INTERNATIONAL SERIES IN PURE AND APPLIED PHYSICS
LEONARD I. SCHIFF, CONSULTING EDITOR

Adler, Bazin, and Schiffer Introduction to General Relativity
Allis and Herlin Thermodynamics and Statistical Mechanics
Becker Introduction to Theoretical Mechanics
Bjorken and Drell Relativistic Quantum Mechanics
Bjorkin and Drell Relativistic Quantum Fields
Chodorow and Susskind Fundamentals of Microwave Electronics
Clark Applied X-rays
Collin Field Theory of Guided Waves
Evans The Atomic Nucleus
Feynman and Hibbs Quantum Mechanics and Path Integrals
Ginzton Microwave Measurements
Green Nuclear Physics
Gurney Introduction to Statistical Mechanics
Hall Introduction to Electron Microscopy
Hardy and Perrin The Principles of Optics
Harnwell Electricity and Electromagnetism
Harnwell and Livingood Experimental Atomic Physics
Henley and Thirring Elementary Quantum Field Theory
Houston Principles of Mathematical Physics
Kennard Kinetic Theory of Gases
Lane Superfluid Physics
Leighton Principles of Modern Physics
Lindsay Mechanical Radiation
Livingston and Blewett Particle Accelerators
Middleton An Introduction to Statistical Communication Theory
Morse Vibration and Sound
Morse and Feshbach Methods of Theoretical Physics
Muskat Physical Principles of Oil Production
Present Kinetic Theory of Gases
Read Dislocations in Crystals
Richtmyer, Kennard, and Lauritsen Introduction to Modern Physics
Schiff Quantum Mechanics
Seitz The Modern Theory of Solids
Slater Introduction to Chemical Physics
Slater Quantum Theory of Atomic Structure, Vol. I
Slater Quantum Theory of Atomic Structure, Vol. II
Slater Quantum Theory of Matter
Slater Quantum Theory of Molecules and Solids, Vol. 1
Slater Quantum Theory of Molecules and Solids, Vol. 2
Slater and Frank Electromagnetism
Slater and Frank Introduction to Theoretical Physics
Slater and Frank Mechanics
Smythe Static and Dynamic Electricity
Stratton Electromagnetic Theory
Tinkham Group Theory and Quantum Mechanics
Townes and Schawlow Microwave Spectroscopy
White Introduction to Atomic Spectra

The late F. K. Richtmyer was Consulting Editor of the series from its inception in 1929 to his death in 1939. Lee A. DuBridge was Consulting Editor from 1939 to 1946; and G. P. Harnwell from 1947 to 1954.

ELECTROMAGNETIC THEORY

BY

JULIUS ADAMS STRATTON
Professor of Physics
Massachusetts Institute of Technology

McGRAW-HILL BOOK COMPANY
NEW YORK AND LONDON
1941

ELECTROMAGNETIC THEORY

PRINTED IN THE UNITED STATES OF AMERICA

323334353637383940 VBVB 89876543

ISBN 07-062150-0

PREFACE

The pattern set nearly 70 years ago by Maxwell's *Treatise on Electricity and Magnetism* has had a dominant influence on almost every subsequent English and American text, persisting to the present day. The *Treatise* was undertaken with the intention of presenting a connected account of the entire known body of electric and magnetic phenomena from the single point of view of Faraday. Thus it contained little or no mention of the hypotheses put forward on the Continent in earlier years by Riemann, Weber, Kirchhoff, Helmholtz, and others. It is by no means clear that the complete abandonment of these older theories was fortunate for the later development of physics. So far as the purpose of the *Treatise* was to disseminate the ideas of Faraday, it was undoubtedly fulfilled; as an exposition of the author's own contributions, it proved less successful. By and large, the theories and doctrines peculiar to Maxwell—the concept of displacement current, the identity of light and electromagnetic vibrations—appeared there in scarcely greater completeness and perhaps in a less attractive form than in the original memoirs. We find that all of the first volume and a large part of the second deal with the stationary state. In fact only a dozen pages are devoted to the general equations of the electromagnetic field, 18 to the propagation of plane waves and the electromagnetic theory of light, and a score more to magnetooptics, all out of a total of 1,000. The mathematical completeness of potential theory and the practical utility of circuit theory have influenced English and American writers in very nearly the same proportion since that day. Only the original and solitary genius of Heaviside succeeded in breaking away from this course.

For an exploration of the fundamental content of Maxwell's equations one must turn again to the Continent. There the work of Hertz, Poincaré, Lorentz, Abraham, and Sommerfeld, together with their associates and successors, has led to a vastly deeper understanding of physical phenomena and to industrial developments of tremendous proportions.

The present volume attempts a more adequate treatment of variable electromagnetic fields and the theory of wave propagation. Some attention is given to the stationary state, but for the purpose of introducing fundamental concepts under simple conditions, and always with a view to later application in the general case. The reader must possess a general knowledge of electricity and magnetism such as may be acquired from an elementary course based on the experimental laws of Coulomb, Ampère, and Faraday, followed by an intermediate course dealing with the more general properties of circuits, with thermionic and electronic devices, and with the elements of electromagnetic machinery, terminating in a formulation of Maxwell's equations. This book takes up at that point. The first chapter contains a general statement of the equations governing fields and potentials, a review of the theory of units, reference material on curvilinear coordinate systems and the elements of tensor analysis, concluding with a formulation of the field equations in a space-time continuum. The second chapter is also general in character, and much of it may be omitted on a first reading. Here one will find a discussion of fundamental field properties that may be deduced without reference to particular coordinate systems. A dimensional analysis of Maxwell's equations leads to basic definitions of the vectors **E** and **B**, and an investigation of the energy relations results in expressions for the mechanical force exerted on elements of charge, current, and neutral matter. In this way a direct connection is established between observable forces and the vectors employed to describe the structure of a field.

In Chaps. III and IV stationary fields are treated as particular cases of the dynamic field equations. The subject of wave propagation is taken up first in Chap. V, which deals with homogeneous plane waves. Particular attention is given to the methods of harmonic analysis, and the problem of dispersion is considered in some detail. Chapters VI and VII treat the propagation of cylindrical and spherical waves in unbounded spaces. A necessary amount of auxiliary material on Bessel functions and spherical harmonics is provided, and consideration is given to vector solutions of the wave equation. The relation of the field to its source, the general theory of radiation, and the outlines of the Kirchhoff-Huygens diffraction theory are discussed in Chap. VIII.

Finally, in Chap. IX, we investigate the effect of plane, cylindrical, and spherical surfaces on the propagation of electromagnetic fields. This chapter illustrates, in fact, the application of the general theory established earlier to problems of practical interest. The reader will find here the more important laws of physical optics, the basic theory governing the propagation of waves along cylindrical conductors, a discussion of cavity oscillations, and an outline of the theory of wave propagation over the earth's surface.

It is regrettable that numerical solutions of special examples could not be given more frequently and in greater detail. Unfortunately the demands on space in a book covering such a broad field made this impractical. The primary objective of the book is a sound exposition of electromagnetic theory, and examples have been chosen with a view to illustrating its principles. No pretense is made of an exhaustive treatment of antenna design, transmission-line characteristics, or similar topics of engineering importance. It is the author's hope that the present volume will provide the fundamental background necessary for a critical appreciation of original contributions in special fields and satisfy the needs of those who are unwilling to accept engineering formulas without knowledge of their origin and limitations.

Each chapter, with the exception of the first two, is followed by a set of problems. There is only one satisfactory way to study a theory, and that is by application to specific examples. The problems have been chosen with this in mind, but they cover also many topics which it was necessary to eliminate from the text. This is particularly true of the later chapters. Answers or references are provided in most cases.

This book deals solely with large-scale phenomena. It is a sore temptation to extend the discussion to that fruitful field which Frenkel terms the "quasi-microscopic state," and to deal with the many beautiful results of the classical electron theory of matter. In the light of contemporary developments, anyone attempting such a program must soon be overcome with misgivings. Although many laws of classical electrodynamics apply directly to submicroscopic domains, one has no basis of selection. The author is firmly convinced that the transition must be made from quantum electrodynamics toward classical theory, rather than in the reverse direction. Whatever form the equations of quantum electrodynamics ultimately assume, their statistical average over large numbers of atoms must lead to Maxwell's equations.

The m.k.s. system of units has been employed exclusively. There is still the feeling among many physicists that this system is being forced upon them by a subversive group of engineers. Perhaps it is, although it was Maxwell himself who first had the idea. At all events, it is a good system, easily learned, and one that avoids endless confusion in practical applications. At the moment there appears to be no doubt of its universal adoption in the near future. Help for the tories among us who hold to the Gaussian system is offered on page 241.

In contrast to the stand taken on the m.k.s. system, the author has no very strong convictions on the matter of rationalized units. Rationalized units have been employed because Maxwell's equations are taken as the starting point rather than Coulomb's law, and it seems reasonable to make the point of departure as simple as possible. As a result of this choice all equations dealing with energy or wave propagation are free from the factor 4π. Such relations are becoming of far greater practical importance than those expressing the potentials and field vectors in terms of their sources.

The use of the time factor $e^{-i\omega t}$ instead of $e^{+i\omega t}$ is another point of mild controversy. This has been done because the time factor is invariably discarded, and it is somewhat more convenient to retain the positive exponent e^{+ikR} for a positive traveling wave. To reconcile any formula with its engineering counterpart, one need only replace $-i$ by $+j$.

The author has drawn upon many sources for his material and is indebted to his colleagues in both the departments of physics and of electrical engineering at the Massachusetts Institute of Technology. Thanks are expressed particularly to Professor M. F. Gardner whose advice on the practical aspects of Laplace transform theory proved invaluable, and to Dr. S. Silver who read with great care a part of the manuscript. In conclusion the author takes this occasion to express his sincere gratitude to Catherine N. Stratton for her constant encouragement during the preparation of the manuscript and untiring aid in the revision of proof.

JULIUS ADAMS STRATTON.

CAMBRIDGE, MASS.,
January, 1941.

CONTENTS

PAGE

CHAPTER II

STRESS AND ENERGY

CHAPTER III

THE ELECTROSTATIC FIELD

PAGE

CHAPTER IV

THE MAGNETOSTATIC FIELD

CHAPTER V

PLANE WAVES IN UNBOUNDED, ISOTROPIC MEDIA

CHAPTER VI

CYLINDRICAL WAVES

CHAPTER VII

SPHERICAL WAVES

ELECTROMAGNETIC THEORY

CHAPTER I

THE FIELD EQUATIONS

A vast wealth of experimental evidence accumulated over the past century leads one to believe that large-scale electromagnetic phenomena are governed by Maxwell's equations. Coulomb's determination of the law of force between charges, the researches of Ampère on the interaction of current elements, and the observations of Faraday on variable fields can be woven into a plausible argument to support this view. The historical approach is recommended to the beginner, for it is the simplest and will afford him the most immediate satisfaction. In the present volume, however, we shall suppose the reader to have completed such a preliminary survey and shall credit him with a general knowledge of the experimental facts and their theoretical interpretation. Electromagnetic theory, according to the standpoint adopted in this book, is the theory of Maxwell's equations. Consequently, we shall postulate these equations at the outset and proceed to deduce the structure and properties of the field together with its relation to the source. No single experiment constitutes proof of a theory. The true test of our initial assumptions will appear in the persistent, uniform correspondence of deduction with observation.

In this first chapter we shall be occupied with the rather dry business of formulating equations and preparing the way for our investigation.

MAXWELL'S EQUATIONS

1.1. The Field Vectors.—By an electromagnetic field let us understand the domain of the four vectors **E** and **B, D** and **H.** These vectors are assumed to be finite throughout the entire field, and at all ordinary points to be continuous functions of position and time, with continuous derivatives. Discontinuities in the field vectors or their derivatives may occur, however, on surfaces which mark an abrupt change in the physical properties of the medium. According to the traditional usage, **E** and **H** are known as the intensities respectively of the electric and magnetic field, **D** is called the electric displacement and **B,** the magnetic induction. Eventually the field vectors must be defined in terms of the experiments by which they can be measured. Until these experiments are formulated, there is no reason to consider one vector more fundamental than another, and we shall apply the word intensity to mean indiscriminately the strength or magnitude of any of the four vectors at a point in space and time.

The source of an electromagnetic field is a distribution of electric charge and current. Since we are concerned only with its macroscopic effects, it may be assumed that this distribution is continuous rather than discrete, and specified as a function of space and time by the density of charge ρ, and by the vector current density $\mathbf{J}$.

We shall now *postulate* that at every ordinary point in space the field vectors are subject to the Maxwell equations:

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \tag{1}$$

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J}. \tag{2}$$

By an ordinary point we shall mean one in whose neighborhood the physical properties of the medium are continuous. It has been noted that the transition of the field vectors and their derivatives across a surface bounding a material body may be discontinuous; such surfaces must, therefore, be excluded until the nature of these discontinuities can be investigated.

