Foundations of Electromagnetic Theory

THIRD EDITION

JOHN R. REITZ
FREDERICK J. MILFORD
ROBERT W. CHRISTY

Foundations of Electromagnetic Theory

THIRD EDITION

JOHN R. REITZ
Scientific Laboratory
Ford Motor Company

FREDERICK J. MILFORD
Battelle Memorial Institute

ROBERT W. CHRISTY
Dartmouth College

ADDISON-WESLEY PUBLISHING COMPANY

Reading, Massachusetts • Menlo Park, California
London • Amsterdam • Don Mills, Ontario • Sydney

This book is in the
Addison-Wesley Series in Physics

 Library of Congress Catalog Card No. 78-18649.

ISBN 0-201-06332-8
ABCDEFGHIJ-MA-79

Preface

Although Maxwell's equations were formulated about one hundred years ago, the subject of electromagnetism has not remained static. Advanced undergraduate students in science, to whom we are directing our attention, today approach the subject with a qualitative understanding of atomic phenomena. At the same time, they have acquired a good background in mathematics and for the first time are in a position to solve some of the important problems of classical physics. The present volume evolved from the teaching of courses in electricity and magnetism to physics majors at Case Institute of Technology and at Dartmouth College. A course in electromagnetism is ideally suited to a development of the concepts of vector analysis, partial differential equations, and boundary-value problems. The sections involving these techniques are written in such a way that little previous knowledge of the subject is required.

We believe that building up electricity and magnetism from the basic experimental laws is the correct approach at the intermediate level, and we have followed this approach through a rigorous exposition of the fundamentals. We have also been careful to include a number of appropriate examples to bridge the gap between the formal development of the subject and the problems. A full understanding of the electric and magnetic fields inside matter can be obtained only after the atomic nature of materials is appreciated. Hence we have used elementary atomic concepts freely in the development of macroscopic theory.

We prefer to discuss the static electric field in a material medium immediately after the vacuum electric field, and we discuss the magnetostatic field similarly. The reader may, however, study both vacuum cases together before considering either electric or magnetic fields in matter, by postponing Chapters 4, 5, 6, 7 (except Sections 7.1 and 7.2), 9, and 10 until after reading Chapter 8 or even Chapter 11. The macroscopic electromagnetic behavior of dielectrics, conductors, magnetic materials, plasmas, and superconductors is treated in separate chapters (Chapters 4, 7, 9, 14, and 15, respectively). A simple discussion of the microscopic theory of these classes of matter (except superconductors) is also provided (in Chapters 5, 7, 10, and 14).

The third edition of the book is changed principally by the addition of more material on *electromagnetic waves*. The two old chapters on Maxwell's equations have been expanded into five chapters. The book is thus adaptable either to a one-semester course or to a two-semester course in which the second semester emphasizes propagation and generation of radiation.

Much of modern physics (and engineering) involves time-dependent electromagnetic fields in which Maxwell's displacement current plays a crucial role. Chapters 16 through 20 develop the application to waves—especially the connection with optics, which is the frequency range that is now succeeding microwaves in technological interest. Chapters 16 and 17 extend the old treatment of the wave equations, introducing the idea of gauge transformations. The notions of complex dielectric function and refractive index are emphasized, with resulting conceptual clarity and simplification of formulas. Chapter 18 expands the treatment of boundary-value problems, to include examples of interest in optical filters and waveguides. Chapter 19 gives the classical microscopic theory of transverse wave propagation in matter (dielectrics, metals, plasmas); it is an extension of Chapters 5 and 7 to time-dependent fields. It also includes a simple discussion of the Kramers-Kronig dispersion relations for a linear response function. Chapter 20, on the generation of radiation by antennas and accelerated charges, includes new material on induction fields, radiation damping, and Thomson scattering.

The material in the rest of the book has been slightly rearranged, so that the discussion of static fields and steady currents is completed before the introduction of Faraday's law of induction, in Chapter 11, followed by its application to slowly varying currents in a-c circuits, plasmas, and superconductors in Chapters 13, 14, and 15. The relativistic formulation of electromagnetism has been put at the end, although it could be read at any point after Chapter 16. Some relativistic aspects are anticipated in new treatments of the magnetic force (Chapter 8) and Faraday's law (Chapter 11).

