of B is chosen with the same considerations as to atomic scale as was the averaging radius a. Discontinuities at the edges of the box introduce some difficulties concerning charges that move in and out of the box because of molecular motion.

The macroscopic volume current density. Electric charge in motion is referred to as *electric current*. Charge motion can be associated with external forces and with microscopic fluctuations in position. Assuming charge q_i has velocity $\mathbf{v}_i(t) = d\mathbf{r}_i(t)/dt$, the charge aggregate has volume current density

$$\mathbf{J}^o(\mathbf{r},t) = \sum_i q_i \mathbf{v}_i(t)\, \delta(\mathbf{r} - \mathbf{r}_i(t)).$$

Spatial averaging gives the macroscopic volume current density

$$\mathbf{J}(\mathbf{r},t) = \langle \mathbf{J}^o(\mathbf{r},t) \rangle = \sum_i q_i \mathbf{v}_i(t) f(\mathbf{r} - \mathbf{r}_i(t)). \tag{1.4}$$

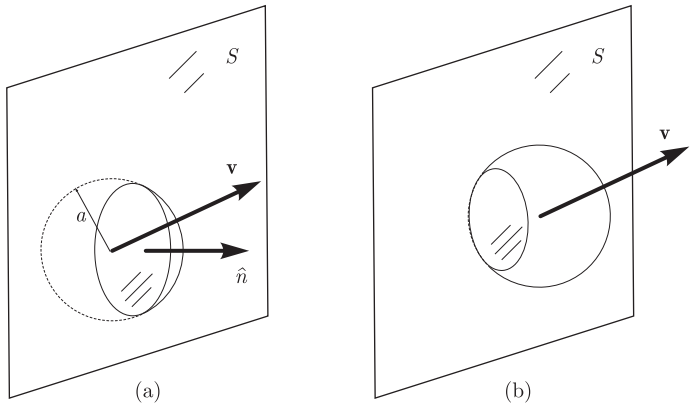

FIGURE 1.1
Intersection of the averaging function of a point charge with a surface S, as the charge crosses S with velocity $\mathbf{v}$: (a) at some time $t = t_1$, and (b) at $t = t_2 > t_1$. The averaging function is represented by a sphere of radius a.

Spatial averaging at time t eliminates currents associated with microscopic motions that are uncorrelated at the scale of the averaging radius (again, we do not consider the magnetic moments of particles). The assumption of a sufficiently large averaging radius leads to

$$\mathbf{J}(\mathbf{r}, t) = \rho(\mathbf{r}, t)\, \mathbf{v}(\mathbf{r}, t). \tag{1.5}$$

The total flux $I(t)$ of current through a surface S is given by

$$I(t) = \int_S \mathbf{J}(\mathbf{r}, t) \cdot \hat{\mathbf{n}}\, dS$$

where $\hat{\mathbf{n}}$ is the unit normal to S. Hence, using (1.4), we have

$$I(t) = \sum_i q_i \frac{d}{dt}(\mathbf{r}_i(t) \cdot \hat{\mathbf{n}}) \int_S f(\mathbf{r} - \mathbf{r}_i(t))\, dS$$

if $\hat{\mathbf{n}}$ stays approximately constant over the extent of the averaging function and S is not in motion. We see that the integral effectively intersects S with the averaging function surrounding each moving point charge (Figure 1.1). The time derivative of $\mathbf{r}_i \cdot \hat{\mathbf{n}}$ represents the velocity at which the averaging function is "carried across" the surface.

Electric current takes a variety of forms, each described by the relation $\mathbf{J} = \rho\mathbf{v}$. Isolated charged particles (positive and negative) and charged insulated bodies moving through space comprise *convection currents*. Negatively-charged electrons moving through the positive background lattice within a conductor comprise a *conduction current*. Empirical evidence suggests that conduction currents are also described by the relation $\mathbf{J} = \sigma\mathbf{E}$ known as *Ohm's law*. A third type of current, called *electrolytic current*, results from the flow of positive or negative ions through a fluid.

1.3.2 Impressed vs. secondary sources

In addition to the simple classification given above we may classify currents as *primary* or *secondary*, depending on the action that sets the charge in motion.

It is helpful to separate primary or "impressed" sources, which are independent of the fields they source, from secondary sources which result from interactions between the sourced fields and the medium in which the fields exist. Most familiar is the conduction current set up in a conducting medium by an externally applied electric field. The impressed source concept is particularly important in circuit theory, where independent voltage sources are modeled as providing primary voltage excitations that are independent of applied load. In this way they differ from the secondary or "dependent" sources that react to the effect produced by the application of primary sources.

In applied electromagnetics the primary source may be so distant that return effects resulting from local interaction of its impressed fields can be ignored. Other examples of primary sources include the applied voltage at the input of an antenna, the current on a probe inserted into a waveguide, and the currents producing a power-line field in which a biological body is immersed.

1.3.3 Surface and line source densities

Because they are spatially averaged effects, macroscopic sources and the fields they source cannot have true spatial discontinuities. However, it is often convenient to work with sources in one or two dimensions. Surface and line source densities are idealizations of actual, continuous macroscopic densities.

The entity we describe as a *surface charge* is a continuous volume charge distributed in a thin layer across some surface S. If the thickness of the layer is small compared to laboratory dimensions, it is useful to assign to each point $\mathbf{r}$ on the surface a quantity describing the amount of charge contained within a cylinder oriented normal to the surface and having infinitesimal cross section dS. We call this quantity the *surface charge density* $\rho_s(\mathbf{r}, t)$, and write the volume charge density as

$$\rho(\mathbf{r}, w, t) = \rho_s(\mathbf{r}, t) f(w, \Delta),$$

where w is distance from S in the normal direction and Δ in some way parameterizes the "thickness" of the charge layer at $\mathbf{r}$. The continuous density function $f(x, \Delta)$ satisfies

$$\int_{-\infty}^{\infty} f(x, \Delta)\, dx = 1$$

and

$$\lim_{\Delta \to 0} f(x, \Delta) = \delta(x).$$

For instance, we might have

$$f(x, \Delta) = \frac{e^{-x^2/\Delta^2}}{\Delta\sqrt{\pi}}. \tag{1.6}$$

With this definition the total charge contained in a cylinder normal to the surface at $\mathbf{r}$ and having cross-sectional area dS is

$$dQ(t) = \int_{-\infty}^{\infty} [\rho_s(\mathbf{r}, t)\, dS]\, f(w, \Delta)\, dw = \rho_s(\mathbf{r}, t)\, dS,$$

and the total charge contained within any cylinder oriented normal to S is

$$Q(t) = \int_S \rho_s(\mathbf{r}, t)\, dS. \tag{1.7}$$

We may describe a line charge as a thin "tube" of volume charge distributed along some contour Γ. The amount of charge contained between two planes normal to the contour and separated by a distance dl is described by the *line charge density* $\rho_l(\mathbf{r}, t)$. The volume charge density associated with the contour is then

$$\rho(\mathbf{r}, \rho, t) = \rho_l(\mathbf{r}, t) f_s(\rho, \Delta),$$

where ρ is the radial distance from the contour in the plane normal to Γ and $f_s(\rho, \Delta)$ is a density function with the properties

$$\int_0^\infty f_s(\rho, \Delta) 2\pi\rho \, d\rho = 1$$

and

$$\lim_{\Delta \to 0} f_s(\rho, \Delta) = \frac{\delta(\rho)}{2\pi\rho}.$$

For example, we might have

$$f_s(\rho, \Delta) = \frac{e^{-\rho^2/\Delta^2}}{\pi\Delta^2}. \tag{1.8}$$

Then the total charge contained between planes separated by a distance dl is

$$dQ(t) = \int_0^\infty [\rho_l(\mathbf{r}, t) \, dl] \, f_s(\rho, \Delta) 2\pi\rho \, d\rho = \rho_l(\mathbf{r}, t) \, dl$$

and the total charge contained between planes placed at the ends of a contour Γ is

$$Q(t) = \int_\Gamma \rho_l(\mathbf{r}, t) \, dl. \tag{1.9}$$

We may define surface and line currents similarly. A surface current is merely a volume current confined to the vicinity of a surface S. The volume current density may be represented using a surface current density function $\mathbf{J}_s(\mathbf{r}, t)$, defined at each point $\mathbf{r}$ on the surface so that

$$\mathbf{J}(\mathbf{r}, w, t) = \mathbf{J}_s(\mathbf{r}, t) f(w, \Delta).$$

Here $f(w, \Delta)$ is some appropriate density function such as (1.6), and the surface current vector obeys $\hat{\mathbf{n}} \cdot \mathbf{J}_s = 0$ where $\hat{\mathbf{n}}$ is normal to S. The total current flowing through a strip of width dl arranged perpendicular to S at $\mathbf{r}$ is

$$dI(t) = \int_{-\infty}^\infty [\mathbf{J}_s(\mathbf{r}, t) \cdot \hat{\mathbf{n}}_l(\mathbf{r}) \, dl] \, f(w, \Delta) \, dw = \mathbf{J}_s(\mathbf{r}, t) \cdot \hat{\mathbf{n}}_l(\mathbf{r}) \, dl$$

where $\hat{\mathbf{n}}_l$ is normal to the strip at $\mathbf{r}$ (and thus also tangential to S at $\mathbf{r}$). The total current passing through a strip intersecting with S along a contour Γ is thus

$$I(t) = \int_\Gamma \mathbf{J}_s(\mathbf{r}, t) \cdot \hat{\mathbf{n}}_l(\mathbf{r}) \, dl.$$

We may describe a line current as a thin "tube" of volume current distributed about some contour Γ and flowing parallel to it. The amount of current passing through a plane normal to the contour is described by the *line current density* $J_l(\mathbf{r}, t)$. The volume current density associated with the contour may be written as

$$\mathbf{J}(\mathbf{r}, \rho, t) = \hat{\mathbf{u}}(\mathbf{r}) J_l(\mathbf{r}, t) f_s(\rho, \Delta),$$

where $\hat{\mathbf{u}}$ is a unit vector along Γ, ρ is the radial distance from the contour in the plane normal to Γ, and $f_s(\rho, \Delta)$ is a density function such as (1.8). The total current passing through any plane normal to Γ at $\mathbf{r}$ is

$$I(t) = \int_0^\infty \left[J_l(\mathbf{r}, t)\hat{\mathbf{u}}(\mathbf{r}) \cdot \hat{\mathbf{u}}(\mathbf{r})\right] f_s(\rho, \Delta) 2\pi\rho \, d\rho = J_l(\mathbf{r}, t).$$

It is often convenient to employ *singular* models for continuous source densities. For instance, it is mathematically simpler to regard a surface charge as residing only in the surface S than to regard it as being distributed about the surface. Of course, the source is then discontinuous since it is zero everywhere outside the surface. We may obtain a representation of such a charge distribution by letting the thickness parameter Δ in the density functions recede to zero, thus concentrating the source into a plane or a line. We describe the limit of the density function in terms of the δ-function. For instance, the volume charge distribution for a surface charge located about the xy-plane is

$$\rho(x, y, z, t) = \rho_s(x, y, t) f(z, \Delta).$$

As $\Delta \to 0$ we have

$$\rho(x, y, z, t) = \rho_s(x, y, t) \lim_{\Delta \to 0} f(z, \Delta) = \rho_s(x, y, t)\delta(z).$$

It is a simple matter to represent singular source densities in this way as long as the surface or line is easily parameterized in terms of constant values of coordinate variables. However, care must be taken to represent the δ-function properly. For instance, the density of charge on the surface of a cone at $\theta = \theta_0$ may be described using the distance normal to this surface, which is given by $r\theta - r\theta_0$:

$$\rho(r, \theta, \phi, t) = \rho_s(r, \phi, t)\delta\left(r[\theta - \theta_0]\right).$$

Using the property $\delta(ax) = \delta(x)/a$, we can also write this as

$$\rho(r, \theta, \phi, t) = \rho_s(r, \phi, t)\frac{\delta(\theta - \theta_0)}{r}.$$

1.3.4 Charge conservation

There are four fundamental conservation laws in physics: conservation of energy, momentum, angular momentum, and charge. These laws are said to be *absolute*; they have never been observed to fail. In that sense they are true empirical laws of physics.

However, in modern physics the fundamental conservation laws have come to represent more than just observed facts. Each law is now associated with a fundamental symmetry of the universe; conversely, each known symmetry is associated with a conservation principle. For example, energy conservation can be shown to arise from the observation that the universe is symmetric with respect to time; the laws of physics do not depend on choice of time origin $t = 0$. Similarly, momentum conservation arises from the observation that the laws of physics are invariant under translation, while angular momentum conservation arises from invariance under rotation.

The law of conservation of charge also arises from a symmetry principle. But instead of being spatial or temporal in character, it is related to the invariance of electrostatic potential. Experiments show that there is no absolute potential, only potential difference. The laws of nature are invariant with respect to what we choose as the "reference"

potential. This in turn is related to the invariance of Maxwell's equations under gauge transforms; the values of the electric and magnetic fields do not depend on which gauge transformation we use to relate the scalar potential Φ to the vector potential $\mathbf{A}$.

We may state the conservation of charge as follows:

The net charge in any closed system remains constant with time.

This does not mean that individual charges cannot be created or destroyed, only that the total charge in any isolated system must remain constant. Thus it is possible for a positron with charge e to annihilate an electron with charge $-e$ without changing the net charge of the system. Only if a system is not closed can its net charge be altered; since moving charge constitutes current, we can say that the total charge within a system depends on the current passing through the surface enclosing the system. This is the essence of the continuity equation. To derive this important result we consider a closed system within which the charge remains constant, and apply the Reynolds transport theorem (see § A.2).

The continuity equation. Consider a region of space occupied by a distribution of charge whose velocity is given by the vector field $\mathbf{v}$. We surround a portion of charge by a surface S and let S deform as necessary to "follow" the charge as it moves. Since S always contains precisely the same charged particles, we have an isolated system for which the time rate of change of total charge must vanish. An expression for the time rate of change is given by the Reynolds transport theorem (A.66); we have[†]

$$\frac{DQ}{Dt} = \frac{D}{Dt}\int_{V(t)} \rho \, dV = \int_{V(t)} \frac{\partial \rho}{\partial t}\, dV + \oint_{S(t)} \rho \mathbf{v} \cdot \mathbf{dS} = 0.$$

The "D/Dt" notation indicates that the volume region $V(t)$ moves with its enclosed particles. Since $\rho\mathbf{v}$ represents current density, we can write

$$\int_{V(t)} \frac{\partial \rho(\mathbf{r}, t)}{\partial t}\, dV + \oint_{S(t)} \mathbf{J}(\mathbf{r}, t) \cdot \mathbf{dS} = 0. \tag{1.10}$$

In this *large-scale form* of the continuity equation, the partial derivative term describes the time rate of change of the charge density for a fixed spatial position $\mathbf{r}$. At any time t, the time rate of change of charge density integrated over a volume is exactly compensated by the total current exiting through the surrounding surface.

We can obtain the continuity equation in *point form* by applying the divergence theorem to the second term of (1.10) to get

$$\int_{V(t)} \left[\frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{J}(\mathbf{r}, t) \right] dV = 0.$$

Since $V(t)$ is arbitrary we can set the integrand to zero to obtain

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = 0. \tag{1.11}$$

[†]Note that in Appendix A we use the symbol $\mathbf{u}$ to represent the velocity of a material and $\mathbf{v}$ to represent the velocity of an artificial surface.

This expression involves the time derivative of ρ with $\mathbf{r}$ fixed. We can also find an expression in terms of the material derivative by using the transport equation (A.67). Enforcing conservation of charge by setting that expression to zero, we have

$$\frac{D\rho(\mathbf{r},t)}{Dt} + \rho(\mathbf{r},t)\,\nabla\cdot\mathbf{v}(\mathbf{r},t) = 0. \tag{1.12}$$

Here $D\rho/Dt$ is the time rate of change of the charge density experienced by an observer moving with the current.

We can state the large-scale form of the continuity equation in terms of a stationary volume. Integrating (1.11) over a stationary volume region V and using the divergence theorem, we find that

$$\int_V \frac{\partial\rho(\mathbf{r},t)}{\partial t}\,dV = -\oint_S \mathbf{J}(\mathbf{r},t)\cdot\mathbf{dS}.$$

Since V is not changing with time we have

$$\frac{dQ(t)}{dt} = \frac{d}{dt}\int_V \rho(\mathbf{r},t)\,dV = -\oint_S \mathbf{J}(\mathbf{r},t)\cdot\mathbf{dS}. \tag{1.13}$$

Hence any increase of total charge within V must be produced by current entering V through S.

Use of the continuity equation. As an example, suppose that in a bounded region of space we have

$$\rho(\mathbf{r},t) = \rho_0 r e^{-\beta t}.$$

We wish to find $\mathbf{J}$ and $\mathbf{v}$, and to verify both versions of the continuity equation in point form. The spherical symmetry of ρ requires that $\mathbf{J} = \hat{\mathbf{r}}J_r$. Application of (1.13) over a sphere of radius a gives

$$4\pi\frac{d}{dt}\int_0^a \rho_0 r e^{-\beta t} r^2\,dr = -4\pi J_r(a)a^2.$$

Hence

$$\mathbf{J} = \hat{\mathbf{r}}\beta\rho_0\frac{r^2}{4}e^{-\beta t}$$

and therefore

$$\nabla\cdot\mathbf{J} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 J_r) = \beta\rho_0 r e^{-\beta t}.$$

The velocity is

$$\mathbf{v} = \frac{\mathbf{J}}{\rho} = \hat{\mathbf{r}}\beta\frac{r}{4},$$

and we have $\nabla\cdot\mathbf{v} = 3\beta/4$. To verify the continuity equations, we compute the time derivatives

$$\frac{\partial\rho}{\partial t} = -\beta\rho_0 r e^{-\beta t},$$

$$\begin{aligned}\frac{D\rho}{Dt} &= \frac{\partial\rho}{\partial t} + \mathbf{v}\cdot\nabla\rho \\ &= -\beta\rho_0 r e^{-\beta t} + \left(\hat{\mathbf{r}}\beta\frac{r}{4}\right)\cdot\left(\hat{\mathbf{r}}\rho_0 e^{-\beta t}\right) \\ &= -\frac{3}{4}\beta\rho_0 r e^{-\beta t}.\end{aligned}$$

Note that the charge density decreases with time less rapidly for a moving observer than for a stationary one (3/4 as fast): the moving observer is following the charge outward, and $\rho \propto r$. Now we can check the continuity equations. First we see

$$\frac{D\rho}{Dt} + \rho\nabla \cdot \mathbf{v} = -\frac{3}{4}\beta\rho_0 r e^{-\beta t} + (\rho_0 r e^{-\beta t})\left(\frac{3}{4}\beta\right) = 0,$$

as required for a moving observer; second we see

$$\frac{\partial\rho}{\partial t} + \nabla \cdot \mathbf{J} = -\beta\rho_0 r e^{-\beta t} + \beta\rho_0 e^{-\beta t} = 0,$$

as required for a stationary observer.