1.2. Charge and Current.—Although the corpuscular nature of electricity is well established, the size of the elementary quantum of charge is too minute to be taken into account as a distinct entity in a strictly macroscopic theory. Obviously the frontier that marks off the domain of large-scale phenomena from those which are microscopic is an arbitrary one. To be sure, a macroscopic element of volume must contain an enormous number of atoms; but that condition alone is an insufficient criterion, for many crystals, including the metals, exhibit frequently a microscopic "grain" or "mosaic" structure which will be excluded from our investigation. We are probably well on the safe side in imposing a limit of one-tenth of a millimeter as the smallest admissible element of length. There are many experiments, such as the scattering of light by particles no larger than 10^{-3} mm. in diameter, which indicate that the macroscopic theory may be pushed well beyond the limit suggested. Nonetheless, we are encroaching here on the proper domain of quantum theory, and it is the quantum theory which must eventually determine the validity of our assumptions in microscopic regions.

Let us suppose that the charge contained within a volume element Δv is Δq. The *charge density* at any point within Δv will be defined by the relation

$$\Delta q = \rho\, \Delta v. \tag{3}$$

Thus by the charge density at a point we mean the average charge per unit volume in the neighborhood of that point. In a strict sense (3) does not define a continuous function of position, for Δv cannot approach zero without limit. Nonetheless we shall assume that ρ can be represented by a function of the coordinates and the time which at ordinary points is continuous and has continuous derivatives. The value of the total charge obtained by integrating that function over a large-scale volume will then differ from the true charge contained therein by a microscopic quantity at most.

Any ordered motion of charge constitutes a current. A current distribution is characterized by a vector field which specifies at each point not only the intensity of the flow but also its direction. As in the study of fluid motion, it is convenient to imagine streamlines traced through the distribution and everywhere tangent to the direction of flow. Consider a surface which is orthogonal to a system of streamlines. The *current density* at any point on this surface is then defined as a vector $\mathbf{J}$ directed along the streamline through the point and equal in magnitude to the charge which in unit time crosses unit area of the surface in the vicinity of the point. On the other hand the current I across *any* surface S is equal to the rate at which charge crosses that surface. If $\mathbf{n}$ is the positive unit normal to an element Δa of S, we have

$$\Delta I = \mathbf{J} \cdot \mathbf{n}\, \Delta a. \tag{4}$$

Since Δa is a macroscopic element of area, Eq. (4) does not define the current density with mathematical rigor as a continuous function of position, but again one may represent the distribution by such a function without incurring an appreciable error. The total current through S is, therefore,

$$I = \int_S \mathbf{J} \cdot \mathbf{n}\, da. \tag{5}$$

Since electrical charge may be either positive or negative, a convention must be adopted as to what constitutes a positive current. If the flow through an element of area consists of positive charges whose velocity vectors form an angle of less than 90 deg. with the positive normal $\mathbf{n}$, the current is said to be positive. If the angle is greater than 90 deg., the current is negative. Likewise if the angle is less than 90 deg. but the charges are negative, the current through the element is negative. In the case of metallic conductors the carriers of electricity are presumably negative electrons, and the direction of the current density vector is therefore opposed to the direction of electron motion.

Let us suppose now that the surface S of Eq. (5) is closed. We shall adhere to the customary convention that *the positive normal to a closed* *surface is drawn outward.* In virtue of the definition of current as the flow of charge across a surface, it follows that the surface integral of the normal component of **J** over S must measure the loss of charge from the region within. There is no experimental evidence to indicate that under ordinary conditions charge may be either created or destroyed in macroscopic amounts. One may therefore write

$$\int_S \mathbf{J} \cdot \mathbf{n}\, da = -\frac{d}{dt}\int_V \rho\, dv, \tag{6}$$

where V is the volume enclosed by S, as a relation expressing the *conservation of charge.* The flow of charge across the surface can originate in two ways. The surface S may be fixed in space and the density ρ be some function of the time as well as of the coordinates; or the charge density may be invariable with time, while the surface moves in some prescribed manner. In this latter event the right-hand integral of (6) is a function of time in virtue of variable limits. If, however, the surface is fixed and the integral convergent, one may replace d/dt by a partial derivative under the sign of integration.

$$\int_S \mathbf{J} \cdot \mathbf{n}\, da = -\int_V \frac{\partial \rho}{\partial t}\, dv. \tag{7}$$

We shall have frequent occasion to make use of the divergence theorem of vector analysis. Let $\mathbf{A}(x, y, z)$ be any vector function of position which together with its first derivatives is continuous throughout a volume V and over the bounding surface S. The surface S is regular but otherwise arbitrary.[1] Then it can be shown that

$$\int_S \mathbf{A} \cdot \mathbf{n}\, da = \int_V \nabla \cdot \mathbf{A}\, dv. \tag{8}$$

As a matter of fact, this relation may be advantageously used as a definition of the divergence. To obtain the value of $\nabla \cdot \mathbf{A}$ at a point P within V, we allow the surface S to shrink about P. When the volume V has become sufficiently small, the integral on the right may be replaced by $V\nabla \cdot \mathbf{A}$, and we obtain

$$\nabla \cdot \mathbf{A} = \lim_{S \to 0} \frac{1}{V}\int_S \mathbf{A} \cdot \mathbf{n}\, da. \tag{9}$$

[1] A regular element of arc is represented in parametric form by the equations $x = x(t)$, $y = y(t)$, $z = z(t)$ such that in the interval $a \leq t \leq b$ x, y, z are continuous, single-valued functions of t with continuous derivatives of all orders unless otherwise restricted. A regular curve is constructed of a finite number of such arcs joined end to end but such that the curve does not cross itself. Thus a regular curve has no double points and is piecewise differentiable. A regular surface element is a portion of surface whose projection on a properly oriented plane is the interior of a regular closed curve. Hence it does not intersect itself. *Cf.* Kellogg, "Foundations of Potential Theory," p. 97, Springer, 1929.

The divergence of a vector at a point is, therefore, to be interpreted as the integral of its normal component over an infinitesimally small surface enclosing that point, divided by the enclosed volume. The flux of a vector through a closed surface is a measure of the sources within; hence the divergence determines their strength at a point. Since S has been shrunk close about P, the value of **A** at every point on the surface may be expressed analytically in terms of the values of **A** and its derivatives at P, and consequently the integral in (9) may be evaluated, leading in the case of rectangular coordinates to

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}. \tag{10}$$

On applying this theorem to (7) the surface integral is transformed to the volume integral

$$\int_V \left(\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t}\right) dv = 0. \tag{11}$$

Now the integrand of (11) is a continuous function of the coordinates and hence there must exist small regions within which the integrand does not change sign. If the integral is to vanish for arbitrary volumes V, it is necessary that the integrand be identically zero. The differential equation

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0 \tag{12}$$

expresses the conservation of charge in the neighborhood of a point. By analogy with an equivalent relation in hydrodynamics, (12) is frequently referred to as the *equation of continuity.*

If at every point within a specified region the charge density is constant, the current passing into the region through the bounding surface must at all times equal the current passing outward. Over the bounding surface S we have

$$\int_S \mathbf{J} \cdot \mathbf{n}\, da = 0, \tag{13}$$

and at every interior point

$$\nabla \cdot \mathbf{J} = 0. \tag{14}$$

Any motion characterized by vector or scalar quantities which are independent of the time is said to be steady, or stationary. A steady-state flow of electricity is thus defined by a vector **J** which at every point within the region is constant in direction and magnitude. In virtue of the divergenceless character of such a current distribution, it follows that in the steady state all streamlines, or current filaments, close upon themselves. The field of the vector **J** is solenoidal.

1.3. Divergence of the Field Vectors.—Two further conditions satisfied by the vectors **B** and **D** may be deduced directly from Maxwell's equations by noting that the divergence of the curl of any vector vanishes identically. We take the divergence of Eq. (1) and obtain

$$\nabla \cdot \frac{\partial \mathbf{B}}{\partial t} = \frac{\partial}{\partial t} \nabla \cdot \mathbf{B} = 0. \tag{15}$$

The commutation of the operators ∇ and $\partial/\partial t$ is admissible, for at an ordinary point **B** and all its derivatives are assumed to be continuous. It follows from (15) that at every point in the field the divergence of **B** is constant. If ever in its past history the field has vanished, this constant must be zero and, since one may reasonably suppose that the initial generation of the field was at a time not infinitely remote, we conclude that

$$\nabla \cdot \mathbf{B} = 0, \tag{16}$$

and the field of **B** is therefore solenoidal.

Likewise the divergence of Eq. (2) leads to

$$\nabla \cdot \mathbf{J} + \frac{\partial}{\partial t} \nabla \cdot \mathbf{D} = 0, \tag{17}$$

or, in virtue of (12), to

$$\frac{\partial}{\partial t} (\nabla \cdot \mathbf{D} - \rho) = 0. \tag{18}$$

If again we admit that at some time in its past or future history the field may vanish, it is necessary that

$$\nabla \cdot \mathbf{D} = \rho. \tag{19}$$

The charges distributed with a density ρ constitute the sources of the vector **D**.

The divergence equations (16) and (19) are frequently included as part of Maxwell's system. It must be noted, however, that if one assumes the conservation of charge, these are not independent relations.

1.4. Integral Form of the Field Equations.—The properties of an electromagnetic field which have been specified by the differential equations (1), (2), (16), and (19) may also be expressed by an equivalent system of integral relations. To obtain this equivalent system, we apply a second fundamental theorem of vector analysis.

According to Stokes' theorem the line integral of a vector taken about a closed contour can be transformed into a surface integral extended

over a surface bounded by the contour. The contour C must either be regular or be resolvable into a finite number of regular arcs, and it is assumed that the otherwise arbitrary surface S bounded by C is two-sided and may be resolved into a finite number of regular elements. The positive side of the surface S is related to the positive direction of circulation on the contour by the usual convention that an observer, moving in a positive sense along C, will have the positive side of S on his left. Then if $\mathbf{A}(x, y, z)$ is any vector function of position, which together with its first derivatives is continuous at all points of S and C, it may be shown that

$$\int_C \mathbf{A} \cdot d\mathbf{s} = \int_S (\nabla \times \mathbf{A}) \cdot \mathbf{n}\, da, \tag{20}$$

where $d\mathbf{s}$ is an element of length along C and $\mathbf{n}$ is a unit vector normal to the positive side of the element of area da. This transformation can also be looked upon as an equation defining the curl. To determine the value of $\nabla \times \mathbf{A}$ at a point P on S, we allow the contour to shrink about P until the enclosed area S is reduced to an infinitesimal element of a plane whose normal is in the direction specified by $\mathbf{n}$. The integral on the right is then equal to $(\nabla \times \mathbf{A}) \cdot \mathbf{n}S$, plus infinitesimals of higher order. The projection of the vector $\nabla \times \mathbf{A}$ in the direction of the normal is, therefore,

$$(\nabla \times \mathbf{A}) \cdot \mathbf{n} = \lim_{C \to 0} \frac{1}{S} \int_C \mathbf{A} \cdot d\mathbf{s}. \tag{21}$$

The curl of a vector at a point is to be interpreted as the line integral of that vector about an infinitesimal path on a surface containing the point, per unit of enclosed area. Since **A** has been assumed analytic in the neighborhood of P, its value at any point on C may be expressed in terms of the values of **A** and its derivatives at P, so that the evaluation of the line integral in (21) about the infinitesimal path can actually be carried out. In particular, if the element S is oriented parallel to the yz-coordinate plane, one finds for the x-component of the curl

$$(\nabla \times \mathbf{A})_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}. \tag{22}$$

Proceeding likewise for the y- and z-components we obtain

$$\nabla \times \mathbf{A} = \mathbf{i}\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) + \mathbf{j}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) + \mathbf{k}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} \tag{23}$$

Let us now integrate the normal component of the vector $\partial\mathbf{B}/\partial t$ over any regular surface S bounded by a closed contour C. From (1) and (20) it follows that

$$\int_C \mathbf{E}\cdot d\mathbf{s} + \int_S \frac{\partial\mathbf{B}}{\partial t}\cdot\mathbf{n}\,da = 0. \tag{24}$$

If the contour is fixed, the operator $\partial/\partial t$ may be brought out from under the sign of integration.

$$\int_C \mathbf{E}\cdot d\mathbf{s} = -\frac{\partial}{\partial t}\int_S \mathbf{B}\cdot\mathbf{n}\,da. \tag{25}$$

By definition, the quantity

$$\Phi = \int_S \mathbf{B}\cdot\mathbf{n}\,da \tag{26}$$

is the magnetic flux, or more specifically the flux of the vector $\mathbf{B}$ through the surface. According to (25) the line integral of the vector $\mathbf{E}$ about any closed, regular curve in the field is equal to the time rate of decrease of the magnetic flux through any surface spanning that curve. The relation between the direction of circulation about a contour and the positive normal to a surface bounded by it is illustrated in Fig. 1. A positive direction about C is chosen arbitrarily and the flux Φ is then positive or negative according to the direction of the lines of $\mathbf{B}$ with respect to the normal. The time rate of change of Φ is in turn positive or negative as the positive flux is increasing or decreasing.