Other changes from earlier editions include the introduction of the Dirac delta function in Chapter 2 and its use to simplify several later derivations. Orthogonal transformations are moved to an Appendix, which can be read in conjunction with Chapter 1 if desired. The del-operator notation is used for vector differentiation. All tables of data and references to other books have been updated, and SI units and notation are used systematically throughout. (Reference is also made, however, to the Gaussian units, since they are widely used in the current physics literature.) A summary section at the end of each chapter identifies key ideas and formulas, and about one hundred and thirty additional problems extend and apply the concepts.

As an aid to the reader, the more difficult problems are labeled with an asterisk. Sections and chapters of the text that are starred are not essential to its further development and may be omitted in an abbreviated study.

Dearborn, Michigan J. R. R.
Columbus, Ohio F. J. M.
Hanover, New Hampshire R. W. C.
January 1979

Contents

Starred sections and chapters may be omitted without loss of continuity

CHAPTER 1 Vector Analysis

In the study of electricity and magnetism a great saving in complexity of notation may be accomplished by using the notation of vector analysis. In providing this valuable shorthand, vector analysis also brings to the forefront the physical ideas involved in equations. It is the purpose of this chapter to give a brief but self-contained exposition of basic vector analysis and to provide the rather utilitarian knowledge of the field which is required for a treatment of electricity and magnetism. Those already familiar with vector analysis will find it a useful review and an introduction to the notation of the text.

1–1 DEFINITIONS

In the study of elementary physics several kinds of quantities have been encountered; in particular, the division into vectors and scalars has been made. For our purposes it will be sufficient to define a scalar as follows:

A scalar is a quantity that is completely characterized by its magnitude.

Examples of scalars are numerous: mass, time, volume, etc. A simple extension of the idea of a scalar is a *scalar field*, i.e., a function of position that is completely specified by its magnitude at all points in space.

A vector may be defined as follows:

A vector is a quantity that is completely characterized by its magnitude and direction.

As examples of vectors we cite position from a fixed origin, velocity, acceleration, force, etc. The generalization to a vector field gives a function of position that is completely specified by its magnitude and direction at all points in space.

These definitions may be refined and extended; in fact, in Appendix I they are replaced by more subtle definitions in terms of transformation properties. In addition, more complicated kinds of quantities, such as tensors, are sometimes encountered. Scalars and vectors will, however, largely suffice for our purposes until Chapter 22.

1-2 VECTOR ALGEBRA

Since the algebra of scalars is familiar to the reader, this algebra will be used to develop vector algebra. In order to proceed with this development it is convenient to have a representation of vectors, for which purpose we introduce a three-dimensional Cartesian coordinate system. This three-dimensional system will be denoted by the three variables x, y, z or, when it is more convenient, x_1, x_2, x_3. With respect to this coordinate system a vector is specified by its x-, y-, and z-components. Thus a vector* **V** is specified by its components V_x, V_y, V_z, where $V_x = |\mathbf{V}| \cos \alpha_1$, $V_y = |\mathbf{V}| \cos \alpha_2$, $V_z = |\mathbf{V}| \cos \alpha_3$, the α's being the angles between **V** and appropriate coordinate axes. The scalar $|\mathbf{V}| = \sqrt{V_x^2 + V_y^2 + V_z^2}$ is the *magnitude* of the vector **V**, or its *length*. In the case of vector fields, each of the components is to be regarded as a function of x, y, and z. It should be emphasized at this point that we introduce a representation of the vectors with respect to a Cartesian coordinate system only for simplicity and ease of understanding; all of the definitions and operations are, in fact, independent of any special choice of coordinates.

The sum of two vectors is defined as the vector whose components are the sums of the corresponding components of the original vectors. Thus if **C** is the sum of **A** and **B**, we write

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \tag{1-1}$$

and

$$C_x = A_x + B_x, \qquad C_y = A_y + B_y, \qquad C_z = A_z + B_z. \tag{1-2}$$

This definition of the vector sum is completely equivalent to the familiar parallelogram rule for vector addition.