The continuity equation in fewer dimensions. The continuity equation can also be used to relate current and charge on a surface or along a line. By conservation of charge we can write

$$\frac{d}{dt}\int_S \rho_s(\mathbf{r}, t)\, dS = -\oint_\Gamma \mathbf{J}_s(\mathbf{r}, t) \cdot \hat{\mathbf{m}}\, dl \tag{1.14}$$

where $\hat{\mathbf{m}}$ is the vector normal to the curve Γ and tangential to the surface S. By the surface divergence theorem (B.20), the corresponding point form is

$$\frac{\partial\rho_s(\mathbf{r}, t)}{\partial t} + \nabla_s \cdot \mathbf{J}_s(\mathbf{r}, t) = 0. \tag{1.15}$$

Here $\nabla_s \cdot \mathbf{J}_s$ is the surface divergence of the vector field $\mathbf{J}_s$. For instance, in rectangular coordinates in the $z = 0$ plane we have

$$\nabla_s \cdot \mathbf{J}_s = \frac{\partial J_{sx}}{\partial x} + \frac{\partial J_{sy}}{\partial y}.$$

In cylindrical coordinates on the cylinder $\rho = a$, we would have

$$\nabla_s \cdot \mathbf{J}_s = \frac{1}{a}\frac{\partial J_{s\phi}}{\partial \phi} + \frac{\partial J_{sz}}{\partial z}.$$

A detailed description of vector operations on a surface may be found in Tai [222], while many identities may be found in Van Bladel [238].

The equation of continuity for a line is easily established by reference to Figure 1.2. Here the net charge exiting the surface during time Δt is given by

$$\Delta t[I(u_2, t) - I(u_1, t)].$$

Thus, the rate of net increase of charge within the system is

$$\frac{dQ(t)}{dt} = \frac{d}{dt}\int \rho_l(\mathbf{r}, t)\, dl = -[I(u_2, t) - I(u_1, t)]. \tag{1.16}$$

The corresponding point form is found by letting the length of the curve approach zero:

$$\frac{\partial I(l, t)}{\partial l} + \frac{\partial \rho_l(l, t)}{\partial t} = 0, \tag{1.17}$$

where l is arc length along the curve. As an example, suppose the line current on a circular loop antenna is approximately

$$I(\phi, t) = I_0 \cos\left(\frac{\omega a}{c}\phi\right)\cos\omega t,$$

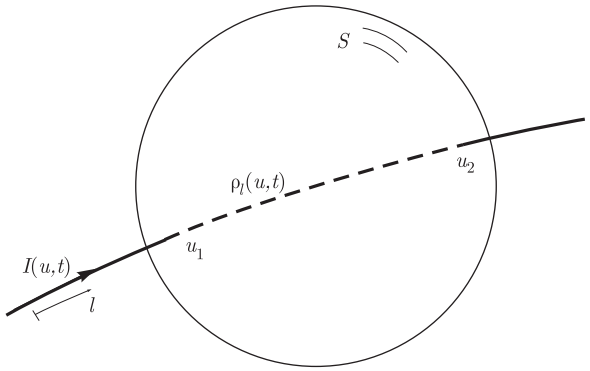

FIGURE 1.2
Linear form of the continuity equation.

where a is the radius of the loop, ω is the frequency of operation, and c is the speed of light. We wish to find the line charge density on the loop. Since $l = a\phi$, we can write

$$I(l,t) = I_0 \cos\left(\frac{\omega l}{c}\right)\cos\omega t.$$

Thus

$$\frac{\partial I(l,t)}{\partial l} = -I_0\frac{\omega}{c}\sin\left(\frac{\omega l}{c}\right)\cos\omega t = -\frac{\partial \rho_l(l,t)}{\partial t}.$$

Integrating with respect to time and ignoring any constant (static) charge, we have

$$\rho(l,t) = \frac{I_0}{c}\sin\left(\frac{\omega l}{c}\right)\sin\omega t$$

or

$$\rho(\phi,t) = \frac{I_0}{c}\sin\left(\frac{\omega a}{c}\phi\right)\sin\omega t.$$

Note that we could have used the chain rule

$$\frac{\partial I(\phi,t)}{\partial l} = \frac{\partial I(\phi,t)}{\partial \phi}\frac{\partial \phi}{\partial l} \quad \text{and} \quad \frac{\partial \phi}{\partial l} = \left[\frac{\partial l}{\partial \phi}\right]^{-1} = \frac{1}{a}$$

to calculate the spatial derivative.

We can apply the volume density continuity equation (1.11) directly to surface and line distributions written in singular notation. For the loop of the previous example, we write the volume current density corresponding to the line current as

$$\mathbf{J}(\mathbf{r},t) = \hat{\boldsymbol{\phi}}\,\delta(\rho - a)\delta(z)I(\phi,t).$$

Substitution into (1.11) then gives

$$\nabla\cdot[\hat{\boldsymbol{\phi}}\delta(\rho-a)\delta(z)I(\phi,t)] = -\frac{\partial\rho(\mathbf{r},t)}{\partial t}.$$

The divergence formula for cylindrical coordinates gives

$$\delta(\rho-a)\delta(z)\frac{\partial I(\phi,t)}{\rho\,\partial\phi} = -\frac{\partial\rho(\mathbf{r},t)}{\partial t}.$$

Next we substitute for $I(\phi, t)$ to get

$$-\frac{I_0}{\rho}\frac{\omega a}{c}\sin\left(\frac{\omega a}{c}\phi\right)\delta(\rho - a)\delta(z)\cos\omega t = -\frac{\partial\rho(\mathbf{r}, t)}{\partial t}.$$

Finally, integrating with respect to time and ignoring any constant term, we have

$$\rho(\mathbf{r}, t) = \frac{I_0}{c}\delta(\rho - a)\delta(z)\sin\left(\frac{\omega a}{c}\phi\right)\sin\omega t,$$

where we have set $\rho = a$ because of the presence of the factor $\delta(\rho - a)$.

1.3.5 Magnetic charge

We take for granted that electric fields are produced by electric charges, whether stationary or in motion. The smallest element of electric charge is the electric *monopole*: a single discretely charged particle from which the electric field diverges. In contrast, experiments show that magnetic fields are created only by currents or by time changing electric fields; hence, magnetic fields have moving electric charge as their source. The elemental source of magnetic field is the magnetic *dipole*, representing a tiny loop of electric current (or a spinning electric particle). The observation made in 1269 by Pierre de Maricourt, that even the smallest magnet has two poles, still holds today.

In a world filled with symmetry at the fundamental level, we find it hard to understand why there should not be a source from which the magnetic field diverges. We would call such a source *magnetic charge*, and the most fundamental quantity of magnetic charge would be exhibited by a *magnetic monopole*. In 1931 Paul Dirac invigorated the search for magnetic monopoles by making the first strong theoretical argument for their existence. Dirac showed that the existence of magnetic monopoles would imply the quantization of electric charge, and would thus provide an explanation for one of the great puzzles of science. Since that time magnetic monopoles have become important players in the "Grand Unified Theories" of modern physics, and in cosmological theories of the origin of the universe.

If magnetic monopoles are ever found to exist, there will be both positive and negative charged particles whose motions will constitute currents. We can define a macroscopic magnetic charge density ρ_m and current density $\mathbf{J}_m$ exactly as we did with electric charge, and use conservation of magnetic charge to provide a continuity equation:

$$\nabla \cdot \mathbf{J}_m(\mathbf{r}, t) + \frac{\partial\rho_m(\mathbf{r}, t)}{\partial t} = 0. \tag{1.18}$$

With these new sources Maxwell's equations become appealingly symmetric. Despite uncertainties about the existence and physical nature of magnetic monopoles, magnetic charge and current have become an integral part of electromagnetic theory. We often use the concept of *fictitious* magnetic sources to make Maxwell's equations symmetric, and then derive various equivalence theorems for use in the solution of important problems. Thus we can put the idea of magnetic sources to use regardless of whether these sources actually exist.

1.4 Problems

1.1 Write the volume charge density for a singular surface charge located on the sphere $r = r_0$, entirely in terms of spherical coordinates. Find the total charge on the sphere.

1.2 Repeat Problem 1.1 for a charged half plane $\phi = \phi_0$.

1.3 Write the volume charge density for a singular surface charge located on the cylinder $\rho = \rho_0$, entirely in terms of cylindrical coordinates. Find the total charge on the cylinder.

1.4 Repeat Problem 1.3 for a charged half plane $\phi = \phi_0$.

1.5 A current flows radially outward from the z-axis of cylindrical coordinates. The volume charge density associated with the current is given by

$$\rho(\mathbf{r}, t) = \rho_0 \rho^2 e^{-\beta t}$$

where ρ_0 and β are constants. (a) Calculate the current density. (b) Show that the following two forms of the continuity equation are satisfied:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0, \qquad \frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0.$$

1.6 Compute the total charge associated with the following volume charge densities:

(a) $\rho = 4r^2 \cos^2\theta\, \delta(\theta - \pi/4)$, $0 \le \phi \le 2\pi$, $0 \le r \le 2$.

(b) $\rho = 4\cos^2\phi\, \delta(\rho - 2)$, $0 \le z \le 2$, $0 \le \phi \le 2\pi$.

(c) $\rho = 4z^3\, \delta(x)\delta(y)$, $0 \le z \le 2$.

(d) $\rho = 4\rho z \delta(\rho - 3)$, $0 \le z \le 2$, $0 \le \phi \le 2\pi$.

(e) $\rho = 5(z + 2)\delta(z)\delta(\rho - 3)$, $0 \le \phi \le \pi$.

1.7 A charge distribution has density $\rho(\mathbf{r}, t) = 4x^2 e^{-4t}$ C/m^3. Calculate the time rate of change of ρ measured by an observer moving with velocity $\mathbf{v} = \hat{\mathbf{x}} Ax$. Find the value of A such that the observer measures the same value of ρ at all times t.

1.8 The charge density in a certain region of free space is given by $\rho(\mathbf{r}, t) = 4r^2 e^{-\beta t}$ C/m^3. Find the time rate of change of the charge density measured by an observer moving with velocity $\mathbf{v} = \hat{\mathbf{r}} 2r$ m/s. For what value of β is the measured time rate of change identically zero?

2

Maxwell's theory of electromagnetism

2.1 The postulate

In 1864, James Clerk Maxwell proposed one of the most successful theories in the history of science. In a famous memoir to the Royal Society [143] he presented nine equations summarizing all known laws on electricity and magnetism. This was more than a mere cataloging of the laws of nature. By postulating the need for an additional term to make the set of equations self-consistent, Maxwell was able to put forth what is still considered a complete theory of macroscopic electromagnetism. The beauty of Maxwell's equations led Boltzmann to ask, "Was it a god who wrote these lines . . . ?" [213].

Since that time authors have struggled to find the best way to present Maxwell's theory. Although it is possible to study electromagnetics from an "empirical–inductive" viewpoint (roughly following the historical order of development beginning with static fields), it is only by postulating the complete theory that we can do justice to Maxwell's vision. His concept of the existence of an electromagnetic "field" (as introduced by Faraday) is fundamental to this theory, and has become one of the most significant principles of modern science.

We find controversy even over the best way to present Maxwell's equations. Maxwell worked at a time before vector notation was completely in place, and thus chose to use scalar variables and equations to represent the fields. Certainly the true beauty of Maxwell's equations emerges when they are written in vector form, and the use of tensors reduces the equations to their underlying physical simplicity. We shall use vector notation in this book because of its wide acceptance by engineers, but we still must decide whether it is more appropriate to present the vector equations in integral or point form.

On one side of this debate, the brilliant mathematician David Hilbert felt that the fundamental natural laws should be posited as axioms, each best described in terms of integral equations [177]. This idea has been championed by Truesdell and Toupin [234]. On the other side, we may quote from the great physicist Arnold Sommerfeld: "The general development of Maxwell's theory must proceed from its differential form; for special problems the integral form may, however, be more advantageous" ([213], p. 23). Special relativity flows naturally from the point forms, with fields easily converted between moving reference frames. For stationary media, it seems to us that the only difference between the two approaches arises in how we handle discontinuities in sources and materials. If we choose to use the point forms of Maxwell's equations, then we must also postulate the boundary conditions at surfaces of discontinuity. This is pointed out clearly by Tai [224], who also notes that if the integral forms are used, then their validity across regions of discontinuity should be stated as part of the postulate.

We have decided to use the point form in this text. In doing so we follow a long history begun by Hertz in 1890 [97] when he wrote down Maxwell's differential equations

as a set of axioms, recognizing the equations as the launching point for the theory of electromagnetism. Also, by postulating Maxwell's equations in point form we can take full advantage of modern developments in the theory of partial differential equations; in particular, the idea of a "well-posed" theory determines what sort of information must be specified to make the postulate useful.

We must also decide which form of Maxwell's differential equations to use as the basis of our postulate. There are several competing forms, each differing on the manner in which materials are considered. The oldest and most widely used form was suggested by Minkowski in 1908 [149]. In the Minkowski form the differential equations contain no mention of the materials supporting the fields; all information about material media is relegated to the constitutive relationships. This places simplicity of the differential equations above intuitive understanding of the behavior of fields in materials. We choose the Maxwell–Minkowski form as the basis of our postulate, primarily for ease of manipulation. But we also recognize the value of other versions of Maxwell's equations. We shall present the basic ideas behind the Boffi form, which places some information about materials into the differential equations (although constitutive relationships are still required). Missing, however, is any information regarding the velocity of a moving medium. By using the polarization and magnetization vectors $\mathbf{P}$ and $\mathbf{M}$ rather than the fields $\mathbf{D}$ and $\mathbf{H}$, it is sometimes easier to visualize the meaning of the field vectors and to understand (or predict) the nature of the constitutive relations.

The Chu and Amperian forms of Maxwell's equations have been promoted as useful alternatives to the Minkowski and Boffi forms. These include explicit information about the velocity of a moving material, and differ somewhat from the Boffi form in the physical interpretation of the electric and magnetic properties of matter. Although each of these models matter in terms of charged particles immersed in free space, magnetization in the Boffi and Amperian forms arises from electric current loops, while the Chu form employs magnetic dipoles. In all three forms polarization is modeled using electric dipoles. For a detailed discussion of the Chu and Amperian forms, the reader should consult the work of Kong [118], Tai [225], Penfield and Haus [166], or Fano, Chu and Adler [78].

Importantly, all of these various forms of Maxwell's equations produce the same values of the physical fields (at least external to the material where the fields are measurable).

We must include several other constituents, besides the field equations, to make the postulate complete. To form a complete field theory we need a source field, a mediating field, and a set of field differential equations. This allows us to mathematically describe the relationship between effect (the mediating field) and cause (the source field). In a well-posed postulate we must also include a set of constitutive relationships and a specification of some field relationship over a bounding surface and at an initial time. If the electromagnetic field is to have physical meaning, we must link it to some observable quantity such as force. Finally, to allow the solution of problems involving mathematical discontinuities we must specify certain boundary, or "jump," conditions.

2.1.1 The Maxwell–Minkowski equations

In Maxwell's macroscopic theory of electromagnetics, the source field consists of the vector field $\mathbf{J}(\mathbf{r}, t)$ (the current density) and the scalar field $\rho(\mathbf{r}, t)$ (the charge density). In Minkowski's form of Maxwell's equations, the mediating field is the *electromagnetic field* consisting of the set of four vector fields $\mathbf{E}(\mathbf{r}, t)$, $\mathbf{D}(\mathbf{r}, t)$, $\mathbf{B}(\mathbf{r}, t)$, and $\mathbf{H}(\mathbf{r}, t)$. The field equations are the four partial differential equations referred to as the *Maxwell–*

Minkowski equations

$$\nabla \times \mathbf{E}(\mathbf{r},t) = -\frac{\partial}{\partial t}\mathbf{B}(\mathbf{r},t), \tag{2.1}$$

$$\nabla \times \mathbf{H}(\mathbf{r},t) = \mathbf{J}(\mathbf{r},t) + \frac{\partial}{\partial t}\mathbf{D}(\mathbf{r},t), \tag{2.2}$$

$$\nabla \cdot \mathbf{D}(\mathbf{r},t) = \rho(\mathbf{r},t), \tag{2.3}$$

$$\nabla \cdot \mathbf{B}(\mathbf{r},t) = 0, \tag{2.4}$$

along with the continuity equation

$$\nabla \cdot \mathbf{J}(\mathbf{r},t) = -\frac{\partial}{\partial t}\rho(\mathbf{r},t). \tag{2.5}$$

Here (2.1) is called *Faraday's law*, (2.2) is called *Ampere's law*, (2.3) is called *Gauss's law*, and (2.4) is called the *magnetic Gauss's law*. For brevity we shall often leave the dependence on $\mathbf{r}$ and t implicit, and refer to the Maxwell–Minkowski equations as simply the "Maxwell equations," or "Maxwell's equations."

Equations (2.1)–(2.5), the point forms of the field equations, describe the relationships between the fields and their sources at each point in space where the fields are continuously differentiable (i.e., the derivatives exist and are continuous). Such points are called *ordinary points*. We shall not attempt to define the fields at other points, but instead seek conditions relating the fields across surfaces containing these points. Normally this is necessary on surfaces across which either sources or material parameters are discontinuous.

The electromagnetic fields carry SI units as follows: $\mathbf{E}$ is measured in Volts per meter (V/m), $\mathbf{B}$ is measured in Teslas (T), $\mathbf{H}$ is measured in Amperes per meter (A/m), and $\mathbf{D}$ is measured in Coulombs per square meter ($\mathrm{C/m^2}$). In older texts we find the units of $\mathbf{B}$ given as Webers per square meter ($\mathrm{Wb/m^2}$) to reflect the role of $\mathbf{B}$ as a flux vector; in that case the Weber ($\mathrm{Wb} = \mathrm{T{\cdot}m^2}$) is regarded as a unit of magnetic flux.

The interdependence of Maxwell's equations. It is often claimed that the divergence equations (2.3) and (2.4) may be derived from the curl equations (2.1) and (2.2). While this is true, it is *not* proper to say that only the two curl equations are required to describe Maxwell's theory. This is because an additional physical assumption, not present in the two curl equations, is required to complete the derivation. Either the divergence equations must be specified, or the values of certain constants that fix the initial conditions on the fields must be specified. It is customary to specify the divergence equations and include them with the curl equations to form the complete set we now call "Maxwell's equations."

To identify the interdependence we take the divergence of (2.1) to get

$$\nabla \cdot (\nabla \times \mathbf{E}) = \nabla \cdot \left(-\frac{\partial \mathbf{B}}{\partial t}\right),$$

hence

$$\frac{\partial}{\partial t}(\nabla \cdot \mathbf{B}) = 0$$

by (B.49). This requires that $\nabla \cdot \mathbf{B}$ be constant with time, say $\nabla \cdot \mathbf{B}(\mathbf{r},t) = C_B(\mathbf{r})$. The constant C_B must be specified as part of the postulate of Maxwell's theory, and the choice we make is subject to experimental validation. We postulate that $C_B(\mathbf{r}) = 0$,

which leads us to (2.4). Note that if we can identify a time prior to which $\mathbf{B}(\mathbf{r}, t) \equiv 0$, then $C_B(\mathbf{r})$ must vanish. For this reason, $C_B(\mathbf{r}) = 0$ and (2.4) are often called the "initial conditions" for Faraday's law [183].