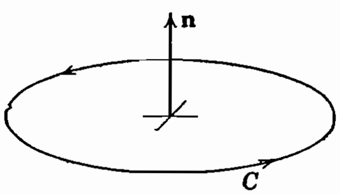

FIG. 1.—Convention relating direction of the positive normal $\mathbf{n}$ to the direction of circulation about a contour C.

We recall that the application of Stokes' theorem to Eq. (1) is valid only if the vector $\mathbf{E}$ and its derivatives are continuous at all points of S and C. Since discontinuities in both $\mathbf{E}$ and $\mathbf{B}$ occur across surfaces marking sudden changes in the physical properties of the medium, the question may be raised as to what extent (25) represents a general law of the electromagnetic field. One might suppose, for example, that the contour linked or pierced a closed iron transformer core. To obviate this difficulty it may be imagined that at the surface of every material body in the field the physical properties vary rapidly but *continuously* within a thin boundary layer from their values just inside to their values just outside the surface. In this manner all discontinuities are eliminated from the field and (25) may be applied to every closed contour.

The experiments of Faraday indicated that the relation (25) holds whatever the cause of flux variation. The partial derivative implies a variable flux density threading a fixed contour, but the total flux can likewise be changed by a deformation of the contour. To take this into account the Faraday law is written generally in the form

$$\int_C \mathbf{E}\cdot d\mathbf{s} = -\frac{d}{dt}\int_S \mathbf{B}\cdot\mathbf{n}\,da. \tag{27}$$

It can be shown that (27) is in fact a consequence of the differential field equations, but the proof must be based on the electrodynamics of moving bodies which will be touched upon in Sec. 1.22.

In like fashion Eq. (2) may be replaced by an equivalent integral relation,

$$\int_C \mathbf{H}\cdot d\mathbf{s} = I + \frac{d}{dt}\int_S \mathbf{D}\cdot\mathbf{n}\,da. \tag{28}$$

where I is the total current linking the contour as defined in (5). In the steady state, the integral on the right is zero and the conduction current I through any regular surface is equal to the line integral of the vector $\mathbf{H}$ about its contour. If, however, the field is variable, the vector $\partial\mathbf{D}/\partial t$ has associated with it a field $\mathbf{H}$ exactly equal to that which would be produced by a current distribution of density

$$\mathbf{J}' = \frac{\partial\mathbf{D}}{\partial t}. \tag{29}$$

To this quantity Maxwell gave the name "displacement current," a term which we shall occasionally employ without committing ourselves as yet to any particular interpretation of the vector $\mathbf{D}$.

The two remaining field equations (16) and (19) can be expressed in an equivalent integral form with the help of the divergence theorem. One obtains

$$\oint_S \mathbf{B}\cdot\mathbf{n}\,da = 0, \tag{30}$$

stating that the total flux of the vector $\mathbf{B}$ crossing any closed, regular surface is zero, and

$$\oint_S \mathbf{D}\cdot\mathbf{n}\,da = \int_V \rho\,dv = q, \tag{31}$$

according to which the flux of the vector $\mathbf{D}$ through a closed surface is equal to the total charge q contained within. The circle through the sign of integration is frequently employed to emphasize the fact that a contour or surface is closed.

MACROSCOPIC PROPERTIES OF MATTER

1.5. The Inductive Capacities ε and μ.—No other assumptions have been made thus far than that an electromagnetic field may be characterized by four vectors **E**, **B**, **D**, and **H**, which at ordinary points satisfy Maxwell's equations, and that the distribution of current which gives rise to this field is such as to ensure the conservation of charge. Between the five vectors **E**, **B**, **D**, **H**, **J** there are but two independent relations, the equations (1) and (2) of the preceding section, and we are therefore obliged to impose further conditions if the system is to be made determinate.

Let us begin with the assumption that at any given point in the field, whether in free space or within matter, the vector **D** may be represented as a function of **E** and the vector **H** as a function of **B**.

$$\mathbf{D} = D(\mathbf{E}), \qquad \mathbf{H} = H(\mathbf{B}). \tag{1}$$

The nature of these functional relations is to be determined solely by the physical properties of the medium in the immediate neighborhood of the specified point. Certain simple relations are of most common occurrence.

1. In *free space*, **D** differs from **E** only by a constant factor, as does **H** from **B**. Following the traditional usage, we shall write

$$\mathbf{D} = \epsilon_0 \mathbf{E}, \qquad \mathbf{H} = \frac{1}{\mu_0}\mathbf{B}. \tag{2}$$

The values and the dimensions of the constants ϵ_0 and μ_0 will depend upon the system of units adopted. In only one of many wholly arbitrary systems does **D** reduce to **E** and **H** to **B** in empty space.

2. If the physical properties of a body in the neighborhood of some interior point are the same in all directions, the body is said to be *isotropic*. At every point in an isotropic medium **D** is parallel to **E** and **H** is parallel to **B**. The relations between the vectors, moreover, are *linear* in almost all the soluble problems of electromagnetic theory. For the isotropic, linear case we put then

$$\mathbf{D} = \epsilon \mathbf{E}, \qquad \mathbf{H} = \frac{1}{\mu}\mathbf{B}. \tag{3}$$

The factors ϵ and μ will be called the inductive capacities of the medium. The dimensionless ratios

$$\kappa_e = \frac{\epsilon}{\epsilon_0}, \qquad \kappa_m = \frac{\mu}{\mu_0}, \tag{4}$$

are independent of the choice of units and will be referred to as the specific inductive capacities. The properties of a *homogeneous* medium are constant from point to point and in this case it is customary to refer to κ_e as the dielectric constant and to κ_m as the permeability. In general, however, one must look upon the inductive capacities as scalar functions of position which characterize the electromagnetic properties of matter in the large.

3. The properties of *anisotropic* matter vary in a different manner along different directions about a point. In this case the vectors **D** and **E**, **H** and **B** are parallel only along certain preferred axes. If it may be assumed that the relations are still linear, as is usually the case, one may express each rectangular component of **D** as a linear function of the three components of **E**.

$$\begin{aligned} D_x &= \epsilon_{11}E_x + \epsilon_{12}E_y + \epsilon_{13}E_z, \\ D_y &= \epsilon_{21}E_x + \epsilon_{22}E_y + \epsilon_{23}E_z, \\ D_z &= \epsilon_{31}E_x + \epsilon_{32}E_y + \epsilon_{33}E_z. \end{aligned} \tag{5}$$

The coefficients ϵ_{jk} of this linear transformation are the components of a symmetric tensor. An analogous relation may be set up between the vectors **H** and **B**, but the occurrence of such a linear anisotropy in what may properly be called macroscopic problems is rare.

The distinction between the microscopic and macroscopic viewpoints is nowhere sharper than in the interpretation of these parameters ϵ and μ, or their tensor equivalents. A microscopic theory must deduce the physical properties of matter from its atomic structure. It must enable one to calculate not only the average field that prevails within a body but also its local value in the neighborhood of a specific atom. It must tell us how the atom will be deformed under the influence of that local field, and how the aggregate effect of these atomic deformations may be represented in the large by such parameters as ϵ and μ.

We, on the other hand, are from the present standpoint sheer behaviorists. Our knowledge of matter is, to use a large word, purely phenomenological. Each substance is to be characterized electromagnetically in terms of a minimum number of parameters. The dependence of the parameters ϵ and μ on such physical variables as density, temperature, and frequency will be established by experiment. Information given by such measurements sheds much light on the internal structure of matter, but the internal structure is not our present concern.

1.6. Electric and Magnetic Polarization.—To describe the electromagnetic state of a sample of matter, it will prove convenient to introduce two additional vectors. We shall *define* the electric and magnetic polarization vectors by the equations

$$\mathbf{P} = \mathbf{D} - \epsilon_0 \mathbf{E}, \qquad \mathbf{M} = \frac{1}{\mu_0}\mathbf{B} - \mathbf{H}. \tag{6}$$

The polarization vectors are thus definitely associated with matter and vanish in free space. By means of these relations let us now eliminate **D** and **H** from the field equations. There results the system

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0,$$

$$(7) \qquad \nabla \times \mathbf{B} - \epsilon_0\mu_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0\left(\mathbf{J} + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M}\right),$$

$$\nabla \cdot \mathbf{B} = 0, \qquad \nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0}(\rho - \nabla \cdot \mathbf{P}),$$

which we are free to interpret as follows: *the presence of rigid material bodies in an electromagnetic field may be completely accounted for by an equivalent distribution of charge of density* $-\nabla \cdot \mathbf{P}$, *and an equivalent distribution of current of density* $\frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M}$.

In isotropic media the polarization vectors are parallel to the corresponding field vectors, and are found experimentally to be proportional to them if ferromagnetic materials are excluded. The electric and magnetic susceptibilities χ_e and χ_m are defined by the relations

$$(8) \qquad \mathbf{P} = \chi_e\epsilon_0 \mathbf{E}, \qquad \mathbf{M} = \chi_m \mathbf{H}.$$

Logically the magnetic polarization **M** should be placed proportional to **B**. Long usage, however, has associated it with **H** and to avoid confusion on a matter which is really of no great importance we adhere to this convention. The susceptibilities χ_e and χ_m defined by (8) are dimensionless ratios whose values are independent of the system of units employed. In due course it will be shown that **E** and **B** are force vectors and in this sense are fundamental. **D** and **H** are derived vectors associated with the state of matter. The polarization vector **P** has the dimensions of **D**, not **E**, while **M** and **H** are dimensionally alike. From (3), (6), and (8) it follows at once that the susceptibilities are related to the specific inductive capacities by the equations

$$(9) \qquad \chi_e = \kappa_e - 1, \qquad \chi_m = \kappa_m - 1.$$

In anisotropic media the susceptibilities are represented by the components of a tensor.

It will be a part of our task in later chapters to formulate experiments by means of which the susceptibility of a substance may be accurately measured. Such measurements show that the electric susceptibility is always positive. In gases it is of the order of 0.0006 (air), but in liquids and solids it may attain values as large as 80 (water). An inherent difference in the nature of the vectors **P** and **M** is indicated by the fact that the magnetic susceptibility χ_m may be either positive or negative. Substances characterized by a positive susceptibility are said to be

paramagnetic, whereas those whose susceptibility is negative are called *diamagnetic*. The metals of the ferromagnetic group, including iron, nickel, cobalt, and their alloys, constitute a particular class of substances of enormous positive susceptibility, the value of which may be of the order of many thousands. In view of the nonlinear relation of **M** to **H** peculiar to these materials, the susceptibility χ_m must now be interpreted as the slope of a tangent to the **M-H** curve at a point corresponding to a particular value of **H**. To include such cases the definition of susceptibility is generalized to

$$(10) \qquad \chi_m = \frac{\partial M}{\partial H}.$$

The susceptibilities of all nonferromagnetic materials, whether paramagnetic or diamagnetic, are so small as to be negligible for most practical purposes.

Thus far it has been assumed that a functional relation exists between the vector **P** or **M** and the applied field, and for this reason they may properly be called the *induced* polarizations. Under certain conditions, however, a magnetic field may be associated with a ferromagnetic body in the absence of any external excitation. The body is then said to be in a state of permanent magnetization. We shall maintain our initial assumption that the field both inside and outside the magnet is completely defined by the vectors **B** and **H**. But now the difference of these two vectors at an interior point is a *fixed* vector $\mathbf{M}_0$, which may be called the intensity of magnetization and which bears no functional relationship to **H**. On the contrary the magnetization $\mathbf{M}_0$ must be interpreted as the source of the field. If an external field is superposed on the field of a permanent magnet, the intensity of magnetization will be augmented by the induced polarization **M**. At any interior point we have, therefore,

$$(11) \qquad \mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M} + \mathbf{M}_0).$$

Of this induced polarization we can only say for the present that it is a function of the resultant **H** prevailing at the same point. The relation of the resultant field within the body to the intensity of an applied field generated by external sources depends not only on the magnetization $\mathbf{M}_0$ but also upon the shape of the body. There will be occasion to examine this matter more carefully in Chap. IV.