Vector subtraction is defined in terms of the negative of a vector, which is the vector whose components are the negatives of the corresponding components of the original vector. Thus if **A** is a vector, $-\mathbf{A}$ is defined by

$$(-\mathbf{A})_x = -A_x, \qquad (-\mathbf{A})_y = -A_y, \qquad (-\mathbf{A})_z = -A_z. \tag{1-3}$$

The operation of subtraction is then defined as the addition of the negative. This is written

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}). \tag{1-4}$$

Since the addition of real numbers is associative and commutative, it follows that vector addition (and subtraction) is also associative and commutative. In vector notation this appears as

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C} = (\mathbf{A} + \mathbf{C}) + \mathbf{B} = \mathbf{A} + \mathbf{B} + \mathbf{C}. \tag{1-5}$$

In other words, the parentheses are not needed, as indicated by the last form.

Proceeding now to the process of multiplication, we note that the simplest product is a scalar times a vector. This operation results in a vector each component of which is the scalar times the corresponding component of the original

* Vector quantities will be denoted by boldface symbols.

vector. If c is a scalar and A a vector, the product $c\mathbf{A}$ is a vector, $\mathbf{B} = c\mathbf{A}$, defined by

$$B_x = cA_x, \qquad B_y = cA_y, \qquad B_z = cA_z. \tag{1-6}$$

It is clear that if A is a *vector field* and c a *scalar field* then **B** is a new vector field which is *not* necessarily a constant multiple of the original field

If, now, two vectors are to be multiplied, there are two possibilities, known as the vector and scalar products. Considering first the scalar product, we note that this name derives from the scalar nature of the product, although the alternative names, inner product and dot product, are sometimes used. The definition of the scalar product, written $\mathbf{A} \cdot \mathbf{B}$, is

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z \tag{1-7}$$

This definition is equivalent to another, and perhaps more familiar, definition, i.e., as the product of the magnitudes of the original vectors times the cosine of the angle between these vectors. If A and **B** are perpendicular to each other,

$$\mathbf{A} \cdot \mathbf{B} = 0.$$

The scalar product is commutative. The length of A is

$$|\mathbf{A}| = \sqrt{\mathbf{A} \cdot \mathbf{A}}.$$

The vector product of two vectors is a vector, which accounts for the name. Alternative names are outer and cross product. The vector product is written $\mathbf{A} \times \mathbf{B}$; if **C** is the vector product of A and **B**, then $\mathbf{C} = \mathbf{A} \times \mathbf{B}$, or

$$C_x = A_y B_z - A_z B_y, \qquad C_y = A_z B_x - A_x B_z, \qquad C_z = A_x B_y - A_y B_x. \tag{1-8}$$

It is important to note that the cross product depends on the order of the factors; interchanging the order introduces a minus sign:

$$\mathbf{B} \times \mathbf{A} = -\mathbf{A} \times \mathbf{B}.$$

Consequently,

$$\mathbf{A} \times \mathbf{A} = 0.$$

This definition is equivalent to the following: the vector product is the product of the magnitudes times the sine of the angle between the original vectors, with the direction given by a right-hand screw rule.*

The vector product may be easily remembered in terms of a determinant. If **i**, **j**, and **k** are unit vectors, i.e., vectors of unit magnitude, in the x-, y-, and z-directions, respectively, then

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}. \tag{1-9}$$

* Let A be rotated into **B** through the smallest possible angle A right-hand screw rotated in this manner will advance in a direction perpendicular to both A and **B**; this direction is the direction of $\mathbf{A} \times \mathbf{B}$.

If this determinant is evaluated by the usual rules, the result is precisely our definition of the cross product.

The algebraic operations discussed above may be combined in many ways. Most of the results so obtained are obvious; however, there are two triple products of sufficient importance to merit explicit mention. The triple scalar product $D = \mathbf{A} \cdot \mathbf{B} \times \mathbf{C}$ is easily found to be given by the determinant

$$D = \mathbf{A} \cdot \mathbf{B} \times \mathbf{C} = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix} = -\mathbf{B} \cdot \mathbf{A} \times \mathbf{C}. \tag{1-10}$$

This product is unchanged by an exchange of dot and cross or by a cyclic permutation of the three vectors, parentheses are not needed, since the cross product of a scalar and a vector is undefined. The other interesting triple product is the triple vector product $\mathbf{D} = \mathbf{A} \times (\mathbf{B} \times \mathbf{C})$. By a repeated application of the definition of the cross product, Eq. (1–8), we find

$$\mathbf{D} = \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}), \tag{1-11}$$

which is frequently known as the *back cab rule*. It should be noted that in the cross product the parentheses are vital; without them the product is not well defined.