Next we take the divergence of (2.2) to find that

$$\nabla \cdot (\nabla \times \mathbf{H}) = \nabla \cdot \mathbf{J} + \frac{\partial}{\partial t}(\nabla \cdot \mathbf{D}).$$

Using (2.5) and (B.49), we obtain

$$\frac{\partial}{\partial t}(\rho - \nabla \cdot \mathbf{D}) = 0$$

and thus $\rho - \nabla \cdot \mathbf{D}$ must be some temporal constant $C_D(\mathbf{r})$. Again, we must postulate the value of C_D as part of the Maxwell theory. We choose $C_D(\mathbf{r}) = 0$ and thus obtain Gauss's law (2.3). If we can identify a time prior to which both $\mathbf{D}$ and ρ are everywhere equal to zero, then $C_D(\mathbf{r})$ must vanish. Hence $C_D(\mathbf{r}) = 0$ and (2.3) may be regarded as "initial conditions" for Ampere's law. Combining the two sets of initial conditions, we find that the curl equations imply the divergence equations as long as we can find a time prior to which all of the fields $\mathbf{E}, \mathbf{D}, \mathbf{B}, \mathbf{H}$ and the sources $\mathbf{J}$ and ρ are equal to zero (since all the fields are related through the curl equations, and the charge and current are related through the continuity equation). Conversely, the empirical evidence supporting the two divergence equations implies that such a time should exist.

Throughout this book we shall refer to the two curl equations as the "fundamental" Maxwell equations, and to the two divergence equations as the "auxiliary" equations. The fundamental equations describe the relationships between the fields while, as we have seen, the auxiliary equations provide a sort of initial condition. This does not imply that the auxiliary equations are of lesser importance; indeed, they are required to establish uniqueness of the fields, to derive the wave equations for the fields, and to properly describe static fields.

Field vector terminology. Various terms are used for the field vectors, sometimes harkening back to the descriptions used by Maxwell himself, and often based on the physical nature of the fields. We are attracted to Sommerfeld's separation of the fields into *entities of intensity* $(\mathbf{E}, \mathbf{B})$ and *entities of quantity* $(\mathbf{D}, \mathbf{H})$. In this system $\mathbf{E}$ is called the *electric field strength*, $\mathbf{B}$ the *magnetic field strength*, $\mathbf{D}$ the *electric excitation*, and $\mathbf{H}$ the *magnetic excitation* [213]. Maxwell separated the fields into a set $(\mathbf{E}, \mathbf{H})$ of vectors that appear within line integrals to give work-related quantities, and a set $(\mathbf{B}, \mathbf{D})$ of vectors that appear within surface integrals to give flux-related quantities; we shall see this clearly when considering the integral forms of Maxwell's equations. By this system, authors such as Jones [112] and Ramo, Whinnery, and Van Duzer [176] call $\mathbf{E}$ the *electric intensity*, $\mathbf{H}$ the *magnetic intensity*, $\mathbf{B}$ the *magnetic flux density*, and $\mathbf{D}$ the *electric flux density*.

Maxwell himself designated names for each of the vector quantities. In his classic paper "A Dynamical Theory of the Electromagnetic Field," [206] Maxwell referred to the quantity we now designate $\mathbf{E}$ as the *electromotive force*, the quantity $\mathbf{D}$ as the *electric displacement* (with a time rate of change given by his now famous "displacement current"), the quantity $\mathbf{H}$ as the *magnetic force*, and the quantity $\mathbf{B}$ as the *magnetic induction* (although he described $\mathbf{B}$ as a density of lines of magnetic force). Maxwell also included a quantity designated *electromagnetic momentum* as an integral part of his theory. We now know this as the vector potential $\mathbf{A}$, which is not generally included as a part of the electromagnetics postulate.

Many authors follow the original terminology of Maxwell, with some slight modifications. For instance, Stratton [217] calls **E** the *electric field intensity*, **H** the *magnetic field intensity*, **D** the *electric displacement*, and **B** the *magnetic induction*. Jackson [105] calls **E** the *electric field*, **H** the *magnetic field*, **D** the *displacement*, and **B** the *magnetic induction*.

Other authors choose freely among combinations of these terms. For instance, Kong [118] calls **E** the *electric field strength*, **H** the *magnetic field strength*, **B** the *magnetic flux density*, and **D** the *electric displacement*. We do not wish to inject further confusion into the issue of nomenclature; still, we find it helpful to use as simple a naming system as possible. We shall refer to **E** as the *electric field*, **H** as the *magnetic field*, **D** as the *electric flux density* and **B** as the *magnetic flux density*. When we use the term *electromagnetic field* we imply the entire set of field vectors $(\mathbf{E}, \mathbf{D}, \mathbf{B}, \mathbf{H})$ used in Maxwell's theory.

Invariance of Maxwell's equations. Maxwell's differential equations are valid for any system in uniform relative motion with respect to the laboratory frame of reference in which we normally do our measurements. The field equations describe the relationships between the source and mediating fields *within that frame of reference*. This property was first proposed for moving material media by Minkowski in 1908 (using the term *covariance*) [149]. For this reason, Maxwell's equations expressed in the form (2.1)–(2.2) are referred to as the *Minkowski form*.

2.1.2 Connection to mechanics

Our postulate must include a connection between the abstract quantities of charge and field and a measurable physical quantity. A convenient means of linking electromagnetics to other classical theories is through mechanics. We postulate that charges experience mechanical forces given by the *Lorentz force equation*. If a small volume element dV contains a total charge $\rho\, dV$, then the force experienced by that charge when moving at velocity $\mathbf{v}$ in an electromagnetic field is

$$d\mathbf{F} = \rho\, dV\, \mathbf{E} + \rho\mathbf{v}\, dV \times \mathbf{B}. \tag{2.6}$$

As with any postulate, we verify this equation through experiment. Note that we write the Lorentz force in terms of charge $\rho\, dV$, rather than charge density ρ, since charge is an invariant quantity under a Lorentz transformation.

The important links between the electromagnetic fields and energy and momentum must also be postulated. We postulate that the quantity

$$\mathbf{S}_{em} = \mathbf{E} \times \mathbf{H} \tag{2.7}$$

represents the transport density of electromagnetic power, and that the quantity

$$\mathbf{g}_{em} = \mathbf{D} \times \mathbf{B} \tag{2.8}$$

represents the transport density of electromagnetic momentum.

2.2 The well-posed nature of the postulate

It is important to investigate whether Maxwell's equations, along with the point form of the continuity equation, suffice as a useful theory of electromagnetics. Certainly we

must agree that a theory is "useful" as long as it is defined as such by the scientists and engineers who employ it. In practice a theory is considered useful if it predicts accurately the behavior of nature under given circumstances, and even a theory that often fails may be useful if it is the best available. We choose here to take a more narrow view and investigate whether the theory is "well-posed."

A mathematical model for a physical problem is said to be *well-posed*, or *correctly set*, if three conditions hold:

1. the model has at least one solution (*existence*);
2. the model has at most one solution (*uniqueness*);
3. the solution is continuously dependent on the data supplied.

The importance of the first condition is obvious: if the electromagnetic model has no solution, it will be of little use to scientists and engineers. The importance of the second condition is equally obvious: if we apply two different solution methods to the same model and get two different answers, the model will not be very helpful in analysis or design work. The third point is more subtle; it is often extended in a practical sense to the following statement:

3′. Small changes in the data supplied produce equally small changes in the solution.

That is, the solution is not sensitive to errors in the data. To make sense of this we must decide which quantity is specified (the independent quantity) and which remains to be calculated (the dependent quantity). Commonly the source field (charge) is taken as the independent quantity, and the mediating (electromagnetic) field is computed from it; in such cases it can be shown that Maxwell's equations are well-posed. Taking the electromagnetic field to be the independent quantity, we can produce situations in which the computed quantity (charge or current) changes wildly with small changes in the specified fields. These situations (called *inverse problems*) are of great importance in remote sensing, where the field is measured and the properties of the object probed are thereby deduced.

At this point we shall concentrate on the "forward" problem of specifying the source field (charge) and computing the mediating field (the electromagnetic field). In this case we may question whether the first of the three conditions (existence) holds. We have twelve unknown quantities (the scalar components of the four vector fields), but only eight equations to describe them (from the scalar components of the two fundamental Maxwell equations and the two scalar auxiliary equations). With fewer equations than unknowns we cannot be sure that a solution exists, and we refer to Maxwell's equations as being *indefinite*. To overcome this problem we must specify more information in the form of constitutive relations among the field quantities $\mathbf{E}$, $\mathbf{B}$, $\mathbf{D}$, $\mathbf{H}$, and $\mathbf{J}$. When these are properly formulated, the number of unknowns and the number of equations are equal and Maxwell's equations are in *definite form*. If we provide more equations than unknowns, the solution may be non-unique. When we model the electromagnetic properties of materials we must supply precisely the right amount of information in the constitutive relations, or our postulate will not be well-posed.

Once Maxwell's equations are in definite form, standard methods for partial differential equations can be used to determine whether the electromagnetic model is well-posed. In a nutshell, the system (2.1)–(2.2) of hyperbolic differential equations is well-posed if and only if we specify $\mathbf{E}$ and $\mathbf{H}$ throughout a volume region V at some time instant and also specify, at all subsequent times,

1. the tangential component of $\mathbf{E}$ over all of the boundary surface S, or
2. the tangential component of $\mathbf{H}$ over all of S, or
3. the tangential component of $\mathbf{E}$ over part of S, and the tangential component of $\mathbf{H}$ over the remainder of S.

Proof of all three of the conditions of well-posedness is quite tedious, but a simplified uniqueness proof is often given in textbooks on electromagnetics. The procedure used by Stratton [217] is reproduced below. The interested reader should refer to Hansen [92] for a discussion of the existence of solutions to Maxwell's equations.

2.2.1 Uniqueness of solutions to Maxwell's equations

Consider a simply connected region of space V bounded by a surface S, where both V and S contain only ordinary points. The fields within V are associated with a current distribution $\mathbf{J}$, which may be internal to V (entirely or in part). By the initial conditions that imply the auxiliary Maxwell's equations, we know there is a time, say $t = 0$, prior to which the current is zero for all time, and thus by causality the fields throughout V are identically zero for all times $t < 0$. We next assume that the fields are specified throughout V at some time $t_0 > 0$, and seek conditions under which they are determined uniquely for all $t > t_0$.

Let the field set $(\mathbf{E}_1, \mathbf{D}_1, \mathbf{B}_1, \mathbf{H}_1)$ be a solution to Maxwell's equations (2.1)–(2.2) associated with the current $\mathbf{J}$ (along with an appropriate set of constitutive relations), and let $(\mathbf{E}_2, \mathbf{D}_2, \mathbf{B}_2, \mathbf{H}_2)$ be a second solution associated with $\mathbf{J}$. To determine the conditions for uniqueness of the fields, we look for a situation that results in $\mathbf{E}_1 = \mathbf{E}_2$, $\mathbf{B}_1 = \mathbf{B}_2$, and so on. The electromagnetic fields must obey

$$\begin{aligned}
\nabla \times \mathbf{E}_1 &= -\frac{\partial \mathbf{B}_1}{\partial t}, \\
\nabla \times \mathbf{H}_1 &= \mathbf{J} + \frac{\partial \mathbf{D}_1}{\partial t}, \\
\nabla \times \mathbf{E}_2 &= -\frac{\partial \mathbf{B}_2}{\partial t}, \\
\nabla \times \mathbf{H}_2 &= \mathbf{J} + \frac{\partial \mathbf{D}_2}{\partial t}.
\end{aligned}$$

Subtracting, we have

$$\nabla \times (\mathbf{E}_1 - \mathbf{E}_2) = -\frac{\partial (\mathbf{B}_1 - \mathbf{B}_2)}{\partial t}, \tag{2.9}$$

$$\nabla \times (\mathbf{H}_1 - \mathbf{H}_2) = \frac{\partial (\mathbf{D}_1 - \mathbf{D}_2)}{\partial t}, \tag{2.10}$$

hence defining $\mathbf{E}_0 = \mathbf{E}_1 - \mathbf{E}_2$, $\mathbf{B}_0 = \mathbf{B}_1 - \mathbf{B}_2$, and so on, we have

$$\mathbf{E}_0 \cdot (\nabla \times \mathbf{H}_0) = \mathbf{E}_0 \cdot \frac{\partial \mathbf{D}_0}{\partial t}, \tag{2.11}$$

$$\mathbf{H}_0 \cdot (\nabla \times \mathbf{E}_0) = -\mathbf{H}_0 \cdot \frac{\partial \mathbf{B}_0}{\partial t}. \tag{2.12}$$

Subtracting again, we have

$$\mathbf{E}_0 \cdot (\nabla \times \mathbf{H}_0) - \mathbf{H}_0 \cdot (\nabla \times \mathbf{E}_0) = \mathbf{H}_0 \cdot \frac{\partial \mathbf{B}_0}{\partial t} + \mathbf{E}_0 \cdot \frac{\partial \mathbf{D}_0}{\partial t},$$

hence

$$-\nabla \cdot (\mathbf{E}_0 \times \mathbf{H}_0) = \mathbf{E}_0 \cdot \frac{\partial \mathbf{D}_0}{\partial t} + \mathbf{H}_0 \cdot \frac{\partial \mathbf{B}_0}{\partial t}$$

by (B.44). Integrating both sides throughout V and using the divergence theorem on the left-hand side, we get

$$-\oint_S (\mathbf{E}_0 \times \mathbf{H}_0) \cdot \mathbf{dS} = \int_V \left(\mathbf{E}_0 \cdot \frac{\partial \mathbf{D}_0}{\partial t} + \mathbf{H}_0 \cdot \frac{\partial \mathbf{B}_0}{\partial t} \right) dV.$$

Breaking S into two arbitrary portions and using (B.6), we obtain

$$\int_{S_1} \mathbf{E}_0 \cdot (\hat{\mathbf{n}} \times \mathbf{H}_0)\, dS - \int_{S_2} \mathbf{H}_0 \cdot (\hat{\mathbf{n}} \times \mathbf{E}_0)\, dS = \int_V \left(\mathbf{E}_0 \cdot \frac{\partial \mathbf{D}_0}{\partial t} + \mathbf{H}_0 \cdot \frac{\partial \mathbf{B}_0}{\partial t} \right) dV.$$

Now if $\hat{\mathbf{n}} \times \mathbf{E}_0 = 0$ or $\hat{\mathbf{n}} \times \mathbf{H}_0 = 0$ over all of S, or some combination of these conditions holds over all of S, then

$$\int_V \left(\mathbf{E}_0 \cdot \frac{\partial \mathbf{D}_0}{\partial t} + \mathbf{H}_0 \cdot \frac{\partial \mathbf{B}_0}{\partial t} \right) dV = 0. \tag{2.13}$$

This expression implies a relationship between $\mathbf{E}_0$, $\mathbf{D}_0$, $\mathbf{B}_0$, and $\mathbf{H}_0$. Since V is arbitrary, we see that one possibility is simply to have $\mathbf{D}_0$ and $\mathbf{B}_0$ constant with time. However, since the fields are identically zero for $t < 0$, if they are constant for all time then those constant values must be zero. Another possibility is to have one of each pair $(\mathbf{E}_0, \mathbf{D}_0)$ and $(\mathbf{H}_0, \mathbf{B}_0)$ equal to zero. Then, by (2.9) and (2.10), $\mathbf{E}_0 = 0$ implies $\mathbf{B}_0 = 0$, and $\mathbf{D}_0 = 0$ implies $\mathbf{H}_0 = 0$. Thus $\mathbf{E}_1 = \mathbf{E}_2$, $\mathbf{B}_1 = \mathbf{B}_2$, and so on, and the solution is unique throughout V. However, we cannot in general rule out more complicated relationships. The number of possibilities depends on the additional constraints on the relationship between $\mathbf{E}_0$, $\mathbf{D}_0$, $\mathbf{B}_0$, and $\mathbf{H}_0$ that we must supply to describe the material supporting the field — i.e., the constitutive relationships. For a simple medium described by the time-constant permittivity ϵ and permeability μ, (2.13) becomes

$$\int_V \left(\mathbf{E}_0 \cdot \epsilon \frac{\partial \mathbf{E}_0}{\partial t} + \mathbf{H}_0 \cdot \mu \frac{\partial \mathbf{H}_0}{\partial t} \right) dV = 0,$$

or

$$\frac{1}{2} \frac{\partial}{\partial t} \int_V (\epsilon \mathbf{E}_0 \cdot \mathbf{E}_0 + \mu \mathbf{H}_0 \cdot \mathbf{H}_0)\, dV = 0.$$

Since the integrand is always positive or zero (and not constant with time, as mentioned above), the only possible conclusion is that $\mathbf{E}_0$ and $\mathbf{H}_0$ must both be zero, and thus the fields are unique.

When establishing more complicated constitutive relations, we must be careful to ensure that they lead to a unique solution, and that the condition for uniqueness is understood. In the case above, the assumption $\hat{\mathbf{n}} \times \mathbf{E}_0\big|_S = 0$ implies that the tangential components of $\mathbf{E}_1$ and $\mathbf{E}_2$ are identical over S — that is, we must give specific values of these quantities on S to ensure uniqueness. A similar statement holds for the condition $\hat{\mathbf{n}} \times \mathbf{H}_0\big|_S = 0$. Requiring that constitutive relations lead to a unique solution is known as *just setting*, and is one of several factors that must be considered, as discussed in the next section.

Uniqueness implies that the electromagnetic state of an isolated region of space may be determined without the knowledge of conditions outside the region. If we wish to solve Maxwell's equations for that region, we need know only the source density within

the region and the values of the tangential fields over the bounding surface. The effects of a complicated external world are thus reduced to the specification of surface fields. This concept has numerous applications to problems in antennas, diffraction, and guided waves.

2.2.2 Constitutive relations

We must supply a set of constitutive relations to complete the conditions for well-posedness. We generally split these relations into two sets. The first describes the relationships between the electromagnetic field quantities, and the second describes mechanical interaction between the fields and resulting secondary sources. All of these relations depend on the properties of the medium supporting the electromagnetic field. Material phenomena are quite diverse, and it is remarkable that the Maxwell–Minkowski equations hold for all phenomena yet discovered. All material effects, from nonlinearity to chirality to temporal dispersion, are described by the constitutive relations.

The specification of constitutive relationships is required in many areas of physical science to describe the behavior of "ideal materials": mathematical models of actual materials encountered in nature. For instance, in continuum mechanics the constitutive equations describe the relationship between material motions and stress tensors [245]. Truesdell and Toupin [234] give an interesting set of "guiding principles" for the concerned scientist to use when constructing constitutive relations. These include consideration of *consistency* (with the basic conservation laws of nature), *coordinate invariance* (independence of coordinate system), *isotropy and aeolotropy* (dependence on, or independence of, orientation), *just setting* (constitutive parameters should lead to a unique solution), *dimensional invariance* (similarity), *material indifference* (non-dependence on the observer), and *equipresence* (inclusion of *all* relevant physical phenomena in *all* of the constitutive relations across disciplines).

The constitutive relations generally involve a set of constitutive parameters and a set of constitutive operators. The constitutive parameters may be as simple as constants of proportionality between the fields or they may be components in a dyadic relationship. The constitutive operators may be linear and integro-differential in nature, or may imply some nonlinear operation on the fields. If the constitutive parameters are spatially constant within a certain region, we term the medium *homogeneous* within that region. If the constitutive parameters vary spatially, the medium is *inhomogeneous*. If the constitutive parameters are constants with time, we term the medium *stationary*; if they are time-changing, the medium is *nonstationary*. If the constitutive operators involve time derivatives or integrals, the medium is said to be *temporally dispersive*; if space derivatives or integrals are involved, the medium is *spatially dispersive*. Examples of all these effects can be found in common materials. It is important to note that the constitutive parameters may depend on other physical properties of the material, such as temperature, mechanical stress, and isomeric state, just as the mechanical constitutive parameters of a material may depend on the electromagnetic properties (principle of equipresence).