1.7. Conducting Media.—To Maxwell's equations there must now be added a third and last empirical relation between the current density and the field. We shall assume that at any point within a liquid or solid the current density is a function of the field **E**.

$$(12) \qquad \mathbf{J} = \mathbf{J}(\mathbf{E}).$$

The distribution of current in an ionized, gaseous medium may depend also on the intensity of the magnetic field, but since electromagnetic phenomena in gaseous discharges are in general governed by a multitude of factors other than those taken into account in the present theory, we shall exclude such cases from further consideration.[1]

Throughout a remarkably wide range of conditions, in both solids and weakly ionized solutions, the relation (12) proves to be linear.

$$\mathbf{J} = \sigma \mathbf{E}. \tag{13}$$

The factor σ is called the *conductivity* of the medium. The distinction between good and poor conductors, or insulators, is relative and arbitrary. All substances exhibit conductivity to some degree but the range of observed values of σ is tremendous. The conductivity of copper, for example, is some 10^7 times as great as that of such a "good" conductor as sea water, and 10^{19} times that of ordinary glass. In Appendix III will be found an abbreviated table of the conductivities of representative materials.

Equation (13) is simply Ohm's law. Let us imagine, for example, a stationary distribution of current throughout the volume of any conducting medium. In virtue of the divergenceless character of the flow this distribution may be represented by closed streamlines. If a and b are two points on a particular streamline and ds is an element of its length, we have

$$\int_a^b \mathbf{E} \cdot d\mathbf{s} = \int_a^b \frac{\mathbf{J}}{\sigma} \cdot d\mathbf{s}. \tag{14}$$

A bundle of adjacent streamlines constitutes a current filament or tube. Since the flow is solenoidal, the current I through every cross section of the filament is the same. Let S be the cross-sectional area of the filament on a plane drawn normal to the direction of flow. S need not be infinitesimal, but is assumed to be so small that over its area the current density is uniform. Then $S\mathbf{J} \cdot d\mathbf{s} = I\, ds$, and

$$\int_a^b \mathbf{E} \cdot d\mathbf{s} = I \int_a^b \frac{1}{\sigma S}\, ds. \tag{15}$$

The factor,

$$R = \int_a^b \frac{1}{\sigma S}\, ds, \tag{16}$$

[1] It is true that to a very slight degree the current distribution in a liquid or solid conductor may be modified by an impressed magnetic field, but the magnitude of this so-called Hall effect is so small that it may be ignored without incurring an appreciable error.

is equal to the resistance of the filament between the points a and b. The resistance of a linear section of homogeneous conductor of uniform cross section S and length l is

$$R = \frac{1}{\sigma} \frac{l}{S}, \tag{17}$$

a formula which is strictly valid only in the case of stationary currents.

Within a region of nonvanishing conductivity there can be no permanent distribution of free charge. This fundamentally important theorem can be easily demonstrated when the medium is homogeneous and such that the relations between **D** and **E** and **J** and **E** are linear. By the equation of continuity,

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = \nabla \cdot \sigma \mathbf{E} + \frac{\partial \rho}{\partial t} = 0. \tag{18}$$

On the other hand in a homogeneous medium

$$\nabla \cdot \mathbf{E} = \frac{1}{\epsilon} \rho, \tag{19}$$

which combined with (18) leads to

$$\frac{\partial \rho}{\partial t} + \frac{\sigma}{\epsilon} \rho = 0. \tag{20}$$

The density of charge at any instant is, therefore,

$$\rho = \rho_0 e^{-\frac{\sigma}{\epsilon} t}, \tag{21}$$

the constant of integration ρ_0 being equal to the density at the time $t = 0$. The initial charge distribution throughout the conductor decays exponentially with the time at every point and in a manner wholly independent of the applied field. If the charge density is initially zero, it remains zero at all times thereafter.

The time

$$\tau = \frac{\epsilon}{\sigma} \tag{22}$$

required for the charge at any point to decay to $1/e$ of its original value is called the *relaxation time*. In all but the poorest conductors τ is exceedingly small. Thus in sea water the relaxation time is about 2×10^{-10} sec.; even in such a poor conductor as distilled water it is not greater than 10^{-6} sec. In the best insulators, such as fused quartz, it may nevertheless assume values exceeding 10^6 sec., an instance of the extraordinary range in the possible values of the parameter σ.

Let us suppose that at $t = 0$ a charge is concentrated within a small spherical region located somewhere in a conducting body. At every other point of the conductor the charge density is zero. The charge within the sphere now begins to fade away exponentially, but according to (21) no charge can reappear anywhere *within* the conductor. What becomes of it? Since the charge is conserved, the decay of charge within the spherical surface must be accompanied by an outward flow, or current. No charge can accumulate at any other interior point; hence the flow must be divergenceless. It will be arrested, however, on the outer surface of the conductor and it is here that we shall rediscover the charge that has been lost from the central sphere. This surface charge makes its appearance at the exact instant that the interior charge begins to decay, for the total charge is constant.

UNITS AND DIMENSIONS

1.8. The M.K.S. or Giorgi System.—An electromagnetic field thus far is no more than a complex of vectors subject to a postulated system of differential equations. To proceed further we must establish the physical dimensions of these vectors and agree on the units in which they are to be measured.

In the customary sense, an "absolute" system of units is one in which every quantity may be measured or expressed in terms of the three fundamental quantities mass, length, and time. Now in electromagnetic theory there is an essential arbitrariness in the matter of dimensions which is introduced with the factors ϵ_0 and μ_0 connecting **D** and **E**, **H** and **B** respectively in free space. No experiment has yet been imagined by means of which dimensions may be attributed to either ϵ_0 or μ_0 as an independent physical entity. On the other hand, it is a direct consequence of the field equations that the quantity

$$(1) \qquad c = \frac{1}{\sqrt{\epsilon_0\mu_0}}$$

shall have the dimensions of a velocity, and every arbitrary choice of ϵ_0 and μ_0 is subject to this restriction. The magnitude of this velocity cannot be calculated a priori, but by suitable experiment it may be measured. The value obtained by the method of Rosa and Dorsey of the Bureau of Standards and corrected by Curtis[1] in 1929 is

$$(2) \qquad c = \frac{1}{\sqrt{\epsilon_0\mu_0}} = 2.99790 \times 10^8 \qquad \text{meters/sec.,}$$

[1] ROSA and DORSEY, A New Determination of the Ratio of the Electrostatic Unit of Electricity, *Bur. Standards, Bull.* 3, p. 433, 1907. CURTIS, *Bur. Standards J. Research*, **3**, 63, 1929.

or for all practical purposes

$$(3) \qquad c = 3 \times 10^8 \qquad \text{meters/sec.}$$

Throughout the early history of electromagnetic theory the absolute *electromagnetic system* of units was employed for all scientific investigations. In this system the centimeter was adopted as the unit of length, the gram as the unit of mass, the second as the unit of time, and as a fourth unit the factor μ_0 was placed arbitrarily equal to unity and considered dimensionless. The dimensions of ϵ_0 were then uniquely determined by (1) and it could be shown that the units and dimensions of every other quantity entering into the theory might be expressed in terms of centimeters, grams, seconds, and μ_0. Unfortunately, this absolute system failed to meet the needs of practice. The units of resistance and of electromotive force were, for example, far too small. To remedy this defect a *practical system* was adopted. Each unit of the practical system had the dimensions of the corresponding electromagnetic unit and differed from it in magnitude by a power of ten which, in the case of voltage and resistance at least, was wholly arbitrary. The practical units have the great advantage of convenient size and they are now universally employed for technical measurements and computations. Since they have been defined as arbitrary multiples of absolute units, they do not, however, constitute an absolute system. Now the quantities mass, length, and time are fundamental solely because the physicist has found it expedient to raise them to that rank. That there are other fundamental quantities is obvious from the fact that all electromagnetic quantities cannot be expressed in terms of these three alone. The restriction of the term "absolute" to systems based on mass, length, and time is, therefore, wholly unwarranted; one should ask only that such a system be self-consistent and that every quantity be defined in terms of a minimum number of basic, independent units. The antipathy of physicists in the past to the practical system of electrical units has been based not on any firm belief in the sanctity of mass, length, and time, but rather on the lack of self-consistency within that system.

Fortunately a most satisfactory solution has been found for this difficulty. In 1901 Giorgi,[1] pursuing an idea originally due to Maxwell, called attention to the fact that the practical system could be converted into an absolute system by an appropriate choice of fundamental units. It is indeed only necessary to choose for the unit of length the inter-

[1] GIORGI: Unità Razionali di Elettromagnetismo, *Atti dell' A.E.I.*, 1901. An historical review of the development of the practical system, including a report of the action taken at the 1935 meeting of the International Electrotechnical Commission and an extensive bibliography is given by Kennelly, *J. Inst. Elec. Engrs.*, **78**, 235–245, 1936. See also GLAZEBROOK, The M.K.S. System of Electrical Units, *J. Inst. Elec. Engrs.*, **78**, pp. 245–247.

national *meter*, for the unit of mass the *kilogram*, for the unit of time the *second*, and as a fourth unit any electrical quantity belonging to the practical system such as the coulomb, the ampere, or the ohm. From the field equations it is then possible to deduce the units and dimensions of every electromagnetic quantity in terms of these four fundamental units. Moreover the derived quantities will be related to each other exactly as in the practical system and may, therefore, be expressed in practical units. In particular it is found that the parameter μ_0 must have the value $4\pi \times 10^{-7}$, whence from (1) the value of ϵ_0 may be calculated. Inversely one might equally well *assume* this value of μ_0 as a fourth basic unit and then deduce the practical series from the field equations.

At a plenary session in June, 1935, the International Electrotechnical Commission adopted unanimously the m.k.s. system of Giorgi. Certain questions, however, still remain to be settled. No official agreement has as yet been reached as to the fourth fundamental unit. Giorgi himself recommended that the ohm, a material standard defined as the resistance of a specified column of mercury under specified conditions of pressure and temperature, be introduced as a basic quantity. If $\mu_0 = 4\pi \times 10^{-7}$ be chosen as the fourth unit and assumed dimensionless, all derived quantities may be expressed in terms of mass, length, and time alone, the dimensions of each being identical with those of the corresponding quantity in the absolute electromagnetic system and differing from them only in the size of the units. This assumption leads, however, to fractional exponents in the dimensions of many quantities, a direct consequence of our arbitrariness in clinging to mass, length, and time as the sole fundamental entities. In the absolute electromagnetic system, for example, the dimensions of charge are $\text{grams}^{\frac{1}{2}} \cdot \text{centimeters}^{\frac{1}{2}}$, an irrationality which can hardly be physically significant. These fractional exponents are entirely eliminated if we choose as a fourth unit the coulomb; for this reason, charge has been advocated at various times as a fundamental quantity quite apart from the question of its magnitude.[1] In the present volume we shall adhere exclusively to the meter-kilogram-second-coulomb system. A subsequent choice by the I.E.C. of some other electrical quantity as basic will in nowise affect the size of our units or the form of the equations.[2]

[1] See the discussion by WALLOT: *Elektrotechnische Zeitschrift*, Nos. 44–46, 1922. Also SOMMERFELD: "Ueber die Electromagnetischen Einheiten," pp. 157–165, Zeeman Verhandelingen, Martinus Nijhoff, The Hague, 1935; *Physik. Z.* **36**, 814–820, 1935.

[2] No ruling has been made as yet on the question of rationalization and opinion seems equally divided in favor and against. If one bases the theory on Maxwell's equations, it seems definitely advantageous to drop the factors 4π which in unrationalized systems stand before the charge and current densities. A rationalized system will be employed in this book.

To demonstrate that the proposed units do constitute a self-consistent system let us proceed as follows. The unit of current in the m.k.s. system is to be the absolute ampere and the unit of resistance is to be the absolute ohm. These quantities are to be such that the work expended per second by a current of 1 amp. passing through a resistance of 1 ohm is 1 joule (absolute). If R is the resistance of a section of conductor carrying a constant current of I amp., the work dissipated in heat in t sec. is

$$(4) \qquad W = I^2Rt \qquad \text{joules.}$$

By means of a calorimeter the heat generated may be measured and thus one determines the relation of the unit of electrical energy to the unit quantity of heat. It is desired that the joule defined by (4) be identical with the joule defined as a unit of mechanical work, so that in the electrical as well as in the mechanical case

$$(5) \qquad 1 \text{ joule} = 0.2389 \quad \text{gram-calorie (mean).}$$

Now we shall *define* the ampere on the basis of the equation of continuity (6), page 4, as the current which transports across any surface 1 coulomb in 1 sec. Then the ohm is a *derived* unit whose magnitude and dimensions are determined by (4):

$$(6) \qquad 1 \text{ ohm} = 1\,\frac{\text{watt}}{\text{ampere}^2} = 1\,\frac{\text{kilogram}\cdot\text{meter}^2}{\text{coulomb}^2\cdot\text{second}},$$

since 1 watt is equal to 1 joule/sec. The resistivity of a medium is defined as the resistance measured between two parallel faces of a unit cube. The reciprocal of this quantity is the conductivity. The dimensions of σ follow from Eq. (17), page 15.

$$(7) \qquad 1 \text{ unit of conductivity} = \frac{1}{\text{ohm}\cdot\text{meter}} = 1\,\frac{\text{coulomb}^2\cdot\text{second}}{\text{kilogram}\cdot\text{meter}^3}.$$

In the United States the reciprocal ohm is usually called the mho, although the name *siemens* has been adopted officially by the I.E.C. The unit of conductivity is therefore 1 siemens/meter.