At this point one might well inquire as to the possibility of vector division. Division of a vector by a scalar can, of course, be defined as multiplication by the reciprocal of the scalar. Division of a vector by another vector, however, is possible only if the two vectors are parallel. On the other hand, it is possible to write general solutions to vector equations and so accomplish something closely akin to division. Consider the equation

$$c = \mathbf{A} \cdot \mathbf{X}, \tag{1-12}$$

where c is a known scalar, $\mathbf{A}$ a known vector, and $\mathbf{X}$ an unknown vector A general solution to this equation is

$$\mathbf{X} = \frac{c\mathbf{A}}{\mathbf{A} \cdot \mathbf{A}} + \mathbf{B}, \tag{1-13}$$

where $\mathbf{B}$ is a vector of arbitrary magnitude that is perpendicular to $\mathbf{A}$, that is, $\mathbf{A} \cdot \mathbf{B} = 0$. What we have done is very nearly to divide c by $\mathbf{A}$; more correctly, we have found the general form of the vector $\mathbf{X}$ that satisfies Eq. (1–12). There is no unique solution, and this fact accounts for the vector $\mathbf{B}$. In the same fashion we may consider the vector equation

$$\mathbf{C} = \mathbf{A} \times \mathbf{X}, \tag{1-14}$$

where $\mathbf{A}$ and $\mathbf{C}$ are known vectors and $\mathbf{X}$ is an unknown vector The general solution of this equation is

$$\mathbf{X} = \frac{\mathbf{C} \times \mathbf{A}}{\mathbf{A} \cdot \mathbf{A}} + k\mathbf{A} \tag{1-15}$$

if $\mathbf{C} \cdot \mathbf{A} = 0$, where k is an arbitrary scalar. If $\mathbf{C} \cdot \mathbf{A} \neq 0$, no solution exists. This again is very nearly the quotient of $\mathbf{C}$ by $\mathbf{A}$; the scalar k takes account of the nonuniqueness of the process. If $\mathbf{X}$ is required to satisfy both (1-12) and (1-14), then the result is unique (if it exists) and given by

$$\mathbf{X} = \frac{\mathbf{C} \times \mathbf{A}}{\mathbf{A} \cdot \mathbf{A}} + \frac{c\mathbf{A}}{\mathbf{A} \cdot \mathbf{A}}. \tag{1-16}$$

1-3 GRADIENT

The extensions of the ideas introduced above to differentiation and integration, i.e., vector calculus, will now be considered. The simplest of these is the relation of a particular vector field to the derivatives of a scalar field. It is convenient first to introduce the idea of the *directional derivative* of a function of several variables. This is just the rate of change of the function in a specified direction. The directional derivative of a scalar function φ is usually denoted by $d\varphi/ds$; it must be understood that $d\mathbf{s}$ represents an infinitesimal displacement in the direction being considered, and that ds is the scalar magnitude of $d\mathbf{s}$. If $d\mathbf{s}$ has the components dx, dy, dz, then

$$\frac{d\varphi}{ds} = \lim_{\Delta s \to 0} \frac{\varphi(x + \Delta x, y + \Delta y, z + \Delta z) - \varphi(x, y, z)}{\Delta s}$$

$$= \frac{\partial\varphi}{\partial x}\frac{dx}{ds} + \frac{\partial\varphi}{\partial y}\frac{dy}{ds} + \frac{\partial\varphi}{\partial z}\frac{dz}{ds}.$$

In order to clarify the idea of a directional derivative, consider a scalar function of two variables. Thus, $\varphi(x, y)$ represents a two-dimensional scalar field. We may plot φ as a function of x and y as is done in Fig. 1-1 for the function $\varphi(x, y) = x^2 + y^2$. The directional derivative at the point x_0, y_0 depends on the direction. If we choose the direction corresponding to $dy/dx = -x_0/y_0$, then we find

$$\left.\frac{d\varphi}{ds}\right|_{x_0, y_0} = \frac{\partial\varphi}{\partial x}\frac{dx}{ds} + \frac{\partial\varphi}{\partial y}\frac{dy}{ds} = \left[2x_0 - 2y_0\frac{x_0}{y_0}\right]\frac{dx}{ds} = 0. \tag{1-17a}$$