Many effects produced by linear constitutive operators, such as those associated with temporal dispersion, have been studied primarily in the frequency domain. In this case temporal derivative and integral operations produce complex constitutive parameters. It is becoming equally important to characterize these effects directly in the time domain for use with direct time-domain field solving techniques such as the finite-difference time-domain (FDTD) method. We shall cover the very basic properties of dispersive media in this section. A detailed description of frequency-domain fields (and a discussion of

complex constitutive parameters) is deferred until later in this book.

It is difficult to find a simple and consistent means for classifying materials by their electromagnetic effects. One way is to separate linear and nonlinear materials, then categorize linear materials by the way in which the fields are coupled through the constitutive relations:

1. *Isotropic* materials are those in which **D** is related to **E**, **B** is related to **H**, and the secondary source current **J** is related to **E**, with the field direction in each pair aligned.

2. In *anisotropic* materials the pairings are the same, but the fields in each pair are generally not aligned.

3. In *biisotropic* materials (such as chiral media) the fields **D** and **B** depend on both **E** and **H**, but with no realignment of **E** or **H**; for instance, **D** is given by the addition of a scalar times **E** plus a second scalar times **H**. Thus the contributions to **D** involve no changes to the directions of **E** and **H**.

4. *Bianisotropic* materials exhibit the most general behavior: **D** and **H** depend on both **E** and **B**, with an arbitrary realignment of either or both of these fields.

In 1888, Roentgen showed experimentally that a material isotropic in its own stationary reference frame exhibits bianisotropic properties when observed from a moving frame. Only recently have materials bianisotropic in their own rest frame been discovered. In 1894 Curie predicted that in a stationary material, based on symmetry, an electric field might produce magnetic effects and a magnetic field might produce electric effects. These effects, coined *magnetoelectric* by Landau and Lifshitz in 1957, were sought unsuccessfully by many experimentalists during the first half of the twentieth century. In 1959 the Soviet scientist I.E. Dzyaloshinskii predicted that, theoretically, the antiferromagnetic material chromium oxide (Cr_2O_3) should display magnetoelectric effects. The magnetoelectric effect was finally observed soon after by D.N. Astrov in a single crystal of Cr_2O_3 using a 10 kHz electric field. Since then the effect has been observed in many different materials. Recently, highly exotic materials with useful electromagnetic properties have been proposed and studied in depth, including chiroplasmas and chiroferrites [247]. As the technology of materials synthesis advances, a host of new and intriguing media will certainly be created.

The most general forms of the constitutive relations between the fields may be written in symbolic form as

$$\mathbf{D} = \mathbf{D}[\mathbf{E}, \mathbf{B}], \tag{2.14}$$

$$\mathbf{H} = \mathbf{H}[\mathbf{E}, \mathbf{B}]. \tag{2.15}$$

That is, **D** and **H** have some mathematically descriptive relationship to **E** and **B**. The specific forms of the relationships may be written in terms of dyadics as [119]

$$c\mathbf{D} = \bar{\mathbf{P}} \cdot \mathbf{E} + \bar{\mathbf{L}} \cdot (c\mathbf{B}), \tag{2.16}$$

$$\mathbf{H} = \bar{\mathbf{M}} \cdot \mathbf{E} + \bar{\mathbf{Q}} \cdot (c\mathbf{B}), \tag{2.17}$$

where each of the quantities $\bar{\mathbf{P}}, \bar{\mathbf{L}}, \bar{\mathbf{M}}, \bar{\mathbf{Q}}$ may be dyadics in the usual sense, or dyadic operators containing space or time derivatives or integrals, or some nonlinear operations on the fields. We may write these expressions as a single matrix equation

$$\begin{bmatrix} c\mathbf{D} \\ \mathbf{H} \end{bmatrix} = [\bar{\mathbf{C}}] \begin{bmatrix} \mathbf{E} \\ c\mathbf{B} \end{bmatrix} \tag{2.18}$$

where the 6×6 matrix

$$[\bar{\mathbf{C}}] = \begin{bmatrix} \bar{\mathbf{P}} & \bar{\mathbf{L}} \\ \bar{\mathbf{M}} & \bar{\mathbf{Q}} \end{bmatrix}.$$

This most general relationship between fields is the property of a bianisotropic material.

We may wonder why $\mathbf{D}$ is not related to $(\mathbf{E}, \mathbf{B}, \mathbf{H})$, $\mathbf{E}$ to $(\mathbf{D}, \mathbf{B})$, etc. The reason is that since the field pairs $(\mathbf{E}, \mathbf{B})$ and $(\mathbf{D}, \mathbf{H})$ convert identically under a Lorentz transformation, a constitutive relation that maps fields as in (2.18) is form invariant, as are the Maxwell–Minkowski equations. That is, although the constitutive parameters may vary numerically between observers moving at different velocities, the form of the relationship given by (2.18) is maintained.

Many authors choose to relate $(\mathbf{D}, \mathbf{B})$ to $(\mathbf{E}, \mathbf{H})$, often because the expressions are simpler and can be more easily applied to specific problems. For instance, in a linear, isotropic material (as shown below) $\mathbf{D}$ is directly proportional to $\mathbf{E}$ and $\mathbf{B}$ is directly proportional to $\mathbf{H}$. To provide the appropriate expression for the constitutive relations, we need only remap (2.18). This gives

$$\mathbf{D} = \bar{\boldsymbol{\epsilon}} \cdot \mathbf{E} + \bar{\boldsymbol{\xi}} \cdot \mathbf{H}, \tag{2.19}$$

$$\mathbf{B} = \bar{\boldsymbol{\zeta}} \cdot \mathbf{E} + \bar{\boldsymbol{\mu}} \cdot \mathbf{H}, \tag{2.20}$$

or

$$\begin{bmatrix} \mathbf{D} \\ \mathbf{B} \end{bmatrix} = \left[\bar{\mathbf{C}}_{EH}\right] \begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix}, \tag{2.21}$$

where the new constitutive parameters $\bar{\boldsymbol{\epsilon}}, \bar{\boldsymbol{\xi}}, \bar{\boldsymbol{\zeta}}, \bar{\boldsymbol{\mu}}$ can be easily found from the original constitutive parameters $\bar{\mathbf{P}}, \bar{\mathbf{L}}, \bar{\mathbf{M}}, \bar{\mathbf{Q}}$. We do note, however, that in the form (2.19)–(2.20) the Lorentz invariance of the constitutive equations is not obvious.

In the following paragraphs we shall characterize some of the most common materials according to these classifications. With this approach effects such as temporal or spatial dispersion are not part of the classification process, but arise from the nature of the constitutive parameters. Hence we shall not dwell on the particulars of the constitutive parameters, but shall concentrate on the form of the constitutive relations.

Constitutive relations for fields in free space. In a vacuum the fields are related by the simple constitutive equations

$$\mathbf{D} = \epsilon_0 \mathbf{E}, \tag{2.22}$$

$$\mathbf{H} = \frac{1}{\mu_0} \mathbf{B}. \tag{2.23}$$

The quantities μ_0 and ϵ_0 are, respectively, the *free-space permeability* and *permittivity constants*. It is convenient to use three numerical quantities to describe the electromagnetic properties of free space — μ_0, ϵ_0, and the speed of light c — and interrelate them through the equation

$$c = 1/(\mu_0 \epsilon_0)^{1/2}.$$

Historically it has been the practice to define μ_0, measure c, and compute ϵ_0. In SI units

$$\begin{aligned} \mu_0 &= 4\pi \times 10^{-7} \text{ H/m}, \\ c &= 2.998 \times 10^{8} \text{ m/s}, \\ \epsilon_0 &= 8.854 \times 10^{-12} \text{ F/m}. \end{aligned}$$

With the two constitutive equations we have enough information to put Maxwell's equations into definite form. Traditionally (2.22) and (2.23) are substituted into (2.1)–(2.2) to give

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.24}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{2.25}$$

These are two vector equations in two vector unknowns (equivalently, six scalar equations in six scalar unknowns).

In terms of the general constitutive relation (2.18), we find that free space is isotropic with

$$\bar{\mathbf{P}} = \bar{\mathbf{Q}} = \frac{1}{\eta_0}\bar{\mathbf{I}}, \qquad \bar{\mathbf{L}} = \bar{\mathbf{M}} = 0,$$

where $\eta_0 = (\mu_0/\epsilon_0)^{1/2}$ is called the *intrinsic impedance of free space*. This emphasizes the fact that free space has, along with c, only a single empirical constant associated with it (i.e., ϵ_0 or η_0). Since no derivative or integral operators appear in the constitutive relations, free space is nondispersive.

Constitutive relations in a linear isotropic material. In a linear isotropic material there is proportionality between $\mathbf{D}$ and $\mathbf{E}$ and between $\mathbf{B}$ and $\mathbf{H}$. The constants of proportionality are the *permittivity* ϵ and the *permeability* μ. If the material is nondispersive, the constitutive relations take the form

$$\mathbf{D} = \epsilon \mathbf{E}, \qquad \mathbf{B} = \mu \mathbf{H},$$

where ϵ and μ may depend on position for inhomogeneous materials. Often the permittivity and permeability are referenced to the permittivity and permeability of free space according to

$$\epsilon = \epsilon_r \epsilon_0, \qquad \mu = \mu_r \mu_0.$$

Here the dimensionless quantities ϵ_r and μ_r are called, respectively, the *relative permittivity* and *relative permeability*.

When dealing with the Maxwell–Boffi equations (§ 2.4) the difference between the material and free space values of $\mathbf{D}$ and $\mathbf{H}$ becomes important. Thus for linear isotropic materials we often write the constitutive relations as

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \epsilon_0 \chi_e \mathbf{E}, \tag{2.26}$$

$$\mathbf{B} = \mu_0 \mathbf{H} + \mu_0 \chi_m \mathbf{H}, \tag{2.27}$$

where the dimensionless quantities $\chi_e = \epsilon_r - 1$ and $\chi_m = \mu_r - 1$ are called, respectively, the *electric* and *magnetic susceptibilities* of the material. In terms of (2.18) we have

$$\bar{\mathbf{P}} = \frac{\epsilon_r}{\eta_0}\bar{\mathbf{I}}, \qquad \bar{\mathbf{Q}} = \frac{1}{\eta_0 \mu_r}\bar{\mathbf{I}}, \qquad \bar{\mathbf{L}} = \bar{\mathbf{M}} = 0.$$

Generally a material will have either its electric or magnetic properties dominant. If $\mu_r = 1$ and $\epsilon_r \neq 1$ then the material is generally called a *perfect dielectric* or a *perfect insulator*, and is said to be an electric material. If $\epsilon_r = 1$ and $\mu_r \neq 1$, the material is said to be a magnetic material.

A linear isotropic material may also have conduction properties. In a *conducting material*, a constitutive relation is generally used to describe the mechanical interaction of field and charge by relating the electric field to a secondary electric current. For a nondispersive isotropic material, the current is aligned with, and proportional to, the electric field; there are no temporal operators in the constitutive relation, which is simply

$$\mathbf{J} = \sigma \mathbf{E}. \tag{2.28}$$

This is known as *Ohm's law*. Here σ is the *conductivity* of the material.

If $\mu_r \approx 1$ and σ is very small, the material is generally called a *good dielectric*. If σ is very large, the material is generally called a *good conductor*. The conditions by which we say the conductivity is "small" or "large" are usually established using the frequency response of the material. Materials that are good dielectrics over broad ranges of frequency include various glasses and plastics such as fused quartz, polyethylene, and teflon. Materials that are good conductors over broad ranges of frequency include common metals such as gold, silver, and copper.

For dispersive linear isotropic materials, the constitutive parameters become nonstationary (time dependent), and the constitutive relations involve time operators. (Note that the name *dispersive* describes the tendency for pulsed electromagnetic waves to spread out, or disperse, in materials of this type.) If we assume that the relationships given by (2.26), (2.27), and (2.28) retain their product form in the frequency domain, then by the convolution theorem we have in the time domain the constitutive relations

$$\mathbf{D}(\mathbf{r}, t) = \epsilon_0 \left(\mathbf{E}(\mathbf{r}, t) + \int_{-\infty}^{t} \chi_e(\mathbf{r}, t - t') \mathbf{E}(\mathbf{r}, t')\, dt' \right), \tag{2.29}$$

$$\mathbf{B}(\mathbf{r}, t) = \mu_0 \left(\mathbf{H}(\mathbf{r}, t) + \int_{-\infty}^{t} \chi_m(\mathbf{r}, t - t') \mathbf{H}(\mathbf{r}, t')\, dt' \right), \tag{2.30}$$

$$\mathbf{J}(\mathbf{r}, t) = \int_{-\infty}^{t} \sigma(\mathbf{r}, t - t') \mathbf{E}(\mathbf{r}, t')\, dt'. \tag{2.31}$$

These expressions were first introduced by Volterra in 1912 [234]. We see that for a linear dispersive material of this type the constitutive operators are time integrals, and that the behavior of $\mathbf{D}(t)$ depends not only on the value of $\mathbf{E}$ at time t, but on its values at all past times. Thus, in dispersive materials there is a "time lag" between the effect of the applied field and the polarization or magnetization that results. In the frequency domain, temporal dispersion is associated with complex values of the constitutive parameters, which, to describe a causal relationship, cannot be constant with frequency. The nonzero imaginary component is identified with the dissipation of electromagnetic energy as heat. Causality is implied by the upper limit being t in the convolution integrals, which indicates that $\mathbf{D}(t)$ cannot depend on future values of $\mathbf{E}(t)$. This assumption leads to a relationship between the real and imaginary parts of the frequency domain constitutive parameters as described through the Kronig–Kramers equations.

Constitutive relations for fields in perfect conductors. In a perfect electric conductor (PEC) or a perfect magnetic conductor (PMC) the fields are exactly specified as the null field:

$$\mathbf{E} = \mathbf{D} = \mathbf{B} = \mathbf{H} = 0.$$

By Ampere's and Faraday's laws we must also have $\mathbf{J} = \mathbf{J}_m = 0$; hence, by the continuity equation, $\rho = \rho_m = 0$.

In addition to the null field, we have the condition that the tangential electric field on the surface of a PEC must be zero. Similarly, the tangential magnetic field on the surface of a PMC must be zero. This implies (§ 2.8.3) that an electric surface current may exist on the surface of a PEC but not on the surface of a PMC, while a magnetic surface current may exist on the surface of a PMC but not on the surface of a PEC.

A PEC may be regarded as the limit of a conducting material as $\sigma \to \infty$. In many practical cases, good conductors such as gold and copper can be assumed to be perfect electric conductors, which greatly simplifies the application of boundary conditions. No physical material is known to behave as a PMC, but the concept is mathematically useful for applying symmetry conditions (in which a PMC is sometimes referred to as a "magnetic wall") and for use in developing equivalence theorems.

Constitutive relations in a linear anisotropic material. In a linear anisotropic material there are relationships between $\mathbf{B}$ and $\mathbf{H}$ and between $\mathbf{D}$ and $\mathbf{E}$, but the field vectors are not aligned as in the isotropic case. We can thus write

$$\mathbf{D} = \bar{\boldsymbol{\epsilon}} \cdot \mathbf{E}, \qquad \mathbf{B} = \bar{\boldsymbol{\mu}} \cdot \mathbf{H}, \qquad \mathbf{J} = \bar{\boldsymbol{\sigma}} \cdot \mathbf{E},$$

where $\bar{\boldsymbol{\epsilon}}$ is called the *permittivity dyadic*, $\bar{\boldsymbol{\mu}}$ is the *permeability dyadic*, and $\bar{\boldsymbol{\sigma}}$ is the *conductivity dyadic*. In terms of the general constitutive relation (2.18) we have

$$\bar{\mathbf{P}} = c\bar{\boldsymbol{\epsilon}}, \qquad \bar{\mathbf{Q}} = \frac{\bar{\boldsymbol{\mu}}^{-1}}{c}, \qquad \bar{\mathbf{L}} = \bar{\mathbf{M}} = 0.$$

Many different types of materials demonstrate anisotropic behavior, including optical crystals, magnetized plasmas, and ferrites. Plasmas and ferrites are examples of *gyrotropic* media. With the proper choice of coordinate system, the frequency-domain permittivity or permeability can be written in matrix form as

$$[\tilde{\bar{\boldsymbol{\epsilon}}}] = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & 0 \\ -\epsilon_{12} & \epsilon_{11} & 0 \\ 0 & 0 & \epsilon_{33} \end{bmatrix}, \qquad [\tilde{\bar{\boldsymbol{\mu}}}] = \begin{bmatrix} \mu_{11} & \mu_{12} & 0 \\ -\mu_{12} & \mu_{11} & 0 \\ 0 & 0 & \mu_{33} \end{bmatrix}. \tag{2.32}$$

Each of the matrix entries may be complex. For the special case of a *lossless* gyrotropic material, the matrices become *hermitian*:

$$[\tilde{\bar{\boldsymbol{\epsilon}}}] = \begin{bmatrix} \epsilon & -j\delta & 0 \\ j\delta & \epsilon & 0 \\ 0 & 0 & \epsilon_3 \end{bmatrix}, \qquad [\tilde{\bar{\boldsymbol{\mu}}}] = \begin{bmatrix} \mu & -j\kappa & 0 \\ j\kappa & \mu & 0 \\ 0 & 0 & \mu_3 \end{bmatrix}, \tag{2.33}$$

where ϵ, ϵ_3, δ, μ, μ_3, and κ are real numbers.

Crystals have received particular attention because of their birefringent properties. A birefringent crystal can be characterized by a symmetric permittivity dyadic that has real permittivity parameters in the frequency domain; equivalently, the constitutive relations do not involve constitutive operators. A coordinate system called the *principal system*, with axes called the *principal axes*, can always be found so that the permittivity dyadic in that system is diagonal:

$$[\tilde{\bar{\boldsymbol{\epsilon}}}] = \begin{bmatrix} \epsilon_x & 0 & 0 \\ 0 & \epsilon_y & 0 \\ 0 & 0 & \epsilon_z \end{bmatrix}.$$

The geometrical structure of a crystal determines the relationship between ϵ_x, ϵ_y, and ϵ_z. If $\epsilon_x = \epsilon_y < \epsilon_z$, then the crystal is *positive uniaxial* (e.g., quartz). If $\epsilon_x = \epsilon_y > \epsilon_z$,

the crystal is *negative uniaxial* (e.g., calcite). If $\epsilon_x \neq \epsilon_y \neq \epsilon_z$, the crystal is *biaxial* (e.g., mica). In uniaxial crystals the z-axis is called the *optical axis*.