The volt will be defined simply as 1 watt/amp., or

$$(8) \qquad 1 \text{ volt} = 1\,\frac{\text{watt}}{\text{ampere}} = 1\,\frac{\text{kilogram}\cdot\text{meter}^2}{\text{coulomb}\cdot\text{second}^2}.$$

Since the unit of current density is 1 amp./meter², we deduce from the relation $\mathbf{J} = \sigma\mathbf{E}$ that

$$(9) \qquad 1 \text{ unit of } \mathbf{E} = 1\,\frac{\text{watt}}{\text{ampere}\cdot\text{meter}} = 1\,\frac{\text{volt}}{\text{meter}} = 1\,\frac{\text{kilogram}\cdot\text{meter}}{\text{coulomb}\cdot\text{second}^2}.$$

The power expended per unit volume by a current of density **J** is therefore $\mathbf{E}\cdot\mathbf{J}$ watts/meter³. It will be noted furthermore that the product of charge and electric field intensity **E** has the dimensions of force. Let a charge of 1 coulomb be placed in an electric field whose intensity is 1 volt/meter.

$$(10) \qquad 1 \text{ coulomb} \times 1\,\frac{\text{volt}}{\text{meter}} = 1\,\frac{\text{joule}}{\text{meter}} = 1\,\frac{\text{kilogram}\cdot\text{meter}}{\text{second}^2}.$$

The unit of force in the m.k.s. system is called the *newton*, and is equivalent to 1 joule/meter, or 10^5 dynes.

The flux of the vector **B** shall be measured in *webers*,

$$(11) \qquad \Phi = \int_S \mathbf{B}\cdot\mathbf{n}\,da \qquad \text{webers},$$

and the intensity of the field **B**, or flux density, may therefore be expressed in webers per square meter. According to (25), page 8,

$$(12) \qquad \int_C \mathbf{E}\cdot d\mathbf{s} = -\frac{d\Phi}{dt} \qquad \frac{\text{webers}}{\text{second}}.$$

The line integral $\int_a^b \mathbf{E}\cdot d\mathbf{s}$ is measured in volts and is usually called the electromotive force (abbreviated e.m.f.) between the points a and b, although its value in a nonstationary field depends on the path of integration. The induced e.m.f. around any closed contour C is, therefore, equal to the rate of decrease of flux threading that contour, so that between the units there exists the relation

$$(13) \qquad 1 \text{ volt} = 1\,\frac{\text{weber}}{\text{second}},$$

or

$$(14) \qquad 1 \text{ weber} = 1\,\frac{\text{joule}}{\text{ampere}} = 1\,\frac{\text{kilogram}\cdot\text{meter}^2}{\text{coulomb}\cdot\text{second}}.$$

It is important to note that the product of current and magnetic flux is an energy. Note also that the product of **B** and a velocity is measured in volts per meter, and is therefore a quantity of the same kind as **E**.

$$(15) \qquad 1 \text{ unit of } \mathbf{B} = 1\,\frac{\text{weber}}{\text{meter}^2} = 1\,\frac{\text{kilogram}}{\text{coulomb}\cdot\text{second}}.$$

$$(16) \qquad 1 \text{ unit of } |\mathbf{B}|\,|\mathbf{v}| = 1\,\frac{\text{weber}}{\text{meter}^2} \times 1\,\frac{\text{meter}}{\text{second}} = 1\,\frac{\text{volt}}{\text{meter}} = 1 \text{ unit of } |\mathbf{E}|.$$

The units which have been deduced thus far constitute an absolute system in the sense that each has been expressed in terms of the four basic quantities, mass, length, time, and charge. That this system is identical with the practical series may be verified by the substitutions

$$(17) \qquad 1 \text{ kilogram} = 10^3 \text{ grams}, \qquad 1 \text{ meter} = 10^2 \text{ centimeters}, \qquad 1 \text{ coulomb} = \tfrac{1}{10} \text{ abcoulomb}.$$

The numerical factors which now appear in each relation are observed to be those that relate the practical units to the absolute electromagnetic units. For example, from (6),

$$(18) \qquad 1 \text{ ohm} = 1\,\frac{\text{kilogram}\cdot\text{meter}^2}{\text{coulomb}^2\cdot\text{second}} = \frac{10^3 \text{ grams}\cdot 10^4 \text{ centimeters}^2}{10^{-2} \text{ abcoulomb}^2\cdot\text{seconds}} = 10^9 \quad \text{abohms};$$

and again from (8),

$$(19) \qquad 1 \text{ volt} = 1\,\frac{\text{kilogram}\cdot\text{meter}^2}{\text{coulomb}\cdot\text{second}^2} = \frac{10^3 \text{ grams}\cdot 10^4 \text{ centimeters}^2}{10^{-1} \text{ abcoulomb}\cdot\text{second}^2} = 10^8 \quad \text{abvolts}.$$

The series must be completed by a determination of the units and dimensions of the vectors **D** and **H**. Since $\mathbf{D} = \epsilon\mathbf{E}$, $\mathbf{H} = \frac{1}{\mu}\mathbf{B}$, it is necessary and sufficient that ϵ_0 and μ_0 be determined such as to satisfy Eq. (2) and such that the proper ratio of practical to absolute units be maintained. We shall represent mass, length, time, and charge by the letters M, L, T, and Q, respectively, and employ the customary symbol $[A]$ as meaning "the dimensions of A." Then from Eq. (31), page 9,

$$(20) \qquad \int_S \mathbf{D}\cdot\mathbf{n}\,da = q \qquad \text{coulombs}$$

and, hence,

$$(21) \qquad [\mathbf{D}] = \frac{\text{coulombs}}{\text{meter}^2} = \frac{Q}{L^2},$$

$$(22) \qquad [\epsilon_0] = \left[\frac{D}{\kappa_e E}\right] = \frac{\text{coulombs}}{\text{volt}\cdot\text{meter}} = \frac{Q^2T^2}{ML^3}.$$

The *farad*, a derived unit of capacity, is defined as the capacity of a conducting body whose potential will be raised 1 volt by a charge of 1 coulomb. It is equal, in other words, to 1 coulomb/volt. The parameter ϵ_0 in the m.k.s. system has dimensions, and may be measured in *farads per meter*.

By analogy with the electrical case, the line integral $\int_a^b \mathbf{H}\cdot d\mathbf{s}$ taken along a specified path is commonly called the magnetomotive force (abbreviated m.m.f.). In a stationary magnetic field

$$\int_C \mathbf{H} \cdot d\mathbf{s} = I \qquad \text{amperes,} \tag{23}$$

where I is the current determined by the flow of charge through any surface spanning the closed contour C. If the field is variable, I must include the displacement current as in (28), page 9. According to (23) a magnetomotive force has the dimensions of current. In practice, however, the current is frequently carried by the turns of a coil or winding which is linked by the contour C. If there are n such turns carrying a current I, the total current threading C is nI ampere-turns and it is customary to express magnetomotive force in these terms, although dimensionally n is a numeric.

$$[\text{m.m.f.}] = \text{ampere-turns}, \tag{24}$$

whence

$$[\mathbf{H}] = \frac{\text{ampere-turns}}{\text{meter}} = \frac{Q}{LT}. \tag{25}$$

It will be observed that the dimensions of $\mathbf{D}$ and those of $\mathbf{H}$ divided by a velocity are identical. For the parameter μ_0 we find

$$[\mu_0] = \left[\frac{B}{\kappa_m H}\right] = \frac{\text{volt} \cdot \text{second}}{\text{ampere} \cdot \text{meter}} = \frac{ML}{Q^2}. \tag{26}$$

As in the case of ϵ_0 it is convenient to express μ_0 in terms of a derived unit, in this case the *henry*, defined as 1 volt-second/amp. (The henry is that inductance in which an induced e.m.f. of 1 volt is generated when the inducing current is varying at the rate of 1 amp./sec.) The parameter μ_0 may, therefore, be measured in *henrys per meter*.

From (22) and (26) it follows now that

$$\left[\frac{1}{\mu_0 \epsilon_0}\right] = \frac{L^2}{T^2}, \tag{27}$$

and hence that our system is indeed dimensionally consistent with Eq. (2). Since it is known that in the rationalized, absolute c.g.s. electromagnetic system μ_0 is equal in magnitude to 4π, Eq. (26) fixes also its magnitude in the m.k.s. system.

$$\mu_0 = 4\pi \frac{\text{gram} \cdot \text{centimeters}}{\text{abcoulombs}^2} = 4\pi \frac{10^{-3}\ \text{kilogram} \cdot 10^{-2}\ \text{meter}}{10^2\ \text{coulombs}^2}, \tag{28}$$

or

$$\mu_0 = 4\pi \times 10^{-7} \frac{\text{kilogram} \cdot \text{meters}}{\text{coulombs}^2} = 1.257 \times 10^{-6} \frac{\text{henry}}{\text{meter}}. \tag{29}$$

The appropriate value of ϵ_0 is then determined from

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} = 2.998 \times 10^8 \frac{\text{meters}}{\text{second}} \tag{2}$$

to be

$$\epsilon_0 = 8.854 \times 10^{-12} \frac{\text{coulomb}^2 \cdot \text{seconds}^2}{\text{kilogram} \cdot \text{meter}^3} = 8.854 \times 10^{-12} \frac{\text{farad}}{\text{meter}}. \tag{30}$$

It is frequently convenient to know the reciprocal values of these factors.

$$\frac{1}{\mu_0} = 0.7958 \times 10^6 \frac{\text{meters}}{\text{henry}}, \qquad \frac{1}{\epsilon_0} = 0.1129 \times 10^{12} \frac{\text{meters}}{\text{farad}}, \tag{31}$$

and the quantities

$$\sqrt{\frac{\mu_0}{\epsilon_0}} = 376.6 \text{ ohms}, \qquad \sqrt{\frac{\epsilon_0}{\mu_0}} = 2.655 \times 10^{-3} \text{ mho}, \tag{32}$$

recur constantly throughout the investigation of wave propagation.

In Appendix I there will be found a summary of the units and dimensions of electromagnetic quantities in terms of mass, length, time, and charge.

THE ELECTROMAGNETIC POTENTIALS

1.9. Vector and Scalar Potentials.—The analysis of an electromagnetic field is often facilitated by the use of auxiliary functions known as potentials. At every ordinary point of space, the field vectors satisfy the system

$$\text{(I)}\quad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \qquad \text{(III)}\quad \nabla \cdot \mathbf{B} = 0,$$

$$\text{(II)}\quad \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J}, \qquad \text{(IV)}\quad \nabla \cdot \mathbf{D} = \rho.$$

According to (III) the field of the vector $\mathbf{B}$ is always solenoidal. Consequently $\mathbf{B}$ can be represented as the curl of another vector $\mathbf{A}_0$.

$$\mathbf{B} = \nabla \times \mathbf{A}_0. \tag{1}$$

However $\mathbf{A}_0$ is not uniquely defined by (1); for $\mathbf{B}$ is equal also to the curl of some vector $\mathbf{A}$,

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{2}$$

where

$$\mathbf{A} = \mathbf{A}_0 - \nabla\psi, \tag{3}$$

and ψ is any arbitrary scalar function of position.

If now $\mathbf{B}$ is replaced in (I) by either (1) or (2), we obtain, respectively,

$$(4) \qquad \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}_0}{\partial t}\right) = 0, \qquad \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0.$$

Thus the fields of the vectors $\mathbf{E} + \frac{\partial \mathbf{A}_0}{\partial t}$ and $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}$ are irrotational and equal to the gradients of two scalar functions ϕ_0 and ϕ.

$$(5) \qquad \mathbf{E} = -\nabla\phi_0 - \frac{\partial \mathbf{A}_0}{\partial t},$$

$$(6) \qquad \mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}.$$

The functions ϕ and ϕ_0 are obviously related by

$$(7) \qquad \phi = \phi_0 + \frac{\partial \psi}{\partial t}.$$

The functions $\mathbf{A}$ are *vector potentials* of the field, and the ϕ are *scalar potentials*. $\mathbf{A}_0$ and ϕ_0 designate one specific pair of potentials from which the field can be derived through (1) and (5). An infinite number of potentials leading to the same field can then be constructed from (3) and (7).