Alternatively, if we choose $dy/dx = y_0/x_0$, we find

$$\left.\frac{d\varphi}{ds}\right|_{x_0, y_0} = \left(2x_0 + 2\frac{y_0^2}{x_0}\right)\sqrt{\frac{x_0^2}{x_0^2 + y_0^2}} = 2\sqrt{x_0^2 + y_0^2}, \tag{1-17b}$$

since $ds = \sqrt{(dx)^2 + (dy)^2}$. As a third possibility choose $dy/dx = \alpha$; then

$$\left.\frac{d\varphi}{ds}\right|_{x_0, y_0} = (2x_0 + 2\alpha y_0)(1 + \alpha^2)^{-1/2}. \tag{1-17c}$$

If this result is differentiated with respect to α and the derivative set equal to zero, then the value of α for which the derivative is a maximum or minimum is found. When we perform these operations, we obtain $\alpha = y_0/x_0$, which simply means

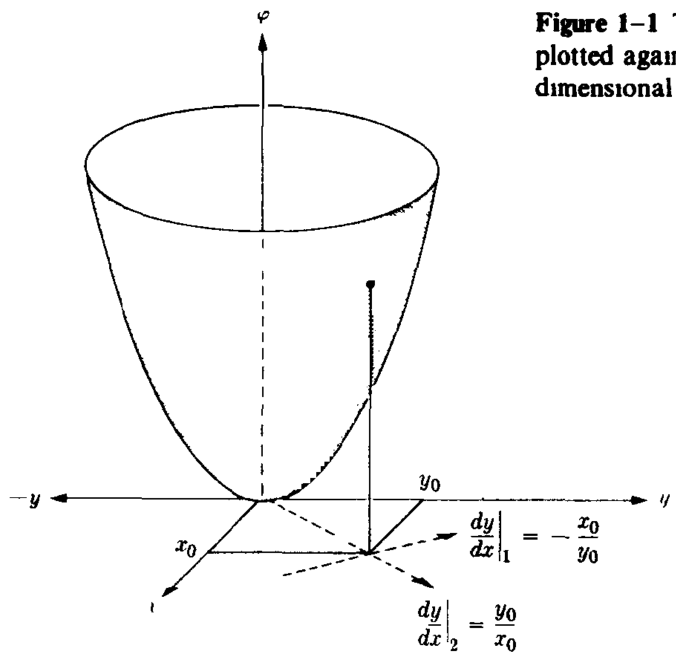

Figure 1–1 The function $\varphi(x, y) = x^2 + y^2$ plotted against x and y in a three-dimensional graph.

that the direction of maximum rate of change of the function $\varphi = x^2 + y^2$ is the radial direction If the direction is radially outward then the maximum is the maximum rate of increase; if it is radially inward it is a maximum rate of decrease or minimum rate of increase. In the direction specified by $dy/dx = -x_0/y_0$ the rate of change of $x^2 + y^2$ is zero. This direction is tangent to the circle $x^2 + y^2 = x_0^2 + y_0^2$. Clearly, on this curve, $\varphi = x^2 + y^2$ does not change. The direction in which $d\varphi/ds$ vanishes gives the direction of the curve $\varphi = constant$ through the point being considered These lines, which are circles for the function $x^2 + y^2$, are completely analogous to the familiar contour lines or lines of constant altitude which appear on topographic maps. Figure 1–2 shows the function $\varphi = x^2 + y^2$ replotted as a contour map.

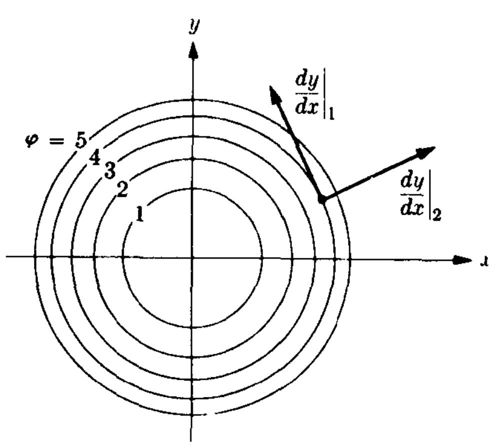

Figure 1–2 The function $\varphi(x, y)$ of Fig. 1–1 expressed as a contour map in two dimensions