If the anisotropic material is dispersive, we can generalize the convolutional form of the isotropic dispersive media to obtain the constitutive relations

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0 \left(\mathbf{E}(\mathbf{r},t) + \int_{-\infty}^{t} \bar{\boldsymbol{\chi}}_e(\mathbf{r},t-t') \cdot \mathbf{E}(\mathbf{r},t')\, dt' \right), \tag{2.34}$$

$$\mathbf{B}(\mathbf{r},t) = \mu_0 \left(\mathbf{H}(\mathbf{r},t) + \int_{-\infty}^{t} \bar{\boldsymbol{\chi}}_m(\mathbf{r},t-t') \cdot \mathbf{H}(\mathbf{r},t')\, dt' \right), \tag{2.35}$$

$$\mathbf{J}(\mathbf{r},t) = \int_{-\infty}^{t} \bar{\boldsymbol{\sigma}}(\mathbf{r},t-t') \cdot \mathbf{E}(\mathbf{r},t')\, dt'. \tag{2.36}$$

Constitutive relations for biisotropic materials. A biisotropic material is an isotropic magnetoelectric material. Here we have $\mathbf{D}$ related to $\mathbf{E}$ and $\mathbf{B}$, and $\mathbf{H}$ related to $\mathbf{E}$ and $\mathbf{B}$, but with no realignment of the fields as in anisotropic (or bianisotropic) materials. Perhaps the simplest example is the *Tellegen medium* devised by B.D.H. Tellegen in 1948 [228], having

$$\mathbf{D} = \epsilon \mathbf{E} + \xi \mathbf{H}, \tag{2.37}$$

$$\mathbf{B} = \xi \mathbf{E} + \mu \mathbf{H}. \tag{2.38}$$

Tellegen proposed that his hypothetical material be composed of small (but macroscopic) ferromagnetic particles suspended in a liquid. This is an example of a *synthetic* material, constructed from ordinary materials to have an exotic electromagnetic behavior. Other examples include artificial dielectrics made from metallic particles imbedded in lightweight foams [74], and *chiral materials* made from small metallic helices suspended in resins [130].

Chiral materials are also biisotropic, and have the constitutive relations

$$\mathbf{D} = \epsilon \mathbf{E} - \chi \frac{\partial \mathbf{H}}{\partial t}, \tag{2.39}$$

$$\mathbf{B} = \mu \mathbf{H} + \chi \frac{\partial \mathbf{E}}{\partial t}, \tag{2.40}$$

where the constitutive parameter χ is called the *chirality parameter*. Note the presence of temporal derivative operators. Alternatively,

$$\mathbf{D} = \epsilon(\mathbf{E} + \beta \nabla \times \mathbf{E}), \tag{2.41}$$

$$\mathbf{B} = \mu(\mathbf{H} + \beta \nabla \times \mathbf{H}), \tag{2.42}$$

by Faraday's and Ampere's laws. Chirality is a natural state of symmetry; many natural substances are chiral materials, including DNA and many sugars. The time derivatives in (2.39)–(2.40) produce rotation of the polarization of time harmonic electromagnetic waves propagating in chiral media.

Constitutive relations in nonlinear media. Nonlinear electromagnetic effects have been studied by scientists and engineers since the beginning of the era of electrical technology. Familiar examples include saturation and hysteresis in ferromagnetic materials and the behavior of p-n junctions in solid-state rectifiers. The invention of the laser extended interest in nonlinear effects to the realm of optics, where phenomena such as

parametric amplification and oscillation, harmonic generation, and magneto-optic interactions have found applications in modern devices [202].

Provided that the external field applied to a nonlinear electric material is small compared to the internal molecular fields, the relationship between $\mathbf{E}$ and $\mathbf{D}$ can be expanded in a Taylor series of the electric field. For an anisotropic material exhibiting no hysteresis effects, the constitutive relation is [150]

$$D_i(\mathbf{r},t) = \epsilon_0 E_i(\mathbf{r},t) + \sum_{j=1}^{3} \chi_{ij}^{(1)} E_j(\mathbf{r},t) + \sum_{j,k=1}^{3} \chi_{ijk}^{(2)} E_j(\mathbf{r},t) E_k(\mathbf{r},t) + \\ + \sum_{j,k,l=1}^{3} \chi_{ijkl}^{(3)} E_j(\mathbf{r},t) E_k(\mathbf{r},t) E_l(\mathbf{r},t) + \cdots \tag{2.43}$$

where the index $i = 1, 2, 3$ refers to the three components of the fields $\mathbf{D}$ and $\mathbf{E}$. The first sum in (2.43) is identical to the constitutive relation for linear anisotropic materials. Thus, $\chi_{ij}^{(1)}$ is identical to the susceptibility dyadic of a linear anisotropic medium considered earlier. The quantity $\chi_{ijk}^{(2)}$ is called the *second-order susceptibility*, and is a three-dimensional matrix (or third rank tensor) describing the nonlinear electric effects quadratic in $\mathbf{E}$. Similarly $\chi_{ijkl}^{(3)}$ is called the *third-order susceptibility*, and is a four-dimensional matrix (or fourth rank tensor) describing the nonlinear electric effects cubic in $\mathbf{E}$. Numerical values of $\chi_{ijk}^{(2)}$ and $\chi_{ijkl}^{(3)}$ are given in Shen [202] for a variety of crystals.

When the material shows hysteresis effects, $\mathbf{D}$ at any point $\mathbf{r}$ and time t is due not only to the value of $\mathbf{E}$ at that point and at that time, but to the values of $\mathbf{E}$ at all points and times. That is, the material displays both temporal and spatial dispersion.

2.3 Maxwell's equations in moving frames

The essence of special relativity is that the mathematical forms of Maxwell's equations are identical in all *inertial reference frames*: frames moving with uniform velocities relative to the *laboratory frame of reference* in which we perform our measurements. This *form invariance* of Maxwell's equations is a specific example of the general physical *principle of covariance*. In the laboratory frame we write the differential equations of Maxwell's theory as

$$\begin{aligned} \nabla \times \mathbf{E}(\mathbf{r},t) &= -\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t}, \\ \nabla \times \mathbf{H}(\mathbf{r},t) &= \mathbf{J}(\mathbf{r},t) + \frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t}, \\ \nabla \cdot \mathbf{D}(\mathbf{r},t) &= \rho(\mathbf{r},t), \\ \nabla \cdot \mathbf{B}(\mathbf{r},t) &= 0, \\ \nabla \cdot \mathbf{J}(\mathbf{r},t) &= -\frac{\partial \rho(\mathbf{r},t)}{\partial t}. \end{aligned}$$

Similarly, in an inertial frame having four-dimensional coordinates $(\mathbf{r}', t')$ we have

$$\nabla' \times \mathbf{E}'(\mathbf{r}',t') = -\frac{\partial \mathbf{B}'(\mathbf{r}',t')}{\partial t'},$$

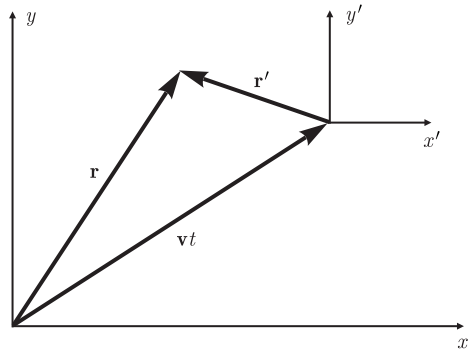

FIGURE 2.1
Primed coordinate system moving with velocity **v** relative to laboratory (unprimed) coordinate system.

$$\begin{aligned}
\nabla' \times \mathbf{H}'(\mathbf{r}', t') &= \mathbf{J}'(\mathbf{r}', t') + \frac{\partial \mathbf{D}'(\mathbf{r}', t')}{\partial t'}, \\
\nabla' \cdot \mathbf{D}'(\mathbf{r}', t') &= \rho'(\mathbf{r}', t'), \\
\nabla' \cdot \mathbf{B}'(\mathbf{r}', t') &= 0, \\
\nabla' \cdot \mathbf{J}'(\mathbf{r}', t') &= -\frac{\partial \rho'(\mathbf{r}', t')}{\partial t'}.
\end{aligned}$$

The primed fields measured in the moving system do *not* have the same numerical values as the unprimed fields measured in the laboratory. To convert between **E** and **E′**, **B** and **B′**, and so on, we must find a way to convert between the coordinates $(\mathbf{r}, t)$ and $(\mathbf{r}', t')$.

2.3.1 Field conversions under Galilean transformation

We shall assume that the primed coordinate system moves with constant velocity **v** relative to the laboratory frame (Figure 2.1). Prior to the early part of the twentieth century, converting between the primed and unprimed coordinate variables was intuitive and obvious: it was thought that time must be measured identically in each coordinate system, and that the relationship between the space variables can be determined simply by the displacement of the moving system at time $t = t'$. Under these assumptions, and under the further assumption that the two systems coincide at time $t = 0$, we can write

$$t' = t, \qquad x' = x - v_x t, \qquad y' = y - v_y t, \qquad z' = z - v_z t,$$

or simply

$$t' = t, \qquad \mathbf{r}' = \mathbf{r} - \mathbf{v}t.$$

This is called a *Galilean transformation*. We can use the chain rule to describe the manner in which differential operations transform, i.e., to relate derivatives with respect to the laboratory coordinates to derivatives with respect to the inertial coordinates. We have, for instance,

$$\begin{aligned}
\frac{\partial}{\partial t} &= \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'} + \frac{\partial x'}{\partial t}\frac{\partial}{\partial x'} + \frac{\partial y'}{\partial t}\frac{\partial}{\partial y'} + \frac{\partial z'}{\partial t}\frac{\partial}{\partial z'} \\
&= \frac{\partial}{\partial t'} - v_x\frac{\partial}{\partial x'} - v_y\frac{\partial}{\partial y'} - v_z\frac{\partial}{\partial z'}
\end{aligned}$$

$$= \frac{\partial}{\partial t'} - (\mathbf{v} \cdot \nabla'). \tag{2.44}$$

Similarly

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial x'}, \qquad \frac{\partial}{\partial y} = \frac{\partial}{\partial y'}, \qquad \frac{\partial}{\partial z} = \frac{\partial}{\partial z'},$$

from which

$$\nabla \times \mathbf{A}(\mathbf{r}, t) = \nabla' \times \mathbf{A}(\mathbf{r}, t), \qquad \nabla \cdot \mathbf{A}(\mathbf{r}, t) = \nabla' \cdot \mathbf{A}(\mathbf{r}, t), \tag{2.45}$$

for each vector field $\mathbf{A}$.

Newton was aware that the laws of mechanics are invariant with respect to Galilean transformations. Do Maxwell's equations also behave in this way? Let us use the Galilean transformation to determine which relationship between the primed and unprimed fields results in form invariance of Maxwell's equations. We first examine $\nabla' \times \mathbf{E}$, the spatial rate of change of the laboratory field with respect to the inertial frame spatial coordinates:

$$\nabla' \times \mathbf{E} = \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\frac{\partial \mathbf{B}}{\partial t'} + (\mathbf{v} \cdot \nabla')\mathbf{B}$$

by (2.45) and (2.44). Rewriting the last term by (B.45) we have

$$(\mathbf{v} \cdot \nabla')\mathbf{B} = -\nabla' \times (\mathbf{v} \times \mathbf{B})$$

since $\mathbf{v}$ is constant and $\nabla' \cdot \mathbf{B} = \nabla \cdot \mathbf{B} = 0$, hence

$$\nabla' \times (\mathbf{E} + \mathbf{v} \times \mathbf{B}) = -\frac{\partial \mathbf{B}}{\partial t'}. \tag{2.46}$$

Similarly

$$\nabla' \times \mathbf{H} = \nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t'} + \nabla' \times (\mathbf{v} \times \mathbf{D}) - \mathbf{v}(\nabla' \cdot \mathbf{D})$$

where $\nabla' \cdot \mathbf{D} = \nabla \cdot \mathbf{D} = \rho$ so that

$$\nabla' \times (\mathbf{H} - \mathbf{v} \times \mathbf{D}) = \frac{\partial \mathbf{D}}{\partial t'} - \rho\mathbf{v} + \mathbf{J}. \tag{2.47}$$

Also

$$\nabla' \cdot \mathbf{J} = \nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} = -\frac{\partial \rho}{\partial t'} + (\mathbf{v} \cdot \nabla')\rho$$

and we may use (B.42) to write

$$(\mathbf{v} \cdot \nabla')\rho = \mathbf{v} \cdot (\nabla'\rho) = \nabla' \cdot (\rho\mathbf{v}),$$

obtaining

$$\nabla' \cdot (\mathbf{J} - \rho\mathbf{v}) = -\frac{\partial \rho}{\partial t'}. \tag{2.48}$$

Equations (2.46), (2.47), and (2.48) show that the forms of Maxwell's equations in the inertial and laboratory frames are identical provided that

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}, \tag{2.49}$$
$$\mathbf{D}' = \mathbf{D}, \tag{2.50}$$
$$\mathbf{H}' = \mathbf{H} - \mathbf{v} \times \mathbf{D}, \tag{2.51}$$
$$\mathbf{B}' = \mathbf{B}, \tag{2.52}$$
$$\mathbf{J}' = \mathbf{J} - \rho\mathbf{v}, \tag{2.53}$$
$$\rho' = \rho. \tag{2.54}$$

That is, (2.49)–(2.54) result in form invariance of Faraday's law, Ampere's law, and the continuity equation under a Galilean transformation. These equations express the fields measured by a moving observer in terms of those measured in the laboratory frame. To convert the opposite way, we need only use the principle of relativity. Neither observer can tell whether he or she is stationary — only that the other observer is moving relative to him or her. To obtain the fields in the laboratory frame we simply change the sign on $\mathbf{v}$ and swap primed with unprimed fields in (2.49)–(2.54):

$$\mathbf{E} = \mathbf{E}' - \mathbf{v} \times \mathbf{B}', \tag{2.55}$$
$$\mathbf{D} = \mathbf{D}', \tag{2.56}$$
$$\mathbf{H} = \mathbf{H}' + \mathbf{v} \times \mathbf{D}', \tag{2.57}$$
$$\mathbf{B} = \mathbf{B}', \tag{2.58}$$
$$\mathbf{J} = \mathbf{J}' + \rho'\mathbf{v}, \tag{2.59}$$
$$\rho = \rho'. \tag{2.60}$$

According to (2.53), a moving observer interprets charge stationary in the laboratory frame as an additional current moving opposite the direction of his or her motion. This seems reasonable. However, while $\mathbf{E}$ depends on both $\mathbf{E}'$ and $\mathbf{B}'$, the field $\mathbf{B}$ is unchanged under the transformation. Why should $\mathbf{B}$ have this special status? In fact, we may uncover an inconsistency among the transformations by considering free space where (2.22) and (2.23) hold: in this case (2.49) gives

$$\mathbf{D}'/\epsilon_0 = \mathbf{D}/\epsilon_0 + \mathbf{v} \times \mu_0\mathbf{H}$$

or

$$\mathbf{D}' = \mathbf{D} + \mathbf{v} \times \mathbf{H}/c^2$$

rather than (2.50). Similarly, from (2.51) we get

$$\mathbf{B}' = \mathbf{B} - \mathbf{v} \times \mathbf{E}/c^2$$

instead of (2.52). Using these, the set of transformations becomes

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}, \tag{2.61}$$
$$\mathbf{D}' = \mathbf{D} + \mathbf{v} \times \mathbf{H}/c^2, \tag{2.62}$$
$$\mathbf{H}' = \mathbf{H} - \mathbf{v} \times \mathbf{D}, \tag{2.63}$$
$$\mathbf{B}' = \mathbf{B} - \mathbf{v} \times \mathbf{E}/c^2, \tag{2.64}$$
$$\mathbf{J}' = \mathbf{J} - \rho\mathbf{v}, \tag{2.65}$$
$$\rho' = \rho. \tag{2.66}$$

These can also be written using dyadic notation as

$$\mathbf{E}' = \bar{\mathbf{I}} \cdot \mathbf{E} + \bar{\boldsymbol{\beta}} \cdot (c\mathbf{B}), \tag{2.67}$$
$$c\mathbf{B}' = -\bar{\boldsymbol{\beta}} \cdot \mathbf{E} + \bar{\mathbf{I}} \cdot (c\mathbf{B}), \tag{2.68}$$

and

$$c\mathbf{D}' = \bar{\mathbf{I}} \cdot (c\mathbf{D}) + \bar{\boldsymbol{\beta}} \cdot \mathbf{H}, \tag{2.69}$$
$$\mathbf{H}' = -\bar{\boldsymbol{\beta}} \cdot (c\mathbf{D}) + \bar{\mathbf{I}} \cdot \mathbf{H}, \tag{2.70}$$

where

$$[\bar{\boldsymbol{\beta}}] = \begin{bmatrix} 0 & -\beta_z & \beta_y \\ \beta_z & 0 & -\beta_x \\ -\beta_y & \beta_x & 0 \end{bmatrix}$$

with $\boldsymbol{\beta} = \mathbf{v}/c$. This set of equations is self-consistent among Maxwell's equations. However, the equations are not consistent with the assumption of a Galilean transformation of the coordinates, and thus Maxwell's equations are not covariant under a Galilean transformation. Maxwell's equations are only covariant under a Lorentz transformation as described in the next section. Expressions (2.61)–(2.64) turn out to be accurate to order v/c, hence are the results of a *first-order Lorentz transformation*. Only when v is an appreciable fraction of c do the field conversions resulting from the first-order Lorentz transformation differ markedly from those resulting from a Galilean transformation; those resulting from the true Lorentz transformation require even higher velocities to differ markedly from the first-order expressions. Engineering accuracy is often accomplished using the Galilean transformation. This pragmatic observation leads to quite a bit of confusion when considering the large-scale forms of Maxwell's equations, as we shall soon see.

2.3.2 Field conversions under Lorentz transformation

To find the proper transformation under which Maxwell's equations are covariant, we must discard our notion that time progresses the same in the primed and the unprimed frames. The proper transformation of coordinates that guarantees covariance of Maxwell's equations is the *Lorentz transformation*

$$ct' = \gamma ct - \gamma \boldsymbol{\beta} \cdot \mathbf{r}, \tag{2.71}$$

$$\mathbf{r}' = \bar{\boldsymbol{\alpha}} \cdot \mathbf{r} - \gamma \boldsymbol{\beta} ct, \tag{2.72}$$

where

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}, \qquad \bar{\boldsymbol{\alpha}} = \bar{\mathbf{I}} + (\gamma - 1)\frac{\boldsymbol{\beta}\boldsymbol{\beta}}{\beta^2}, \qquad \beta = |\boldsymbol{\beta}|.$$

This is obviously more complicated than the Galilean transformation; only as $\boldsymbol{\beta} \to 0$ are the Lorentz and Galilean transformations equivalent.