Let us suppose that the medium is homogeneous and isotropic, and that ϵ and μ are independent of field intensity.

$$(8) \qquad \mathbf{D} = \epsilon\mathbf{E}, \qquad \mathbf{B} = \mu\mathbf{H}.$$

In terms of the potentials

$$(9) \qquad \mathbf{D} = -\epsilon\left(\nabla\phi + \frac{\partial \mathbf{A}}{\partial t}\right), \qquad \mathbf{H} = \frac{1}{\mu}\nabla \times \mathbf{A},$$

which upon substitution into (II) and (IV) give

$$(10) \qquad \nabla \times \nabla \times \mathbf{A} + \mu\epsilon\nabla\frac{\partial \phi}{\partial t} + \mu\epsilon\frac{\partial^2 \mathbf{A}}{\partial t^2} = \mu\mathbf{J},$$

$$(11) \qquad \nabla^2\phi + \nabla \cdot \frac{\partial \mathbf{A}}{\partial t} = -\frac{1}{\epsilon}\rho.$$

All particular solutions of (10) and (11) lead to the same electromagnetic field when subjected to identical boundary conditions. They differ among themselves by the arbitrary function ψ. Let us impose now upon $\mathbf{A}$ and ϕ the supplementary condition

$$(12) \qquad \nabla \cdot \mathbf{A} + \mu\epsilon\frac{\partial \phi}{\partial t} = 0.$$

To do this it is only necessary that ψ shall satisfy

$$(13) \qquad \nabla^2\psi - \mu\epsilon\frac{\partial^2 \psi}{\partial t^2} = \nabla \cdot \mathbf{A}_0 + \mu\epsilon\frac{\partial \phi_0}{\partial t},$$

where ϕ_0 and $\mathbf{A}_0$ are particular solutions of (10) and (11). The potentials ϕ and $\mathbf{A}$ are now uniquely defined and are solutions of the equations

$$(14) \qquad \nabla \times \nabla \times \mathbf{A} - \nabla\nabla \cdot \mathbf{A} + \mu\epsilon\frac{\partial^2 \mathbf{A}}{\partial t^2} = \mu\mathbf{J},$$

$$(15) \qquad \nabla^2\phi - \mu\epsilon\frac{\partial^2 \phi}{\partial t^2} = -\frac{1}{\epsilon}\rho.$$

Equation (14) reduces to the same form as (15) when use is made of the vector identity

$$(16) \qquad \nabla \times \nabla \times \mathbf{A} = \nabla\nabla \cdot \mathbf{A} - \nabla \cdot \nabla\mathbf{A}.$$

The last term of (16) can be interpreted as the Laplacian operating on the *rectangular* components of $\mathbf{A}$. In this case

$$(17) \qquad \nabla^2\mathbf{A} - \mu\epsilon\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu\mathbf{J}.$$

The expansion of the operator $\nabla \cdot \nabla\mathbf{A}$ in curvilinear systems will be discussed in Sec. 1.16, page 50.

The relations (2) and (6) for the vectors $\mathbf{B}$ and $\mathbf{E}$ are by no means general. To them may be added any particular solution of the *homogeneous* equations

$$(Ia) \quad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \qquad (IIIa) \quad \nabla \cdot \mathbf{B} = 0,$$

$$(IIa) \quad \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = 0, \qquad (IVa) \quad \nabla \cdot \mathbf{D} = 0.$$

From the symmetry of this system it is at once evident that it can be satisfied identically by

$$(18) \qquad \mathbf{D} = -\nabla \times \mathbf{A}^*, \qquad \mathbf{H} = -\nabla\phi^* - \frac{\partial \mathbf{A}^*}{\partial t},$$

from which we construct

$$(19) \qquad \mathbf{E} = -\frac{1}{\epsilon}\nabla \times \mathbf{A}^*, \qquad \mathbf{B} = -\mu\left(\nabla\phi^* + \frac{\partial \mathbf{A}^*}{\partial t}\right).$$

The new potentials are subject only to the conditions

$$(20) \qquad \begin{aligned} \nabla^2\mathbf{A}^* - \mu\epsilon\frac{\partial^2 \mathbf{A}^*}{\partial t^2} &= 0, \\ \nabla^2\phi^* - \mu\epsilon\frac{\partial^2 \phi^*}{\partial t^2} &= 0, \\ \nabla \cdot \mathbf{A}^* + \mu\epsilon\frac{\partial \phi^*}{\partial t} &= 0. \end{aligned}$$

A general solution of the inhomogeneous system (I) to (IV) is, therefore,

$$\mathbf{B} = \nabla \times \mathbf{A} - \mu \frac{\partial \mathbf{A}^*}{\partial t} - \mu \nabla \phi^*, \tag{21}$$

$$\mathbf{E} = -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} - \frac{1}{\epsilon} \nabla \times \mathbf{A}^*, \tag{22}$$

provided μ and ϵ are constant.

The functions ϕ^* and $\mathbf{A}^*$ are potentials of a source distribution which is entirely external to the region considered. Usually ϕ^* and $\mathbf{A}^*$ are put equal to zero and the potentials of all charges, both distant and local, are represented by ϕ and $\mathbf{A}$.

At any point where the charge and current densities are zero a *possible* field is $\phi_0 = 0$, $\mathbf{A}_0 = 0$. The function ψ is now any solution of the homogeneous equation

$$\nabla^2 \psi - \mu\epsilon \frac{\partial^2 \psi}{\partial t^2} = 0. \tag{23}$$

Since at the same point the scalar potential ϕ satisfies the same equation, ψ may be chosen such that ϕ vanishes. *In this case the field can be expressed in terms of a vector potential alone.*

$$\mathbf{B} = \nabla \times \mathbf{A}, \qquad \mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}, \tag{24}$$

$$\nabla^2 \mathbf{A} - \mu\epsilon \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0, \qquad \nabla \cdot \mathbf{A} = 0. \tag{25}$$

Concerning the units and dimensions of these new quantities, we note first that $\mathbf{E}$ is measured in volts/meter and that the scalar potential ϕ is therefore to be measured in volts. If q is a charge measured in coulombs, it follows that the product $q\phi$ represents an energy expressed in joules. From the relation $\mathbf{B} = \nabla \times \mathbf{A}$ it is clear that the vector potential $\mathbf{A}$ may be expressed in webers/meter, but equally well in either volt-seconds/meter or in joules/ampere. The product of current and vector potential is therefore an energy. The dimensions of $\mathbf{A}^*$ are found to be coulombs/meter, while ϕ^* will be measured in ampere-turns.

1.10. Homogeneous Conducting Media.—In view of the extreme brevity of the relaxation time it may be assumed that the density of free charge is always zero in the interior of a conductor. The field equations for a homogeneous, isotropic medium then reduce to

$$\text{(I}b\text{)}\quad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \qquad \text{(III}b\text{)}\quad \nabla \cdot \mathbf{B} = 0,$$

$$\text{(II}b\text{)}\quad \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} - \sigma \mathbf{E} = 0, \qquad \text{(IV}b\text{)}\quad \nabla \cdot \mathbf{D} = 0.$$

We are now free to express either $\mathbf{B}$ or $\mathbf{D}$ in terms of a vector potential. In the first alternative we have

$$\mathbf{B} = \nabla \times \mathbf{A}, \qquad \mathbf{E} = -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t}. \tag{26}$$

If the vector and scalar potentials are subjected to the relation

$$\nabla \cdot \mathbf{A} + \mu\epsilon \frac{\partial \phi}{\partial t} + \mu\sigma\phi = 0, \tag{27}$$

a possible electromagnetic field may be constructed from any pair of solutions of the equations

$$\nabla^2 \mathbf{A} - \mu\epsilon \frac{\partial^2 \mathbf{A}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{A}}{\partial t} = 0, \tag{28}$$

$$\nabla^2 \phi - \mu\epsilon \frac{\partial^2 \phi}{\partial t^2} - \mu\sigma \frac{\partial \phi}{\partial t} = 0. \tag{29}$$

As in the preceding paragraph one will note that the field vectors are invariant to changes in the potentials satisfying the relations

$$\phi = \phi_0 + \frac{\partial \psi}{\partial t}, \qquad \mathbf{A} = \mathbf{A}_0 - \nabla \psi, \tag{30}$$

where ϕ_0, $\mathbf{A}_0$ are the potentials of a possible field and ψ is an arbitrary scalar function. In order that $\mathbf{A}$ and ϕ satisfy (27) it is only necessary that ψ be subjected to the additional condition

$$\nabla^2 \psi - \mu\epsilon \frac{\partial^2 \psi}{\partial t^2} - \mu\sigma \frac{\partial \psi}{\partial t} = \nabla \cdot \mathbf{A}_0 + \mu\epsilon \frac{\partial \phi_0}{\partial t} + \mu\sigma\phi_0. \tag{31}$$

To a particular solution of (31) one is free to add any solution of the homogeneous equation

$$\nabla^2 \psi - \mu\epsilon \frac{\partial^2 \psi}{\partial t^2} - \mu\sigma \frac{\partial \psi}{\partial t} = 0. \tag{32}$$

Frequently it is convenient to choose ψ such that the scalar potential vanishes. The field within the conductor is then determined by a single vector $\mathbf{A}$.

$$\mathbf{B} = \nabla \times \mathbf{A}, \qquad \mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}, \tag{33}$$

$$\nabla^2 \mathbf{A} - \mu\epsilon \frac{\partial^2 \mathbf{A}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{A}}{\partial t} = 0, \qquad \nabla \cdot \mathbf{A} = 0. \tag{34}$$

The field may also be defined in terms of potentials ϕ^* and $\mathbf{A}^*$ by

$$\mathbf{D} = -\nabla \times \mathbf{A}^*, \qquad \mathbf{H} = -\nabla \phi^* - \frac{\partial \mathbf{A}^*}{\partial t} - \frac{\sigma}{\epsilon} \mathbf{A}^*. \tag{35}$$

If ϕ^* and $\mathbf{A}^*$ are to satisfy (28) and (29), it is necessary that they be related by

$$\nabla \cdot \mathbf{A}^* + \mu\epsilon \frac{\partial \phi^*}{\partial t} = 0. \tag{36}$$

The field defined by (35) is invariant to all transformations of the potentials of the type

$$\phi^* = \phi_0^* + \frac{\partial \psi^*}{\partial t} + \frac{\sigma}{\epsilon}\psi^*, \qquad \mathbf{A}^* = \mathbf{A}_0^* - \nabla\psi^*, \tag{37}$$

where as above ϕ_0^* and $\mathbf{A}_0^*$ are the potentials of any possible electromagnetic field. To ensure the relation (36) it is only necessary that ψ^* be chosen such as to satisfy

$$\nabla^2\psi^* - \mu\epsilon\frac{\partial^2\psi^*}{\partial t^2} - \mu\sigma\frac{\partial\psi^*}{\partial t} = \nabla \cdot \mathbf{A}_0^* + \mu\epsilon\frac{\partial\phi_0^*}{\partial t}. \tag{38}$$

Finally, by a proper choice of ψ^* the scalar potential ϕ^* may be made to vanish.

$$\mathbf{D} = -\nabla \times \mathbf{A}^*, \qquad \mathbf{H} = -\frac{\partial \mathbf{A}^*}{\partial t} - \frac{\sigma}{\epsilon}\mathbf{A}^*, \tag{39}$$

$$\nabla^2\mathbf{A}^* - \mu\epsilon\frac{\partial^2\mathbf{A}^*}{\partial t^2} - \mu\sigma\frac{\partial\mathbf{A}^*}{\partial t} = 0, \qquad \nabla \cdot \mathbf{A}^* = 0. \tag{40}$$

1.11. The Hertz Vectors, or Polarization Potentials.—We have seen that the integration of Maxwell's equations may be reduced to the determination of a vector and a scalar potential, which in homogeneous media satisfy one and the same differential equation. It was shown by Hertz[1] that it is possible under ordinary conditions to define an electromagnetic field in terms of a single vector function.

Let us confine ourselves for the present to regions of an isotropic, homogeneous medium within which there are neither conduction currents nor free charges. The field equations then reduce to the homogeneous system (I*a*)–(IV*a*). We assume, for reasons which will become apparent, that the vector potential $\mathbf{A}$ is proportional to the time derivative of a vector $\mathbf{\Pi}$.

$$\mathbf{A} = \mu\epsilon\frac{\partial \mathbf{\Pi}}{\partial t}. \tag{41}$$

Consequently,

$$\mathbf{B} = \mu\epsilon\nabla \times \frac{\partial \mathbf{\Pi}}{\partial t}, \qquad \mathbf{E} = -\nabla\phi - \mu\epsilon\frac{\partial^2 \mathbf{\Pi}}{\partial t^2}, \tag{42}$$

[1] HERTZ, *Ann. Physik*, **36**, 1, 1888. The general solution is due to Righi: *Bologna Mem.*, (5) **9**, 1, 1901, and *Il Nuovo Cimento*, (5) **2**, 2, 1901.

and, when in turn this expression for $\mathbf{E}$ is introduced into (II*a*), it is found that

$$\frac{\partial}{\partial t}\left(\nabla \times \nabla \times \mathbf{\Pi} + \nabla\phi + \mu\epsilon\frac{\partial^2\mathbf{\Pi}}{\partial t^2}\right) = 0. \tag{43}$$

We recall that at points where there is no charge, the scalar function ϕ is wholly arbitrary so long as it satisfies an equation such as (23). In the present instance it will be chosen such that

$$\phi = -\nabla \cdot \mathbf{\Pi}. \tag{44}$$

Then upon integrating (43) with respect to the time, we obtain

$$\nabla \times \nabla \times \mathbf{\Pi} - \nabla\nabla \cdot \mathbf{\Pi} + \mu\epsilon\frac{\partial^2\mathbf{\Pi}}{\partial t^2} = \text{constant}. \tag{45}$$

The particular value of the constant does not affect the determination of the field and we are therefore free to place it equal to zero. Equation (IV*a*) is also satisfied, for the divergence of the curl of any vector vanishes identically. Then we may state that *every solution of the vector equation*

$$\nabla \times \nabla \times \mathbf{\Pi} - \nabla\nabla \cdot \mathbf{\Pi} + \mu\epsilon\frac{\partial^2\mathbf{\Pi}}{\partial t^2} = 0 \tag{46}$$

determines an electromagnetic field through

$$\mathbf{B} = \mu\epsilon\nabla \times \frac{\partial\mathbf{\Pi}}{\partial t}, \qquad \mathbf{E} = \nabla\nabla \cdot \mathbf{\Pi} - \mu\epsilon\frac{\partial^2\mathbf{\Pi}}{\partial t^2}. \tag{47}$$

The condition that ϕ shall satisfy (23) is fulfilled in virtue of (46). One may replace (46) by

$$\nabla^2\mathbf{\Pi} - \mu\epsilon\frac{\partial^2\mathbf{\Pi}}{\partial t^2} = 0, \tag{48}$$

provided ∇^2 is understood to operate on the *rectangular* components of $\mathbf{\Pi}$.

Since the vector $\mathbf{D}$ as well as $\mathbf{B}$ is solenoidal in a charge-free region, an alternative solution can be constructed of the form

$$\mathbf{A}^* = \mu\epsilon\frac{\partial\mathbf{\Pi}^*}{\partial t}, \qquad \phi^* = -\nabla \cdot \mathbf{\Pi}^*, \tag{49}$$

$$\mathbf{D} = -\mu\epsilon\nabla \times \frac{\partial\mathbf{\Pi}^*}{\partial t}, \qquad \mathbf{H} = \nabla\nabla \cdot \mathbf{\Pi}^* - \mu\epsilon\frac{\partial^2\mathbf{\Pi}^*}{\partial t^2}, \tag{50}$$

where $\mathbf{\Pi}^*$ is any solution of (46) or (48).

From these results we conclude that the electromagnetic field within a region throughout which ϵ and μ are constant, ρ and $\mathbf{J}$ equal to zero, may be resolved into two partial fields, the one derived from the vector $\mathbf{\Pi}$

and the other from the vector $\mathbf{\Pi}^*$. The origin of these fields lies exterior to the region. To determine the physical significance of the Hertz vectors it is now necessary to relate them to their sources; in other words, we must find the *inhomogeneous* equations from which (48) is derived.

Let us express the vector **D** in terms of **E** and the electric polarization **P**. According to (6), page 11, $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$. Then in place of (II*a*) and (IV*a*), we must now write

$$(51)\qquad \nabla \times \mathbf{H} - \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \frac{\partial \mathbf{P}}{\partial t}, \qquad \nabla \cdot \mathbf{E} = -\frac{1}{\epsilon_0} \nabla \cdot \mathbf{P}.$$

It may be verified without difficulty that these two equations, as well as (I*a*) and (III*a*), are still identically satisfied by (47), provided only that ϵ be replaced by ϵ_0 and that $\mathbf{\Pi}$ be now any solution of

$$(52)\qquad \nabla^2 \mathbf{\Pi} - \mu\epsilon_0 \frac{\partial^2 \mathbf{\Pi}}{\partial t^2} = -\frac{1}{\epsilon_0} \mathbf{P}.$$

The source of the vector $\mathbf{\Pi}$ *and the electromagnetic field derived from it is a distribution of electric polarization* **P**. In due course we shall interpret the vector **P** as the electric dipole moment per unit volume of the medium. Since $\mathbf{\Pi}$ is associated with a distribution of electric dipoles, the partial field which it defines is sometimes said to be of *electric type*, and $\mathbf{\Pi}$ itself may be called the electric polarization potential.

In like manner it can be shown that the field associated with $\mathbf{\Pi}^*$ is set up by a distribution of magnetic polarization. According to (6), page 11, the vector **B** is related to **H** by $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$, which when introduced into (I*a*) and (III*a*) gives

$$(53)\qquad \nabla \times \mathbf{E} + \mu_0 \frac{\partial \mathbf{H}}{\partial t} = -\mu_0 \frac{\partial \mathbf{M}}{\partial t}, \qquad \nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M}.$$

Then these equations, as well as (II*a*) and (IV*a*), are satisfied identically by (50) if we replace there μ by μ_0 and prescribe that $\mathbf{\Pi}^*$ shall be a solution of

$$(54)\qquad \nabla^2 \mathbf{\Pi}^* - \mu_0\epsilon \frac{\partial^2 \mathbf{\Pi}^*}{\partial t^2} = -\mathbf{M}.$$

We shall show later that the polarization **M** may be interpreted as the density of a distribution of magnetic moment. The partial field derived from $\mathbf{\Pi}^*$ may be imagined to have its origin in magnetic dipoles and is said to be a field of *magnetic type.*

The electric polarization **P** may be induced in the dielectric by the field **E**, but it may also contain a part whose magnitude is controlled by wholly external factors. In the practical application of the theory one is interested usually only in this independent part $\mathbf{P}_0$, which will be shown to represent the electric moment of dipole oscillators activated by external power sources. The same is true for the magnetic polarization. To represent these conditions we shall write (6), page 11, in the modified form

$$(55)\qquad \mathbf{D} = \epsilon\mathbf{E} + \mathbf{P}_0, \qquad \mathbf{H} = \frac{1}{\mu}\mathbf{B} - \mathbf{M}_0,$$

in which $\mathbf{P}_0$ and $\mathbf{M}_0$ are prescribed and independent of **E** and **H**, *and where the induced polarizations of the medium have again been absorbed into the parameters* ϵ *and* μ. Then the electromagnetic field due to these distributions of $\mathbf{P}_0$ and $\mathbf{M}_0$ is determined by

$$(56)\qquad \mathbf{E} = \nabla\nabla \cdot \mathbf{\Pi} - \mu\epsilon \frac{\partial^2 \mathbf{\Pi}}{\partial t^2} - \mu\nabla \times \frac{\partial \mathbf{\Pi}^*}{\partial t},$$

$$(57)\qquad \mathbf{H} = \epsilon\nabla \times \frac{\partial \mathbf{\Pi}}{\partial t} + \nabla\nabla \cdot \mathbf{\Pi}^* - \mu\epsilon \frac{\partial^2 \mathbf{\Pi}^*}{\partial t^2},$$

when $\mathbf{\Pi}$ and $\mathbf{\Pi}^*$ are solutions of

$$(58)\qquad \nabla^2 \mathbf{\Pi} - \mu\epsilon \frac{\partial^2 \mathbf{\Pi}}{\partial t^2} = -\frac{1}{\epsilon} \mathbf{P}_0, \qquad \nabla^2 \mathbf{\Pi}^* - \mu\epsilon \frac{\partial^2 \mathbf{\Pi}^*}{\partial t^2} = -\mathbf{M}_0.$$

In virtue of the second of Eqs. (58) and of the identity (16) we may also write (57) as

$$(59)\qquad \mathbf{H} = \epsilon\nabla \times \frac{\partial \mathbf{\Pi}}{\partial t} + \nabla \times \nabla \times \mathbf{\Pi}^* - \mathbf{M}_0.$$

Since $\mathbf{B} = \nabla \times \mathbf{A}$, it is evident from this last relation that the vector potential **A** may be derived from the Hertzian vectors by putting

$$(60)\qquad \mathbf{A} = \mu\epsilon \frac{\partial \mathbf{\Pi}}{\partial t} + \mu\nabla \times \mathbf{\Pi}^* - \nabla\psi,$$

where ψ is an arbitrary scalar function. The associated potential ϕ is

$$(61)\qquad \phi = -\nabla \cdot \mathbf{\Pi} + \frac{\partial \psi}{\partial t},$$

with ψ subject only to the condition that it satisfy

$$(62)\qquad \nabla^2\psi - \mu\epsilon \frac{\partial^2 \psi}{\partial t^2} = 0.$$

The extension of these equations to a homogeneous conducting medium follows without difficulty. The reader will verify by direct substitution that the system (I*b*)–(IV*b*), in a medium which is *free of* *fixed polarization* $\mathbf{P}_0$ *and* $\mathbf{M}_0$, is satisfied by

$$\mathbf{E} = \nabla \times \nabla \times \mathbf{\Pi} - \mu \nabla \times \frac{\partial \mathbf{\Pi}^*}{\partial t}, \tag{63}$$

$$\mathbf{H} = \nabla \times \left(\epsilon \frac{\partial \mathbf{\Pi}}{\partial t} + \sigma \mathbf{\Pi}\right) + \nabla \times \nabla \times \mathbf{\Pi}^*, \tag{64}$$

$$\begin{gathered} \nabla \times \nabla \times \mathbf{\Pi} - \nabla\nabla \cdot \mathbf{\Pi} + \mu\epsilon \frac{\partial^2 \mathbf{\Pi}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{\Pi}}{\partial t} = 0, \\ \nabla \times \nabla \times \mathbf{\Pi}^* - \nabla\nabla \cdot \mathbf{\Pi}^* + \mu\epsilon \frac{\partial^2 \mathbf{\Pi}^*}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{\Pi}^*}{\partial t} = 0. \end{gathered} \tag{65}$$

1.12. Complex Field Vectors and Potentials.—It has been shown by Silberstein, Bateman, and others that the equations satisfied by the fields and potentials may be reduced to a particularly compact form by the construction of a complex vector whose real and imaginary parts are formed from the vectors defining the magnetic and electric fields.[1] The procedure has no apparent physical significance but frequently facilitates analysis.

Consider again a homogeneous, isotropic medium in which $\mathbf{D} = \epsilon\mathbf{E}$, $\mathbf{B} = \mu\mathbf{H}$. If now we define $\mathbf{Q}$ as a complex field vector by

$$\mathbf{Q} = \mathbf{B} + i\sqrt{\epsilon\mu}\,\mathbf{E}, \tag{66}$$

the Maxwell equations (I)–(IV) reduce to

$$\nabla \times \mathbf{Q} + i\sqrt{\epsilon\mu}\frac{\partial \mathbf{Q}}{\partial t} = \mu\mathbf{J}, \qquad \nabla \cdot \mathbf{Q} = i\sqrt{\frac{\mu}{\epsilon}}\,\rho. \tag{67}$$

The vector operation $\nabla \times \mathbf{Q}$ may be eliminated from (67) by the simple expedient of taking the curl of both members. By the identity (16) we obtain

$$\nabla\nabla \cdot \mathbf{Q} - \nabla^2\mathbf{Q} + i\sqrt{\epsilon\mu}\,\nabla \times \frac{\partial \mathbf{Q}}{\partial t} = \mu\nabla \times \mathbf{J}, \tag{68}$$

which, on replacing the curl and divergence of $\mathbf{Q}$ by their values from (67), reduces to

$$\nabla^2\mathbf{Q} - \epsilon\mu\frac{\partial^2 \mathbf{Q}}{\partial t^2} = -\mu\left(\nabla \times \mathbf{J} - i\sqrt{\epsilon\mu}\frac{\partial \mathbf{J}}{\partial t} - i\frac{1}{\sqrt{\epsilon\mu}}\nabla\rho\right). \tag{69}$$

When this last equation is resolved into its real and imaginary com-

[1] SILBERSTEIN, *Ann. phys.*, **22**, **24**, 1907. Also *Phil. Mag.* (6) **23**, 790, 1912. BATEMAN, "Electrical and Optical Wave Motion," Chap. I, Cambridge University Press.

ponents, one obtains the equations satisfied individually by the vectors $\mathbf{E}$ and $\mathbf{H}$.

$$\nabla^2\mathbf{H} - \epsilon\mu\frac{\partial^2 \mathbf{H}}{\partial t^2} = -\nabla \times \mathbf{J}, \tag{70}$$

$$\nabla^2\mathbf{E} - \epsilon\mu\frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu\frac{\partial \mathbf{J}}{\partial t} + \frac{1}{\epsilon}\nabla\rho. \tag{71}$$

Next, let us define $\mathbf{Q}$ in terms of complex vector and scalar potentials $\mathbf{L}$ and Φ by the equation

$$\mathbf{Q} = \nabla \times \mathbf{L} - i\sqrt{\epsilon\mu}\frac{\partial \mathbf{L}}{\partial t} - i\sqrt{\epsilon\mu}\,\nabla\Phi, \tag{72}$$

subject to the condition

$$\nabla \cdot \mathbf{L} + \epsilon\mu\frac{\partial \Phi}{\partial t} = 0. \tag{73}$$

It will be verified without difficulty that (72) is an integral of (67) provided the complex potentials satisfy the equations

$$\nabla^2\mathbf{L} - \epsilon\mu\frac{\partial^2 \mathbf{L}}{\partial t^2} = -\mu\mathbf{J}, \tag{74}$$

$$\nabla^2\Phi - \epsilon\mu\frac{\partial^2 \Phi}{\partial t^2} = -\frac{1}{\epsilon}\rho. \tag{75}$$

If the real and imaginary parts of these potentials are written in the form

$$\mathbf{L} = \mathbf{A} - i\sqrt{\frac{\mu}{\epsilon}}\mathbf{A}^*, \qquad \Phi = \phi - i\sqrt{\frac{\mu}{\epsilon}}\phi^*, \tag{76}$$

and substituted into (72), one finds again after separation of reals and imaginaries the general expressions for the field vectors deduced in Eqs. (21) and (22).