Not surprisingly, field conversions between inertial reference frames are more complicated with the Lorentz transformation than with the Galilean transformation. For simplicity we assume that the velocity of the moving frame has only an x-component: $\mathbf{v} = \hat{\mathbf{x}}v$. Later we can generalize this to any direction. Equations (2.71) and (2.72) become

$$x' = x + (\gamma - 1)x - \gamma vt, \tag{2.73}$$

$$y' = y, \tag{2.74}$$

$$z' = z, \tag{2.75}$$

$$ct' = \gamma ct - \gamma \frac{v}{c} x, \tag{2.76}$$

and the chain rule gives

$$\frac{\partial}{\partial x} = \gamma \frac{\partial}{\partial x'} - \gamma \frac{v}{c^2} \frac{\partial}{\partial t'}, \tag{2.77}$$

$$\frac{\partial}{\partial y} = \frac{\partial}{\partial y'}, \tag{2.78}$$

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial z'}, \tag{2.79}$$

$$\frac{\partial}{\partial t} = -\gamma v \frac{\partial}{\partial x'} + \gamma \frac{\partial}{\partial t'}. \tag{2.80}$$

We begin by examining Faraday's law in the laboratory frame. In component form we have

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -\frac{\partial B_x}{\partial t}, \tag{2.81}$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -\frac{\partial B_y}{\partial t}, \tag{2.82}$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -\frac{\partial B_z}{\partial t}. \tag{2.83}$$

These become

$$\frac{\partial E_z}{\partial y'} - \frac{\partial E_y}{\partial z'} = \gamma v \frac{\partial B_x}{\partial x'} - \gamma \frac{\partial B_x}{\partial t'}, \tag{2.84}$$

$$\frac{\partial E_x}{\partial z'} - \gamma \frac{\partial E_z}{\partial x'} + \gamma \frac{v}{c^2} \frac{\partial E_z}{\partial t'} = \gamma v \frac{\partial B_y}{\partial x'} - \gamma \frac{\partial B_y}{\partial t'}, \tag{2.85}$$

$$\gamma \frac{\partial E_y}{\partial x'} - \gamma \frac{v}{c^2} \frac{\partial E_y}{\partial t'} - \frac{\partial E_x}{\partial y'} = \gamma v \frac{\partial B_z}{\partial x'} - \gamma \frac{\partial B_z}{\partial t'}, \tag{2.86}$$

after we use (2.77)–(2.80) to convert the derivatives in the laboratory frame to derivatives with respect to the moving frame coordinates. To simplify (2.84) we consider

$$\nabla \cdot \mathbf{B} = \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0.$$

Converting the laboratory frame coordinates to the moving frame coordinates, we have

$$\gamma \frac{\partial B_x}{\partial x'} - \gamma \frac{v}{c^2} \frac{\partial B_x}{\partial t'} + \frac{\partial B_y}{\partial y'} + \frac{\partial B_z}{\partial z'} = 0$$

or

$$-\gamma v \frac{\partial B_x}{\partial x'} = -\gamma \frac{v^2}{c^2} \frac{\partial B_x}{\partial t'} + v \frac{\partial B_y}{\partial y'} + v \frac{\partial B_z}{\partial z'}.$$

Substituting this into (2.84) and rearranging (2.85) and (2.86), we obtain

$$\frac{\partial}{\partial y'} \gamma (E_z + v B_y) - \frac{\partial}{\partial z'} \gamma (E_y - v B_z) = -\frac{\partial B_x}{\partial t'},$$

$$\frac{\partial E_x}{\partial z'} - \frac{\partial}{\partial x'} \gamma (E_z + v B_y) = -\frac{\partial}{\partial t'} \gamma \left(B_y + \frac{v}{c^2} E_z \right),$$

$$\frac{\partial}{\partial x'} \gamma (E_y - v B_z) - \frac{\partial E_x}{\partial y'} = -\frac{\partial}{\partial t'} \gamma \left(B_z - \frac{v}{c^2} E_y \right).$$

Comparison with (2.81)–(2.83) shows that form invariance of Faraday's law under the Lorentz transformation requires

$$E_x' = E_x, \qquad E_y' = \gamma (E_y - v B_z), \qquad E_z' = \gamma (E_z + v B_y),$$

and

$$B_x' = B_x, \qquad B_y' = \gamma \left(B_y + \frac{v}{c^2} E_z \right), \qquad B_z' = \gamma \left(B_z - \frac{v}{c^2} E_y \right).$$

To generalize $\mathbf{v}$ to any direction, we simply note that the components of the fields parallel to the velocity direction are identical in the moving and laboratory frames, while the components perpendicular to the velocity direction convert according to a simple cross product rule. After similar analyses with Ampere's and Gauss's laws (see Problem 2.2), we find that

$$\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel}, \qquad \mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel}, \qquad \mathbf{D}'_{\parallel} = \mathbf{D}_{\parallel}, \qquad \mathbf{H}'_{\parallel} = \mathbf{H}_{\parallel},$$

$$\mathbf{E}'_{\perp} = \gamma(\mathbf{E}_{\perp} + \boldsymbol{\beta} \times c\mathbf{B}_{\perp}), \tag{2.87}$$
$$c\mathbf{B}'_{\perp} = \gamma(c\mathbf{B}_{\perp} - \boldsymbol{\beta} \times \mathbf{E}_{\perp}), \tag{2.88}$$
$$c\mathbf{D}'_{\perp} = \gamma(c\mathbf{D}_{\perp} + \boldsymbol{\beta} \times \mathbf{H}_{\perp}), \tag{2.89}$$
$$\mathbf{H}'_{\perp} = \gamma(\mathbf{H}_{\perp} - \boldsymbol{\beta} \times c\mathbf{D}_{\perp}), \tag{2.90}$$

and

$$\mathbf{J}'_{\parallel} = \gamma(\mathbf{J}_{\parallel} - \rho\mathbf{v}), \tag{2.91}$$
$$\mathbf{J}'_{\perp} = \mathbf{J}_{\perp}, \tag{2.92}$$
$$c\rho' = \gamma(c\rho - \boldsymbol{\beta} \cdot \mathbf{J}), \tag{2.93}$$

where the symbols $\parallel$ and $\perp$ designate the components of the field parallel and perpendicular to $\mathbf{v}$, respectively.

These conversions are self-consistent, and the Lorentz transformation is the transformation under which Maxwell's equations are covariant. If $v^2 \ll c^2$, then $\gamma \approx 1$ and to first order (2.87)–(2.93) reduce to (2.61)–(2.66). If $v/c \ll 1$, then the first-order fields reduce to the Galilean fields (2.49)–(2.54).

To convert in the opposite direction, we can swap primed and unprimed fields and change the sign on $\mathbf{v}$:

$$\mathbf{E}_{\perp} = \gamma(\mathbf{E}'_{\perp} - \boldsymbol{\beta} \times c\mathbf{B}'_{\perp}), \tag{2.94}$$
$$c\mathbf{B}_{\perp} = \gamma(c\mathbf{B}'_{\perp} + \boldsymbol{\beta} \times \mathbf{E}'_{\perp}), \tag{2.95}$$
$$c\mathbf{D}_{\perp} = \gamma(c\mathbf{D}'_{\perp} - \boldsymbol{\beta} \times \mathbf{H}'_{\perp}), \tag{2.96}$$
$$\mathbf{H}_{\perp} = \gamma(\mathbf{H}'_{\perp} + \boldsymbol{\beta} \times c\mathbf{D}'_{\perp}), \tag{2.97}$$

and

$$\mathbf{J}_{\parallel} = \gamma(\mathbf{J}'_{\parallel} + \rho'\mathbf{v}), \tag{2.98}$$
$$\mathbf{J}_{\perp} = \mathbf{J}'_{\perp}, \tag{2.99}$$
$$c\rho = \gamma(c\rho' + \boldsymbol{\beta} \cdot \mathbf{J}'). \tag{2.100}$$

The conversion formulas can be written much more succinctly in dyadic notation:

$$\mathbf{E}' = \gamma\bar{\boldsymbol{\alpha}}^{-1} \cdot \mathbf{E} + \gamma\bar{\boldsymbol{\beta}} \cdot (c\mathbf{B}), \tag{2.101}$$
$$c\mathbf{B}' = -\gamma\bar{\boldsymbol{\beta}} \cdot \mathbf{E} + \gamma\bar{\boldsymbol{\alpha}}^{-1} \cdot (c\mathbf{B}), \tag{2.102}$$

$$c\mathbf{D}' = \gamma\bar{\boldsymbol{\alpha}}^{-1} \cdot (c\mathbf{D}) + \gamma\bar{\boldsymbol{\beta}} \cdot \mathbf{H}, \tag{2.103}$$
$$\mathbf{H}' = -\gamma\bar{\boldsymbol{\beta}} \cdot (c\mathbf{D}) + \gamma\bar{\boldsymbol{\alpha}}^{-1} \cdot \mathbf{H}, \tag{2.104}$$

and

$$c\rho' = \gamma(c\rho - \boldsymbol{\beta} \cdot \mathbf{J}), \tag{2.105}$$
$$\mathbf{J}' = \bar{\boldsymbol{\alpha}} \cdot \mathbf{J} - \gamma\boldsymbol{\beta}c\rho, \tag{2.106}$$

where $\bar{\boldsymbol{\alpha}}^{-1} \cdot \bar{\boldsymbol{\alpha}} = \bar{\mathbf{I}}$, and thus $\bar{\boldsymbol{\alpha}}^{-1} = \bar{\boldsymbol{\alpha}} - \gamma\boldsymbol{\beta}\boldsymbol{\beta}$.

Maxwell's equations are covariant under a Lorentz transformation but not under a Galilean transformation; the laws of mechanics are invariant under a Galilean transformation but not under a Lorentz transformation. How then should we analyze interactions between electromagnetic fields and particles or materials? Einstein realized that the laws of mechanics needed revision to make them Lorentz covariant: in fact, under his theory of special relativity all physical laws should demonstrate Lorentz covariance. Interestingly, charge is then Lorentz invariant, whereas mass is not (recall that invariance refers to a quantity, whereas covariance refers to the form of a natural law). We shall not attempt to describe all the ramifications of special relativity, but instead refer the reader to any of the excellent and readable texts on the subject, including those by Bohm [17], Einstein [70], and Born [21], and to the nice historical account by Miller [149]. However, we shall examine the importance of Lorentz invariants in electromagnetic theory.

Lorentz invariants. Although the electromagnetic fields are not Lorentz invariant (e.g., the numerical value of $\mathbf{E}$ measured by one observer differs from that measured by another observer in uniform relative motion), several quantities do give identical values regardless of the velocity of motion. Most fundamental are the speed of light and the quantity of electric charge which, unlike mass, is the same in all frames of reference. Other important *Lorentz invariants* include $\mathbf{E} \cdot \mathbf{B}$, $\mathbf{H} \cdot \mathbf{D}$, and the quantities

$$\begin{aligned} \mathbf{B} \cdot \mathbf{B} - \mathbf{E} \cdot \mathbf{E}/c^2, \\ \mathbf{H} \cdot \mathbf{H} - c^2 \mathbf{D} \cdot \mathbf{D}, \\ \mathbf{B} \cdot \mathbf{H} - \mathbf{E} \cdot \mathbf{D}, \\ c\mathbf{B} \cdot \mathbf{D} + \mathbf{E} \cdot \mathbf{H}/c. \end{aligned}$$

(See Problem 2.3.) To see the importance of these quantities, consider the special case of fields in empty space. If $\mathbf{E} \cdot \mathbf{B} = 0$ in one reference frame, then it is zero in all reference frames. Then if $\mathbf{B} \cdot \mathbf{B} - \mathbf{E} \cdot \mathbf{E}/c^2 = 0$ in any reference frame, the ratio of E to B is always c^2 regardless of the reference frame in which the fields are measured. This is the characteristic of a plane wave in free space.

If $\mathbf{E} \cdot \mathbf{B} = 0$ and $c^2B^2 > E^2$, then we can find a reference frame using the conversion formulas (2.101)–(2.106) (see Problem 2.5) in which the electric field is zero but the magnetic field is nonzero. In this case we call the fields *purely magnetic* in any reference frame, even if both $\mathbf{E}$ and $\mathbf{B}$ are nonzero. Similarly, if $\mathbf{E} \cdot \mathbf{B} = 0$ and $c^2B^2 < E^2$ then we can find a reference frame in which the magnetic field is zero but the electric field is nonzero. We call fields of this type *purely electric*.

The Lorentz force is not Lorentz invariant. Consider a point charge at rest in the laboratory frame. While we measure only an electric field in the laboratory frame, an inertial observer measures both electric and magnetic fields. A test charge Q in the laboratory frame experiences the Lorentz force $\mathbf{F} = Q\mathbf{E}$; in an inertial frame the same charge experiences $\mathbf{F}' = Q\mathbf{E}' + Q\mathbf{v} \times \mathbf{B}'$ (see Problem 2.6). The conversion formulas show that $\mathbf{F}$ and $\mathbf{F}'$ are not identical.

We see that both $\mathbf{E}$ and $\mathbf{B}$ are integral components of the electromagnetic field: the separation of the field into electric and magnetic components depends on the motion of the reference frame in which measurements are made. This has obvious implications when considering static electric and magnetic fields.

Derivation of Maxwell's equations from Coulomb's law. Consider a point charge at rest in the laboratory frame. If the magnetic component of force on this charge arises

naturally through motion of an inertial reference frame, and if this force can be expressed in terms of Coulomb's law in the laboratory frame, then perhaps the magnetic field can be derived directly from Coulomb's law and the Lorentz transformation. Perhaps it is possible to derive all of Maxwell's theory with Coulomb's law and Lorentz invariance as the only postulates.

Several authors, notably Purcell [175] and Elliott [73], have used this approach. However, Jackson [105] has pointed out that many additional assumptions are required to deduce Maxwell's equations beginning with Coulomb's law. Feynman [81] is critical of the approach, pointing out that we must introduce a vector potential which adds to the scalar potential from electrostatics in order to produce an entity that transforms according to the laws of special relativity. In addition, the assumption of Lorentz invariance seems to involve circular reasoning since the Lorentz transformation was originally introduced to make Maxwell's equations covariant. But Lucas and Hodgson [135] point out that the Lorentz transformation can be deduced from other fundamental principles (such as causality and the isotropy of space), and that the postulate of a vector potential is reasonable. Schwartz [198] gives a detailed derivation of Maxwell's equations from Coulomb's law, outlining the necessary assumptions.

Transformation of constitutive relations. Minkowski's interest in the covariance of Maxwell's equations was aimed not merely at the relationship between fields in different moving frames of reference, but at an understanding of the electrodynamics of moving media. He wished to ascertain the effect of a moving material body on the electromagnetic fields in some region of space. By proposing the covariance of Maxwell's equations in materials as well as in free space, he extended Maxwell's theory to moving material bodies.

We have seen in (2.101)–(2.104) that $(\mathbf{E}, c\mathbf{B})$ and $(c\mathbf{D}, \mathbf{H})$ convert identically under a Lorentz transformation. Since the most general form of the constitutive relations relate $c\mathbf{D}$ and $\mathbf{H}$ to the field pair $(\mathbf{E}, c\mathbf{B})$ (see § 2.2.2) as

$$\begin{bmatrix} c\mathbf{D} \\ \mathbf{H} \end{bmatrix} = [\bar{\mathbf{C}}] \begin{bmatrix} \mathbf{E} \\ c\mathbf{B} \end{bmatrix},$$

this form of the constitutive relations must be Lorentz covariant. That is, in the reference frame of a moving material we have

$$\begin{bmatrix} c\mathbf{D}' \\ \mathbf{H}' \end{bmatrix} = [\bar{\mathbf{C}}'] \begin{bmatrix} \mathbf{E}' \\ c\mathbf{B}' \end{bmatrix},$$

and should be able to convert $[\bar{\mathbf{C}}']$ to $[\bar{\mathbf{C}}]$. We should be able to find the constitutive matrix describing the relationships among the fields observed in the laboratory frame.

It is somewhat laborious to obtain the constitutive matrix $[\bar{\mathbf{C}}]$ for an arbitrary moving medium. Detailed expressions for isotropic, bianisotropic, gyrotropic, and uniaxial media are given by Kong [118]. The rather complicated expressions can be written in a more compact form if we consider the expressions for $\mathbf{B}$ and $\mathbf{D}$ in terms of the pair $(\mathbf{E}, \mathbf{H})$. For a linear isotropic material such that $\mathbf{D}' = \epsilon'\mathbf{E}'$ and $\mathbf{B}' = \mu'\mathbf{H}'$ in the moving frame, the relationships in the laboratory frame are [118]

$$\mathbf{B} = \mu'\bar{\mathbf{A}} \cdot \mathbf{H} - \boldsymbol{\Omega} \times \mathbf{E}, \tag{2.107}$$

$$\mathbf{D} = \epsilon'\bar{\mathbf{A}} \cdot \mathbf{E} + \boldsymbol{\Omega} \times \mathbf{H}, \tag{2.108}$$

where

$$\bar{\mathbf{A}} = \frac{1-\beta^2}{1-n^2\beta^2}\left[\bar{\mathbf{I}} - \frac{n^2-1}{1-\beta^2}\boldsymbol{\beta}\boldsymbol{\beta}\right], \tag{2.109}$$

$$\boldsymbol{\Omega} = \frac{n^2-1}{1-n^2\beta^2}\frac{\boldsymbol{\beta}}{c}, \tag{2.110}$$

and where $n = c(\mu'\epsilon')^{1/2}$ is the optical index of the medium. A moving material that is isotropic in its own moving reference frame is bianisotropic in the laboratory frame. If, for instance, we tried to measure the relationship between the fields of a moving isotropic fluid, but used instruments that were stationary in our laboratory (e.g., attached to our measurement bench) we would find that $\mathbf{D}$ depends not only on $\mathbf{E}$ but also on $\mathbf{H}$, and that $\mathbf{D}$ aligns with neither $\mathbf{E}$ nor $\mathbf{H}$. That a moving material isotropic in its own frame of reference is bianisotropic in the laboratory frame was known long ago. Roentgen showed experimentally in 1888 that a dielectric moving through an electric field becomes magnetically polarized, while H.A. Wilson showed in 1905 that a dielectric moving through a magnetic field becomes electrically polarized [160].

If $v^2/c^2 \ll 1$, we can consider the form of the constitutive equations for a first-order Lorentz transformation. Ignoring terms to order v^2/c^2 in (2.109) and (2.110), we obtain $\bar{\mathbf{A}} = \bar{\mathbf{I}}$ and $\boldsymbol{\Omega} = \mathbf{v}(n^2-1)/c^2$. Then, by (2.107) and (2.108),

$$\mathbf{B} = \mu'\mathbf{H} - (n^2-1)\frac{\mathbf{v}\times\mathbf{E}}{c^2}, \tag{2.111}$$

$$\mathbf{D} = \epsilon'\mathbf{E} + (n^2-1)\frac{\mathbf{v}\times\mathbf{H}}{c^2}. \tag{2.112}$$

We can also derive these from the first-order field conversion equations (2.61)–(2.64). From (2.61) and (2.62) we have

$$\mathbf{D}' = \mathbf{D} + \mathbf{v}\times\mathbf{H}/c^2 = \epsilon'\mathbf{E}' = \epsilon'(\mathbf{E} + \mathbf{v}\times\mathbf{B}).$$

Eliminating $\mathbf{B}$ via (2.64), we have

$$\mathbf{D} + \mathbf{v}\times\mathbf{H}/c^2 = \epsilon'\mathbf{E} + \epsilon'\mathbf{v}\times(\mathbf{v}\times\mathbf{E}/c^2) + \epsilon'\mathbf{v}\times\mathbf{B}' = \epsilon'\mathbf{E} + \epsilon'\mathbf{v}\times\mathbf{B}'$$

where we have neglected terms of order v^2/c^2. Since $\mathbf{B}' = \mu'\mathbf{H}' = \mu'(\mathbf{H} - \mathbf{v}\times\mathbf{D})$, we have

$$\mathbf{D} + \mathbf{v}\times\mathbf{H}/c^2 = \epsilon'\mathbf{E} + \epsilon'\mu'\mathbf{v}\times\mathbf{H} - \epsilon'\mu'\mathbf{v}\times\mathbf{v}\times\mathbf{D}.$$

Using $n^2 = c^2\mu'\epsilon'$ and neglecting the last term since it is of order v^2/c^2, we obtain

$$\mathbf{D} = \epsilon'\mathbf{E} + (n^2-1)\frac{\mathbf{v}\times\mathbf{H}}{c^2},$$

which is identical to the expression (2.112) obtained by approximating the exact result to first order. Similar steps produce (2.111). In a Galilean frame where $v/c \ll 1$, the expressions reduce to $\mathbf{D} = \epsilon'\mathbf{E}$ and $\mathbf{B} = \mu'\mathbf{H}$, and the isotropy of the fields is preserved.

For a conducting medium having

$$\mathbf{J}' = \sigma'\mathbf{E}'$$

in a moving reference frame, Cullwick [53] shows that in the laboratory frame

$$\mathbf{J} = \sigma'\gamma[\bar{\mathbf{I}} - \boldsymbol{\beta}\boldsymbol{\beta}]\cdot\mathbf{E} + \sigma'\gamma c\boldsymbol{\beta}\times\mathbf{B}.$$

For $v \ll c$ we can set $\gamma \approx 1$ and see that

$$\mathbf{J} = \sigma'(\mathbf{E} + \mathbf{v}\times\mathbf{B})$$

to first order.

Constitutive relations in deforming or rotating media. The transformations discussed in the previous paragraphs hold for media in uniform relative motion. When a material body undergoes deformation or rotation, the concepts of special relativity are not directly applicable. However, authors such as Pauli [165] and Sommerfeld [213] have maintained that Minkowski's theory is *approximately* valid for deforming or rotating media if $\mathbf{v}$ is taken to be the *instantaneous* velocity at each point within the body. The reasoning is that at any instant in time each point within the body has a velocity $\mathbf{v}$ that may be associated with some inertial reference frame (generally different for each point). Thus the constitutive relations for the material at that point, within some small time interval taken about the observation time, may be assumed to be those of a stationary material, and the relations measured by an observer within the laboratory frame may be computed using the inertial frame for that point. This *instantaneous rest-frame* theory is most accurate at small accelerations $d\mathbf{v}/dt$. Van Bladel [237] outlines its shortcomings. See also Anderson [3] and Mo [152] for detailed discussions of the electromagnetic properties of material media in accelerating frames of reference.

2.4 The Maxwell–Boffi equations

In any version of Maxwell's theory, the mediating field is the electromagnetic field described by four field vectors. In Minkowski's form of Maxwell's equations we use $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{H}$. As an alternative consider the electromagnetic field as represented by the vector fields $\mathbf{E}$, $\mathbf{B}$, $\mathbf{P}$, and $\mathbf{M}$, and described by

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.113}$$

$$\nabla \times (\mathbf{B}/\mu_0 - \mathbf{M}) = \mathbf{J} + \frac{\partial}{\partial t}(\epsilon_0 \mathbf{E} + \mathbf{P}), \tag{2.114}$$

$$\nabla \cdot (\epsilon_0 \mathbf{E} + \mathbf{P}) = \rho, \tag{2.115}$$

$$\nabla \cdot \mathbf{B} = 0. \tag{2.116}$$

These *Maxwell–Boffi equations* are named after L. Boffi, who formalized them for moving media [16]. The quantity $\mathbf{P}$ is the *polarization vector*, and $\mathbf{M}$ is the *magnetization vector*. The use of $\mathbf{P}$ and $\mathbf{M}$ in place of $\mathbf{D}$ and $\mathbf{H}$ is sometimes called an application of the *principle of Ampere and Lorentz* [234].

Let us examine the ramification of using (2.113)–(2.116) as the basis for a postulate of electromagnetics. These equations are similar to the Maxwell–Minkowski equations used earlier; must we rebuild all the underpinning of a new postulate, or can we use our original arguments based on the Minkowski form? For instance, how do we invoke uniqueness if we no longer have the field $\mathbf{H}$? What represents the flux of energy, formerly found using $\mathbf{E} \times \mathbf{H}$? And, importantly, are (2.113)–(2.114) form invariant under a Lorentz transformation?

It turns out that the set of vector fields $(\mathbf{E}, \mathbf{B}, \mathbf{P}, \mathbf{M})$ is merely a linear mapping of the set $(\mathbf{E}, \mathbf{D}, \mathbf{B}, \mathbf{H})$. As pointed out by Tai [225], any linear mapping of the four field vectors from Minkowski's form onto any other set of four field vectors will preserve the covariance of Maxwell's equations. Boffi chose to keep $\mathbf{E}$ and $\mathbf{B}$ intact and to introduce only two new fields; he could have kept $\mathbf{H}$ and $\mathbf{D}$ instead, or used a mapping that introduced four completely new fields (as did Chu). Many authors retain $\mathbf{E}$ and $\mathbf{H}$.

This is somewhat more cumbersome since these vectors do not convert as a pair under a Lorentz transformation. A discussion of the idea of field vector "pairing" appears in § 2.6.

The usefulness of the Boffi form lies in the specific mapping chosen. Comparison of (2.113)–(2.116) to (2.1)–(2.4) quickly reveals that

$$\mathbf{P} = \mathbf{D} - \epsilon_0 \mathbf{E}, \tag{2.117}$$

$$\mathbf{M} = \mathbf{B}/\mu_0 - \mathbf{H}. \tag{2.118}$$

We see that $\mathbf{P}$ is the difference between $\mathbf{D}$ in a material and $\mathbf{D}$ in free space, while $\mathbf{M}$ is the difference between $\mathbf{H}$ in free space and $\mathbf{H}$ in a material. In free space, $\mathbf{P} = \mathbf{M} = 0$.

Equivalent polarization and magnetization sources. The Boffi formulation provides a new way to regard $\mathbf{E}$ and $\mathbf{B}$. Maxwell grouped $(\mathbf{E}, \mathbf{H})$ as a pair of "force vectors" to be associated with line integrals (or curl operations in the point forms of his equations), and $(\mathbf{D}, \mathbf{B})$ as a pair of "flux vectors" associated with surface integrals (or divergence operations). That is, $\mathbf{E}$ is interpreted as belonging to the computation of "emf" as a line integral, while $\mathbf{B}$ is interpreted as a density of magnetic "flux" passing through a surface. Similarly, $\mathbf{H}$ yields the "mmf" about some closed path and $\mathbf{D}$ the electric flux through a surface. The introduction of $\mathbf{P}$ and $\mathbf{M}$ allows us to also regard $\mathbf{E}$ as a flux vector and $\mathbf{B}$ as a force vector — in essence, allowing the two fields $\mathbf{E}$ and $\mathbf{B}$ to take on the duties that required four fields in Minkowski's form. To see this, we rewrite the Maxwell–Boffi equations as

$$\begin{aligned}
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\
\nabla \times \frac{\mathbf{B}}{\mu_0} &= \left(\mathbf{J} + \nabla \times \mathbf{M} + \frac{\partial \mathbf{P}}{\partial t} \right) + \frac{\partial \epsilon_0 \mathbf{E}}{\partial t}, \\
\nabla \cdot (\epsilon_0 \mathbf{E}) &= (\rho - \nabla \cdot \mathbf{P}), \\
\nabla \cdot \mathbf{B} &= 0,
\end{aligned}$$

and compare them to the Maxwell–Minkowski equations for sources in free space:

$$\begin{aligned}
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\
\nabla \times \frac{\mathbf{B}}{\mu_0} &= \mathbf{J} + \frac{\partial \epsilon_0 \mathbf{E}}{\partial t}, \\
\nabla \cdot (\epsilon_0 \mathbf{E}) &= \rho, \\
\nabla \cdot \mathbf{B} &= 0.
\end{aligned}$$

The forms are preserved if we identify $\partial \mathbf{P}/\partial t$ and $\nabla \times \mathbf{M}$ as new types of current density, and $\nabla \cdot \mathbf{P}$ as a new type of charge density. We define

$$\mathbf{J}_P = \frac{\partial \mathbf{P}}{\partial t} \tag{2.119}$$

as an *equivalent polarization current* density, and

$$\mathbf{J}_M = \nabla \times \mathbf{M}$$

as an *equivalent magnetization current* density (sometimes called the *equivalent Amperian currents of magnetized matter* [234]). We define

$$\rho_P = -\nabla \cdot \mathbf{P}$$

as an *equivalent polarization charge* density (sometimes called the *Poisson–Kelvin* equivalent charge distribution [234]). Then the Maxwell–Boffi equations become simply

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.120}$$

$$\nabla \times \frac{\mathbf{B}}{\mu_0} = (\mathbf{J} + \mathbf{J}_M + \mathbf{J}_P) + \frac{\partial \epsilon_0 \mathbf{E}}{\partial t}, \tag{2.121}$$

$$\nabla \cdot (\epsilon_0 \mathbf{E}) = (\rho + \rho_P), \tag{2.122}$$

$$\nabla \cdot \mathbf{B} = 0. \tag{2.123}$$

Here is the new view. A material can be viewed as composed of charged particles of matter immersed in free space. When these charges are properly considered as "equivalent" polarization and magnetization charges, all field effects (describable through flux and force vectors) can be handled by the two fields **E** and **B**. Whereas in Minkowski's form **D** diverges from ρ, in Boffi's form **E** diverges from a *total* charge density consisting of ρ and ρ_P. Whereas in the Minkowski form **H** curls around **J**, in the Boffi form **B** curls around the total current density consisting of **J**, $\mathbf{J}_M$, and $\mathbf{J}_P$.

This view was pioneered by Lorentz, who by 1892 considered matter as consisting of bulk molecules in a vacuum that would respond to an applied electromagnetic field [149]. The resulting motion of the charged particles of matter then became another source term for the "fundamental" fields **E** and **B**. Using this reasoning he was able to reduce the fundamental Maxwell equations to two equations in two unknowns, demonstrating a simplicity appealing to many (including Einstein). Of course, to apply this concept we must be able to describe how the charged particles respond to an applied field. Simple microscopic models of the constituents of matter are generally used: some combination of electric and magnetic dipoles, or of loops of electric and magnetic current.

The Boffi equations are mathematically appealing since they now specify both the curl and divergence of the two field quantities **E** and **B**. By the Helmholtz theorem we know that a field vector is uniquely specified when both its curl and divergence are given. But this assumes that the equivalent sources produced by **P** and **M** are true source fields in the same sense as **J**. We have precluded this by insisting in Chapter 1 that the source field must be independent of the mediating field it sources. If we view **P** and **M** as merely a mapping from the original vector fields of Minkowski's form, we still have four vector fields with which to contend. And with these must also be a mapping of the constitutive relationships, which now link the fields **E**, **B**, **P**, and **M**. Rather than argue the actual physical existence of the equivalent sources, we note that a real benefit of the new view is that under certain circumstances the equivalent source quantities can be determined through physical reasoning. Hence we can create physical models of **P** and **M** and deduce their links to **E** and **B**. We may then find it easier to understand and deduce the constitutive relationships. However we do not in general consider **E** and **B** to be in any way more "fundamental" than **D** and **H**.

Covariance of the Boffi form. Because of the linear relationships (2.117) and (2.118), covariance of the Maxwell–Minkowski equations carries over to the Maxwell–Boffi equations. However, the conversion between fields in different moving reference frames will now involve **P** and **M**. Since Faraday's law is unchanged in the Boffi form, we still have

$$\mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel}, \tag{2.124}$$

$$\mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel}, \tag{2.125}$$

$$\mathbf{E}'_\perp = \gamma(\mathbf{E}_\perp + \boldsymbol{\beta} \times c\mathbf{B}_\perp), \tag{2.126}$$
$$c\mathbf{B}'_\perp = \gamma(c\mathbf{B}_\perp - \boldsymbol{\beta} \times \mathbf{E}_\perp). \tag{2.127}$$

To see how $\mathbf{P}$ and $\mathbf{M}$ convert, we note that in the laboratory frame $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$ and $\mathbf{H} = \mathbf{B}/\mu_0 - \mathbf{M}$, while in the moving frame $\mathbf{D}' = \epsilon_0\mathbf{E}' + \mathbf{P}'$ and $\mathbf{H}' = \mathbf{B}'/\mu_0 - \mathbf{M}'$. Thus

$$\mathbf{P}'_\parallel = \mathbf{D}'_\parallel - \epsilon_0\mathbf{E}'_\parallel = \mathbf{D}_\parallel - \epsilon_0\mathbf{E}_\parallel = \mathbf{P}_\parallel$$

and

$$\mathbf{M}'_\parallel = \mathbf{B}'_\parallel/\mu_0 - \mathbf{H}'_\parallel = \mathbf{B}_\parallel/\mu_0 - \mathbf{H}_\parallel = \mathbf{M}_\parallel.$$

For the perpendicular components

$$\mathbf{D}'_\perp = \gamma(\mathbf{D}_\perp + \boldsymbol{\beta} \times \mathbf{H}_\perp/c) = \epsilon_0\mathbf{E}'_\perp + \mathbf{P}'_\perp = \epsilon_0\left[\gamma(\mathbf{E}_\perp + \boldsymbol{\beta} \times c\mathbf{B}_\perp)\right] + \mathbf{P}'_\perp;$$

substitution of $\mathbf{H}_\perp = \mathbf{B}_\perp/\mu_0 - \mathbf{M}_\perp$ then gives

$$\mathbf{P}'_\perp = \gamma(\mathbf{D}_\perp - \epsilon_0\mathbf{E}_\perp) - \gamma\epsilon_0\boldsymbol{\beta} \times c\mathbf{B}_\perp + \gamma\boldsymbol{\beta} \times \mathbf{B}_\perp/(c\mu_0) - \gamma\boldsymbol{\beta} \times \mathbf{M}_\perp/c$$

or

$$c\mathbf{P}'_\perp = \gamma(c\mathbf{P}_\perp - \boldsymbol{\beta} \times \mathbf{M}_\perp).$$

Similarly,

$$\mathbf{M}'_\perp = \gamma(\mathbf{M}_\perp + \boldsymbol{\beta} \times c\mathbf{P}_\perp).$$

Hence

$$\mathbf{E}'_\parallel = \mathbf{E}_\parallel, \quad \mathbf{B}'_\parallel = \mathbf{B}_\parallel, \quad \mathbf{P}'_\parallel = \mathbf{P}_\parallel, \quad \mathbf{M}'_\parallel = \mathbf{M}_\parallel, \quad \mathbf{J}'_\perp = \mathbf{J}_\perp, \tag{2.128}$$

and

$$\mathbf{E}'_\perp = \gamma(\mathbf{E}_\perp + \boldsymbol{\beta} \times c\mathbf{B}_\perp), \tag{2.129}$$
$$c\mathbf{B}'_\perp = \gamma(c\mathbf{B}_\perp - \boldsymbol{\beta} \times \mathbf{E}_\perp), \tag{2.130}$$
$$c\mathbf{P}'_\perp = \gamma(c\mathbf{P}_\perp - \boldsymbol{\beta} \times \mathbf{M}_\perp), \tag{2.131}$$
$$\mathbf{M}'_\perp = \gamma(\mathbf{M}_\perp + \boldsymbol{\beta} \times c\mathbf{P}_\perp), \tag{2.132}$$
$$\mathbf{J}'_\parallel = \gamma(\mathbf{J}_\parallel - \rho\mathbf{v}). \tag{2.133}$$

In the case of the first-order Lorentz transformation we can set $\gamma \approx 1$ to obtain

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}, \tag{2.134}$$
$$\mathbf{B}' = \mathbf{B} - \frac{\mathbf{v} \times \mathbf{E}}{c^2}, \tag{2.135}$$
$$\mathbf{P}' = \mathbf{P} - \frac{\mathbf{v} \times \mathbf{M}}{c^2}, \tag{2.136}$$
$$\mathbf{M}' = \mathbf{M} + \mathbf{v} \times \mathbf{P}, \tag{2.137}$$
$$\mathbf{J}' = \mathbf{J} - \rho\mathbf{v}. \tag{2.138}$$

To convert from the moving frame to the laboratory frame we simply swap primed with unprimed fields and let $\mathbf{v} \to -\mathbf{v}$.

As a simple example, consider a linear isotropic medium having

$$\mathbf{D}' = \epsilon_0\epsilon'_r\mathbf{E}', \qquad \mathbf{B}' = \mu_0\mu'_r\mathbf{H}',$$

in a moving reference frame. From (2.117) we have

$$\mathbf{P}' = \epsilon_0\epsilon_r'\mathbf{E}' - \epsilon_0\mathbf{E}' = \epsilon_0\chi_e'\mathbf{E}'$$

where $\chi_e' = \epsilon_r' - 1$ is the electric susceptibility of the moving material. Similarly (2.118) yields

$$\mathbf{M}' = \frac{\mathbf{B}'}{\mu_0} - \frac{\mathbf{B}'}{\mu_0\mu_r'} = \frac{\mathbf{B}'\chi_m'}{\mu_0\mu_r'}$$

where $\chi_m' = \mu_r' - 1$ is the magnetic susceptibility of the moving material. How are $\mathbf{P}$ and $\mathbf{M}$ related to $\mathbf{E}$ and $\mathbf{B}$ in the laboratory frame? For simplicity, we consider the first-order expressions. From (2.136) we have

$$\mathbf{P} = \mathbf{P}' + \frac{\mathbf{v}\times\mathbf{M}'}{c^2} = \epsilon_0\chi_e'\mathbf{E}' + \frac{\mathbf{v}\times\mathbf{B}'\chi_m'}{\mu_0\mu_r'c^2}.$$

Substituting for $\mathbf{E}'$ and $\mathbf{B}'$ from (2.134) and (2.135), and using $\mu_0c^2 = 1/\epsilon_0$, we have

$$\mathbf{P} = \epsilon_0\chi_e'(\mathbf{E} + \mathbf{v}\times\mathbf{B}) + \epsilon_0\frac{\chi_m'}{\mu_r'}\mathbf{v}\times\left(\mathbf{B} - \frac{\mathbf{v}\times\mathbf{E}}{c^2}\right).$$

Neglecting the last term since it varies as v^2/c^2, we get

$$\mathbf{P} = \epsilon_0\chi_e'\mathbf{E} + \epsilon_0\left(\chi_e' + \frac{\chi_m'}{\mu_r'}\right)\mathbf{v}\times\mathbf{B}. \tag{2.139}$$

Similarly,

$$\mathbf{M} = \frac{\chi_m'}{\mu_0\mu_r'}\mathbf{B} - \epsilon_0\left(\chi_e' + \frac{\chi_m'}{\mu_r'}\right)\mathbf{v}\times\mathbf{E}. \tag{2.140}$$

2.5 Large-scale form of Maxwell's equations

We can write Maxwell's equations in a form that incorporates the spatial variation of the field in a certain region of space. To do this, we integrate the point form of Maxwell's equations over a region of space, then perform some succession of manipulations until we arrive at a form that provides us some benefit in our work with electromagnetic fields. The results are particularly useful for understanding the properties of electric and magnetic circuits, and for predicting the behavior of electrical machinery.

We shall consider two important situations: a mathematical surface that moves with constant velocity $\mathbf{v}$ and with constant shape, and a surface that moves and deforms arbitrarily.

2.5.1 Surface moving with constant velocity

Consider an open surface S moving with constant velocity $\mathbf{v}$ relative to the laboratory frame (Figure 2.2). Assume every point on the surface is an ordinary point. At any instant t we can express the relationship between the fields at points on S in either frame. In the laboratory frame we have

$$\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}, \qquad \nabla\times\mathbf{H} = \frac{\partial\mathbf{D}}{\partial t} + \mathbf{J},$$

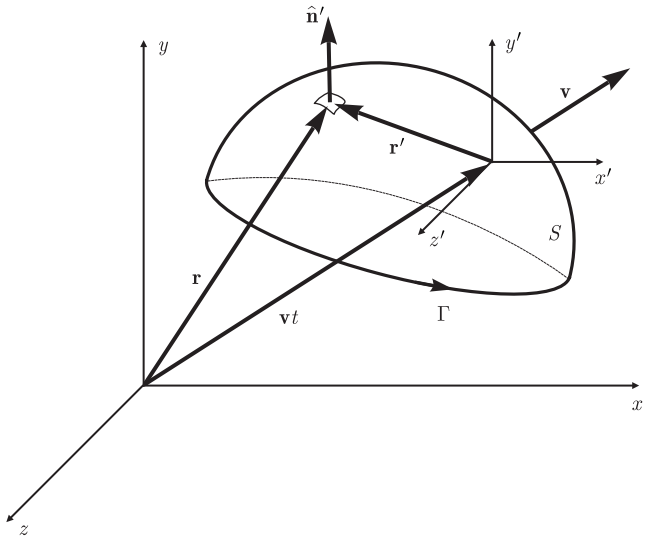

FIGURE 2.2
Open surface having velocity $\mathbf{v}$ relative to laboratory (unprimed) coordinate system. Surface is non-deforming.

while in the moving frame

$$\nabla' \times \mathbf{E}' = -\frac{\partial \mathbf{B}'}{\partial t'}, \qquad \nabla' \times \mathbf{H}' = \frac{\partial \mathbf{D}'}{\partial t'} + \mathbf{J}'.$$

If we integrate over S and use Stokes's theorem, we get for the laboratory frame

$$\oint_\Gamma \mathbf{E} \cdot \mathbf{dl} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{dS}, \tag{2.141}$$

$$\oint_\Gamma \mathbf{H} \cdot \mathbf{dl} = \int_S \frac{\partial \mathbf{D}}{\partial t} \cdot \mathbf{dS} + \int_S \mathbf{J} \cdot \mathbf{dS}, \tag{2.142}$$

and for the moving frame

$$\oint_{\Gamma'} \mathbf{E}' \cdot \mathbf{dl}' = -\int_{S'} \frac{\partial \mathbf{B}'}{\partial t'} \cdot \mathbf{dS}', \tag{2.143}$$

$$\oint_{\Gamma'} \mathbf{H}' \cdot \mathbf{dl}' = \int_{S'} \frac{\partial \mathbf{D}'}{\partial t'} \cdot \mathbf{dS}' + \int_{S'} \mathbf{J}' \cdot \mathbf{dS}'. \tag{2.144}$$

Here boundary contour Γ has sense determined by the right-hand rule. We use the notation Γ', S', etc., to indicate that all integrations for the moving frame are computed using space and time variables in that frame. Equation (2.141) is the *integral form of Faraday's law*, while (2.142) is the *integral form of Ampere's law.*

Faraday's law states that the net circulation of $\mathbf{E}$ about a contour Γ (sometimes called the *electromotive force* or *emf*) is determined by the flux of the time-rate of change of the flux vector $\mathbf{B}$ passing through the surface bounded by Γ. Ampere's law states that the circulation of $\mathbf{H}$ (sometimes called the *magnetomotive force* or *mmf*) is determined by the flux of the current $\mathbf{J}$ plus the flux of the time-rate of change of the flux vector $\mathbf{D}$. It is the term containing $\partial \mathbf{D}/\partial t$ that Maxwell recognized as necessary to make his equations

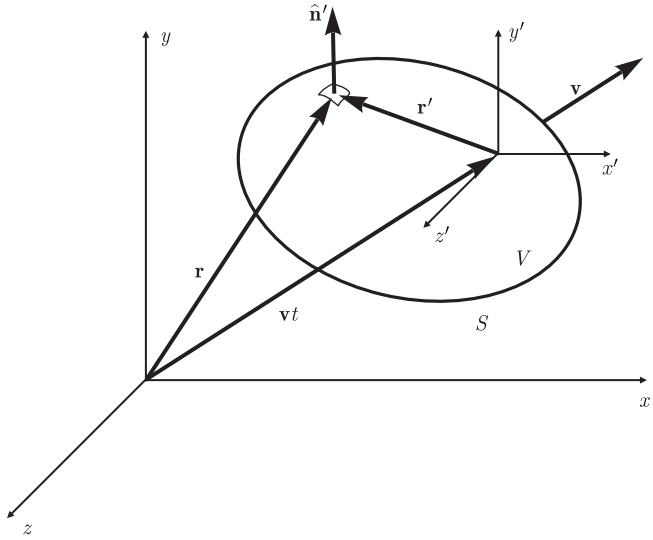

FIGURE 2.3
Non-deforming volume region having velocity $\mathbf{v}$ relative to laboratory (unprimed) coordinate system.

consistent; since it has units of current, it is often referred to as the *displacement current* term.

Equations (2.141)–(2.142) are the large-scale or integral forms of Maxwell's equations. They are the integral-form equivalents of the point forms, and are form invariant under Lorentz transformation. If we express the fields in terms of the moving reference frame, we can write

$$\oint_{\Gamma'} \mathbf{E}' \cdot \mathbf{dl}' = -\frac{d}{dt'} \int_{S'} \mathbf{B}' \cdot \mathbf{dS}', \tag{2.145}$$

$$\oint_{\Gamma'} \mathbf{H}' \cdot \mathbf{dl}' = \frac{d}{dt'} \int_{S'} \mathbf{D}' \cdot \mathbf{dS}' + \int_{S'} \mathbf{J}' \cdot \mathbf{dS}'. \tag{2.146}$$

These hold for a stationary surface, since the surface would be stationary to an observer who moves with it. We are therefore justified in removing the partial derivative from the integral. Although the surfaces and contours considered here are purely mathematical, they often coincide with actual physical boundaries. The surface may surround a moving material medium, for instance, or the contour may conform to a wire moving in an electrical machine.

We can also convert the auxiliary equations to large-scale form. Consider a volume region V surrounded by a surface S that moves with velocity $\mathbf{v}$ relative to the laboratory frame (Figure 2.3). Integrating the point form of Gauss's law over V we have

$$\int_V \nabla \cdot \mathbf{D}\, dV = \int_V \rho\, dV.$$

Using the divergence theorem and recognizing that the integral of charge density is total charge, we obtain

$$\oint_S \mathbf{D} \cdot \mathbf{dS} = \int_V \rho\, dV = Q(t) \tag{2.147}$$

where $Q(t)$ is the total charge contained within V at time t. This large-scale form of Gauss's law states that the total flux of $\mathbf{D}$ passing through a closed surface is identical to the electric charge Q contained within. Similarly,

$$\oint_S \mathbf{B} \cdot \mathbf{dS} = 0 \tag{2.148}$$

is the large-scale magnetic field Gauss's law. It states that the total flux of $\mathbf{B}$ passing through a closed surface is zero, since there are no magnetic charges contained within (i.e., magnetic charge does not exist).

Since charge is an invariant quantity, the large-scale forms of the auxiliary equations take the same form in a moving reference frame:

$$\oint_{S'} \mathbf{D}' \cdot \mathbf{dS}' = \int_{V'} \rho' \, dV' = Q(t) \tag{2.149}$$

and

$$\oint_{S'} \mathbf{B}' \cdot \mathbf{dS}' = 0. \tag{2.150}$$

The large-scale forms of the auxiliary equations may be derived from the large-scale forms of Faraday's and Ampere's laws. To obtain Gauss's law, we let the open surface in Ampere's law become a closed surface. Then $\oint \mathbf{H} \cdot \mathbf{dl}$ vanishes, and application of the large-scale form of the continuity equation (1.10) produces (2.147). The magnetic Gauss's law (2.148) is found from Faraday's law (2.141) by a similar transition from an open surface to a closed surface.

The values obtained from the expressions (2.141)–(2.142) will *not* match those obtained from (2.143)–(2.144), and we can use the Lorentz transformation field conversions to study how they differ. That is, we can write either side of the laboratory equations in terms of the moving reference frame fields, or vice versa. For most engineering applications where $v/c \ll 1$ this is not done via the Lorentz transformation field relations, but rather via the Galilean approximations to these relations (see Tai [226] for details on using the Lorentz transformation field relations). We consider the most common situation in the next section.

Kinematic form of the large-scale Maxwell equations. Confusion can result from the fact that the large-scale forms of Maxwell's equations can be written in a number of ways. A popular formulation of Faraday's law, the *emf formulation*, revolves around the concept of electromotive force. Unfortunately, various authors offer different definitions of emf in a moving circuit.

Consider a non-deforming contour in space, moving with constant velocity $\mathbf{v}$ relative to the laboratory frame (Figure 2.4). In terms of the laboratory fields we have the large-scale form of Faraday's law (2.141). The flux term on the right-hand side of this equation can be written differently by employing the Helmholtz transport theorem (A.63). If a non-deforming surface S moves with uniform velocity $\mathbf{v}$ relative to the laboratory frame, and a vector field $\mathbf{A}(\mathbf{r}, t)$ is expressed in the stationary frame, then the time derivative of the flux of $\mathbf{A}$ through S is

$$\frac{d}{dt} \int_S \mathbf{A} \cdot \mathbf{dS} = \int_S \left[\frac{\partial \mathbf{A}}{\partial t} + \mathbf{v}(\nabla \cdot \mathbf{A}) - \nabla \times (\mathbf{v} \times \mathbf{A}) \right] \cdot \mathbf{dS}. \tag{2.151}$$

Using this with (2.141) we have

$$\oint_\Gamma \mathbf{E} \cdot \mathbf{dl} = -\frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{dS} + \int_S \mathbf{v}(\nabla \cdot \mathbf{B}) \cdot \mathbf{dS} - \int_S \nabla \times (\mathbf{v} \times \mathbf{B}) \cdot \mathbf{dS}.$$

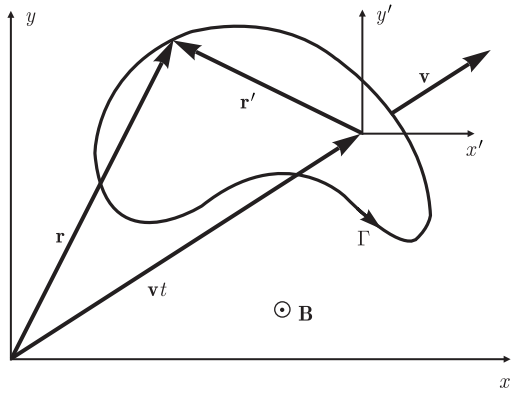

FIGURE 2.4
Non-deforming closed contour moving with velocity **v** through a magnetic field **B** given in the laboratory (unprimed) coordinate system.

Remembering that $\nabla \cdot \mathbf{B} = 0$ and using Stokes's theorem on the last term, we obtain

$$\oint_\Gamma (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{dl} = -\frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{dS} = -\frac{d\Psi(t)}{dt} \tag{2.152}$$

where the *magnetic flux*

$$\int_S \mathbf{B} \cdot \mathbf{dS} = \Psi(t)$$

represents the flux of **B** through S. Following Sommerfeld [213], we may set

$$\mathbf{E}^* = \mathbf{E} + \mathbf{v} \times \mathbf{B}$$

to obtain the *kinematic form of Faraday's law*

$$\oint_\Gamma \mathbf{E}^* \cdot \mathbf{dl} = -\frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{dS} = -\frac{d\Psi(t)}{dt}. \tag{2.153}$$

(The asterisk should not be confused with the notation for complex conjugate.)

Much confusion arises from the similarity between (2.153) and (2.145). In fact, these expressions are different and give different results. This is because $\mathbf{B}'$ in (2.145) is measured *in the frame of the moving circuit*, while **B** in (2.153) is measured in the frame of the laboratory. Further confusion arises from various definitions of emf. Many authors (e.g., Hermann Weyl [249]) define emf to be the circulation of $\mathbf{E}^*$. In that case the emf is equal to the negative time rate of change of the flux of the *laboratory frame* magnetic field **B** through S. Since the Lorentz force experienced by a charge q moving *with the contour* is given by $q\mathbf{E}^* = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, this emf is the circulation of Lorentz force per unit charge along the contour. If the contour is aligned with a conducting circuit, then in some cases this emf can be given physical interpretation as the work required to move a charge around the entire circuit through the conductor against the Lorentz force. Unfortunately the usefulness of this definition of emf is lost if the time or space rate of change of the fields is so large that no true loop current can be established (hence Kirchoff's law cannot be employed). Such a problem must be treated as an electromagnetic "scattering" problem with consideration given to retardation effects.

Detailed discussions of the physical interpretation of $\mathbf{E}^*$ in the definition of emf are given by Scanlon [193] and Cullwick [53].

Other authors choose to define emf as the circulation of the electric field *in the frame of the moving contour*. In this case the circulation of $\mathbf{E}'$ in (2.145) is the emf, and is related to the flux of the magnetic field *in the frame of the moving circuit*. As pointed out above, the result differs from that based on the Lorentz force. If we wish, we can also write this emf in terms of the fields expressed in the laboratory frame. To do this we must convert $\partial\mathbf{B}'/\partial t'$ to the laboratory fields using the rules for a Lorentz transformation. The result, given by Tai [226], is quite complicated and involves both the magnetic *and* electric laboratory-frame fields.

The moving-frame emf as computed from the Lorentz transformation is rarely used as a working definition of emf, mostly because circuits moving at relativistic velocities are seldom used by engineers. Unfortunately, more confusion arises for the case $v \ll c$, since for a Galilean frame the Lorentz-force and moving-frame emfs become identical. This is apparent if we use (2.52) to replace $\mathbf{B}'$ with the laboratory frame field $\mathbf{B}$, and (2.49) to replace $\mathbf{E}'$ with the combination of laboratory frame fields $\mathbf{E} + \mathbf{v} \times \mathbf{B}$. Then (2.145) becomes

$$\oint_\Gamma \mathbf{E}' \cdot \mathbf{dl} = \oint_\Gamma (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{dl} = -\frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{dS},$$

which is identical to (2.153). For circuits moving with low velocity then, the circulation of $\mathbf{E}'$ can be interpreted as work per unit charge. As an added bit of confusion, the term

$$\oint_\Gamma (\mathbf{v} \times \mathbf{B}) \cdot \mathbf{dl} = \int_S \nabla \times (\mathbf{v} \times \mathbf{B}) \cdot \mathbf{dS}$$

is sometimes called *motional emf*, since it is the component of the circulation of $\mathbf{E}^*$ that is directly attributable to the motion of the circuit.

Although less commonly done, we can also rewrite Ampere's law (2.142) using (2.151). This gives

$$\oint_\Gamma \mathbf{H} \cdot \mathbf{dl} = \int_S \mathbf{J} \cdot \mathbf{dS} + \frac{d}{dt} \int_S \mathbf{D} \cdot \mathbf{dS} - \int_S (\mathbf{v} \nabla \cdot \mathbf{D}) \cdot \mathbf{dS} + \int_S \nabla \times (\mathbf{v} \times \mathbf{D}) \cdot \mathbf{dS}.$$

Using $\nabla \cdot \mathbf{D} = \rho$ and using Stokes's theorem on the last term, we obtain

$$\oint_\Gamma (\mathbf{H} - \mathbf{v} \times \mathbf{D}) \cdot \mathbf{dl} = \frac{d}{dt} \int_S \mathbf{D} \cdot \mathbf{dS} + \int_S (\mathbf{J} - \rho \mathbf{v}) \cdot \mathbf{dS}.$$

Finally, letting $\mathbf{H}^* = \mathbf{H} - \mathbf{v} \times \mathbf{D}$ and $\mathbf{J}^* = \mathbf{J} - \rho \mathbf{v}$ we can write the *kinematic form of Ampere's law*:

$$\oint_\Gamma \mathbf{H}^* \cdot \mathbf{dl} = \frac{d}{dt} \int_S \mathbf{D} \cdot \mathbf{dS} + \int_S \mathbf{J}^* \cdot \mathbf{dS}. \tag{2.154}$$

In a Galilean frame where we use (2.49)–(2.54), we see that (2.154) is identical to

$$\oint_\Gamma \mathbf{H}' \cdot \mathbf{dl} = \frac{d}{dt'} \int_S \mathbf{D}' \cdot \mathbf{dS} + \int_S \mathbf{J}' \cdot \mathbf{dS} \tag{2.155}$$

where the primed fields are measured in the frame of the moving contour. This equivalence does *not* hold when the Lorentz transformation is used to represent the primed fields.

Alternative form of the large-scale Maxwell equations. We can write Maxwell's equations in an alternative large-scale form involving only surface and volume integrals. This will be useful later for establishing the field jump conditions across a material or source discontinuity. Again we begin with Maxwell's equations in point form, but instead of integrating them over an open surface we integrate over a volume region V moving with velocity $\mathbf{v}$ (Figure 2.3). In the laboratory frame this gives

$$\int_V (\nabla \times \mathbf{E})\, dV = -\int_V \frac{\partial \mathbf{B}}{\partial t}\, dV,$$
$$\int_V (\nabla \times \mathbf{H})\, dV = \int_V \left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}\right) dV.$$

An application of curl theorem (B.24) then gives

$$\oint_S (\hat{\mathbf{n}} \times \mathbf{E})\, dS = -\int_V \frac{\partial \mathbf{B}}{\partial t}\, dV, \tag{2.156}$$
$$\oint_S (\hat{\mathbf{n}} \times \mathbf{H})\, dS = \int_V \left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}\right) dV. \tag{2.157}$$

Similar results are obtained for the fields in the moving frame:

$$\oint_{S'} (\hat{\mathbf{n}}' \times \mathbf{E}')\, dS' = -\int_{V'} \frac{\partial \mathbf{B}'}{\partial t'}\, dV',$$
$$\oint_{S'} (\hat{\mathbf{n}}' \times \mathbf{H}')\, dS' = \int_{V'} \left(\frac{\partial \mathbf{D}'}{\partial t'} + \mathbf{J}'\right) dV'.$$

These large-scale forms are an alternative to (2.141)–(2.144). They are also form-invariant under a Lorentz transformation.

An alternative to the kinematic formulation of (2.153) and (2.154) can be achieved by applying a kinematic identity for a moving volume region. If V is surrounded by a surface S that moves with velocity $\mathbf{v}$ relative to the laboratory frame, and if a vector field $\mathbf{A}$ is measured in the laboratory frame, then the vector form of the general transport theorem (A.68) states that

$$\frac{d}{dt}\int_V \mathbf{A}\, dV = \int_V \frac{\partial \mathbf{A}}{\partial t}\, dV + \oint_S \mathbf{A}(\mathbf{v} \cdot \hat{\mathbf{n}})\, dS. \tag{2.158}$$

Applying this to (2.156) and (2.157) we have

$$\oint_S [\hat{\mathbf{n}} \times \mathbf{E} - (\mathbf{v} \cdot \hat{\mathbf{n}})\mathbf{B}]\, dS = -\frac{d}{dt}\int_V \mathbf{B}\, dV, \tag{2.159}$$
$$\oint_S [\hat{\mathbf{n}} \times \mathbf{H} + (\mathbf{v} \cdot \hat{\mathbf{n}})\mathbf{D}]\, dS = \int_V \mathbf{J}\, dV + \frac{d}{dt}\int_V \mathbf{D}\, dV. \tag{2.160}$$

We can also apply (2.158) to the large-scale form of the continuity equation (1.10) and obtain the expression for a volume region moving with velocity $\mathbf{v}$:

$$\oint_S (\mathbf{J} - \rho\mathbf{v}) \cdot \mathbf{dS} = -\frac{d}{dt}\int_V \rho\, dV.$$