If the free currents and charges are everywhere zero in the region under consideration, Eq. (67) reduces to

$$\nabla \times \mathbf{Q} + i\sqrt{\epsilon\mu}\frac{\partial \mathbf{Q}}{\partial t} = 0, \qquad \nabla \cdot \mathbf{Q} = 0. \tag{77}$$

The electromagnetic field may now be expressed in terms of a single complex Hertzian vector $\mathbf{\Gamma}$.

$$\mathbf{Q} = \mu\epsilon\nabla \times \frac{\partial \mathbf{\Gamma}}{\partial t} + i\sqrt{\mu\epsilon}\,\nabla \times \nabla \times \mathbf{\Gamma}, \tag{78}$$

where $\mathbf{\Gamma}$ is any solution of

$$\nabla^2\mathbf{\Gamma} - \epsilon\mu\frac{\partial^2 \mathbf{\Gamma}}{\partial t^2} = 0. \tag{79}$$

If, finally, $\mathbf{\Gamma}$ is defined as

$$\mathbf{\Gamma} = \mathbf{\Pi} - i\sqrt{\frac{\mu}{\epsilon}}\,\mathbf{\Pi}^* \qquad (80)$$

and substituted into (78), one finds again after separation into real and imaginary parts exactly the expressions (47) and (50) for the electric and magnetic field vectors.

When the medium is conducting, the field equations are no longer symmetrical and the method fails. The difficulty may be overcome if the field varies harmonically. The time then enters explicitly as a factor such as $e^{\pm i\omega t}$. After differentiating with respect to time, the system (I*b*)–(IV*b*) may be made symmetrical by introducing a *complex inductive capacity* $\epsilon' = \epsilon \pm i\dfrac{\sigma}{\omega}$.

BOUNDARY CONDITIONS

1.13. Discontinuities in the Field Vectors.—The validity of the field equations has been postulated only for ordinary points of space; that is to say, for points in whose neighborhood the physical properties of the medium vary continuously. However, across any surface which bounds one body or medium from another there occur sharp changes in the parameters ϵ, μ, and σ. On a macroscopic scale these changes may usually be considered discontinuous and hence the field vectors themselves may be expected to exhibit corresponding discontinuities.

Let us imagine at the start that the surface S which bounds medium (1) from medium (2) has been replaced by a very thin transition layer within which the parameters ϵ, μ, σ vary rapidly but *continuously* from their values near S in (1) to their values near S in (2). Within this layer, as within the media (1) and (2), the field vectors and their first derivatives are continuous, bounded functions of position and time. Through the layer we now draw a small right cylinder, as indicated in Fig. 2*a*. The elements of the cylinder are normal to S and its ends lie in the surfaces of the layer so that they are separated by just the layer thickness Δl. Fixing our attention first on the field of the vector $\mathbf{B}$, we have

$$\oint \mathbf{B}\cdot\mathbf{n}\,da = 0, \qquad (1)$$

when integrated over the walls and ends of the cylinder. If the base, whose area is Δa, is made sufficiently small, it may be assumed that $\mathbf{B}$ has a constant value over each end. Neglecting differentials of higher order we may approximate (1) by

$$(\mathbf{B}\cdot\mathbf{n}_1 + \mathbf{B}\cdot\mathbf{n}_2)\Delta a + \text{contributions of the walls} = 0. \qquad (2)$$

The contribution of the walls to the surface integral is directly proportional to Δl. Now let the transition layer shrink into the surface S. In the limit, as $\Delta l \rightarrow 0$, the ends of the cylinder lie just on either side of S and the contribution from the walls becomes vanishingly small. The value of $\mathbf{B}$ at a point on S in medium (1) will be denoted by $\mathbf{B}_1$, while

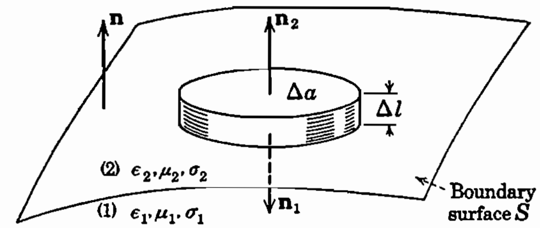

FIG. 2*a*.—For the normal boundary condition.

the corresponding value of $\mathbf{B}$ just across the surface in (2) will be denoted by $\mathbf{B}_2$. We shall also indicate the positive normal to S by a unit vector $\mathbf{n}$ drawn from (1) into (2). According to this convention medium (1) lies on the negative side of S, medium (2) on the positive side, and $\mathbf{n}_1 = -\mathbf{n}$. Then as $\Delta l \rightarrow 0$, $\Delta a \rightarrow 0$,

$$(\mathbf{B}_2 - \mathbf{B}_1)\cdot\mathbf{n} = 0; \qquad (3)$$

the transition of the normal component of $\mathbf{B}$ *across any surface of discontinuity in the medium is continuous.* Equation (3) is a direct consequence of the condition $\nabla\cdot\mathbf{B} = 0$, and is sometimes called the surface divergence.

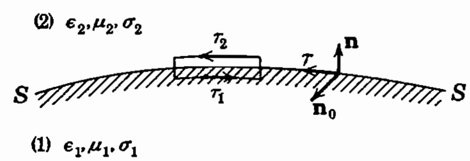

FIG. 2*b*.—For the tangential boundary condition.

The vector $\mathbf{D}$ may be treated in the same manner, but in this case the surface integral of the normal component over a closed surface is equal to the total charge contained within it.

$$\oint \mathbf{D}\cdot\mathbf{n}\,da = q. \qquad (4)$$

The charge is distributed throughout the transition layer with a density ρ. As the ends of the cylinder shrink together, the total charge q remains constant, for it cannot be destroyed, and

$$q = \rho\,\Delta l\,\Delta a. \qquad (5)$$

In the limit as $\Delta l \rightarrow 0$, the volume density ρ becomes infinite. It is then convenient to replace the product $\rho\,\Delta l$ by a *surface density* ω, defined as the charge per unit area. The transition of the normal component of the vector $\mathbf{D}$ across any surface S is now given by

$$(\mathbf{D}_2 - \mathbf{D}_1) \cdot \mathbf{n} = \omega. \tag{6}$$

The presence of a layer of charge on S results in an abrupt change in the normal component of $\mathbf{D}$, *the amount of the discontinuity being equal to the surface density measured in coulombs per square meter.*

Turning now to the behavior of the tangential components we replace the cylinder of Fig. 2*a* by a rectangular path drawn as in Fig. 2*b*. The sides of the rectangle of length Δs lie in either face of the transition layer and the ends which penetrate the layer are equal in length to its thickness Δl. This rectangle constitutes a contour C_0 about which

$$\int_{C_0} \mathbf{E} \cdot d\mathbf{s} + \int_{S_0} \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n}_0 \, da = 0, \tag{7}$$

where S_0 is the area of the rectangle and $\mathbf{n}_0$ its positive normal. The direction of this positive normal is determined, as in Fig. 1, page 8, by the direction of circulation about C_0. Let $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ be unit vectors in the direction of circulation along the lower and upper sides of the rectangle as shown. Neglecting differentials of higher order, one may approximate (7) by

$$(\mathbf{E} \cdot \boldsymbol{\tau}_1 + \mathbf{E} \cdot \boldsymbol{\tau}_2) \Delta s + \text{contributions from ends} = -\frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n}_0 \, \Delta s \, \Delta l. \tag{8}$$

As the layer contracts to the surface S, the contributions from the segments at the ends, which are proportional to Δl, become vanishingly small. If $\mathbf{n}$ is again the positive normal to S drawn from (1) into (2), we may define the unit tangent vector $\boldsymbol{\tau}$ by

$$\boldsymbol{\tau} = \mathbf{n}_0 \times \mathbf{n}. \tag{9}$$

Since

$$\mathbf{n}_0 \times \mathbf{n} \cdot \mathbf{E} = \mathbf{n}_0 \cdot \mathbf{n} \times \mathbf{E}, \tag{10}$$

we have in the limit as $\Delta l \to 0$, $\Delta s \to 0$,

$$\mathbf{n}_0 \cdot \left[\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) + \lim_{\Delta l \to 0} \left(\frac{\partial \mathbf{B}}{\partial t} \Delta l \right) \right] = 0. \tag{11}$$

The orientation of the rectangle — and hence also of $\mathbf{n}_0$ — is entirely arbitrary, from which it follows that the bracket in (11) must equal zero, or

$$\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = -\lim_{\Delta l \to 0} \frac{\partial \mathbf{B}}{\partial t} \Delta l. \tag{12}$$

The field vectors and their derivatives have been assumed to be bounded; consequently the right-hand side of (12) vanishes with Δl.

$$\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0. \tag{13}$$

The transition of the tangential components of the vector $\mathbf{E}$ *through a surface of discontinuity is continuous.*

The behavior of $\mathbf{H}$ at the boundary may be deduced immediately from (12) and the field equation

$$\int_{C_0} \mathbf{H} \cdot d\mathbf{s} - \int_{S_0} \frac{\partial \mathbf{D}}{\partial t} \cdot \mathbf{n}_0 \, da = \int_{S_0} \mathbf{J} \cdot \mathbf{n}_0 \, da. \tag{14}$$

We have

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \lim_{\Delta l \to 0} \left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \right) \Delta l. \tag{15}$$

The first term on the right of (15) vanishes as $\Delta l \to 0$ because $\mathbf{D}$ and its derivatives are bounded. If the current density $\mathbf{J}$ is finite, the second term vanishes as well. It may happen, however, that the current $I = \mathbf{J} \cdot \mathbf{n}_0 \, \Delta s \, \Delta l$ through the rectangle is squeezed into an infinitesimal layer on the surface S as the sides are brought together. It is convenient to represent this surface current by a surface density $\mathbf{K}$ defined as the limit of the product $\mathbf{J} \, \Delta \, l$ as $\Delta l \to 0$ and $\mathbf{J} \to \infty$. Then

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{K}. \tag{16}$$

When the conductivities of the contiguous media are finite, there can be no surface current, for $\mathbf{E}$ is bounded and hence the product $\sigma \mathbf{E} \, \Delta l$ vanishes with Δl. In this case, which is the usual one,

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = 0, \quad \text{(finite conductivity)}. \tag{17}$$

Not infrequently, however, it is necessary to assume the conductivity of a body to be infinite in order to simplify the analysis of its field. One must then apply (16) as a boundary condition rather than (17).

Summarizing, we are now able to supplement the field equations by four relations which determine the transition of an electromagnetic field from one medium to another separated by a surface of discontinuity.

$$\begin{aligned} \mathbf{n} \cdot (\mathbf{B}_2 - \mathbf{B}_1) &= 0, \qquad & \mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) &= \mathbf{K}, \\ \mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) &= 0, \qquad & \mathbf{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) &= \omega. \end{aligned} \tag{18}$$

From them follow immediately the conditions for the transition of the normal components of $\mathbf{E}$ and $\mathbf{H}$.

$$\mathbf{n} \cdot \left(\mathbf{H}_2 - \frac{\mu_1}{\mu_2} \mathbf{H}_1 \right) = 0, \qquad \mathbf{n} \cdot \left(\mathbf{E}_2 - \frac{\epsilon_1}{\epsilon_2} \mathbf{E}_1 \right) = \frac{\omega}{\epsilon_2}. \tag{19}$$

Likewise the tangential components of **D** and **B** must satisfy

$$\mathbf{n} \times \left(\mathbf{D}_2 - \frac{\epsilon_2}{\epsilon_1}\mathbf{D}_1\right) = 0, \qquad \mathbf{n} \times \left(\mathbf{B}_2 - \frac{\mu_2}{\mu_1}\mathbf{B}_1\right) = \mu_2\mathbf{K}. \tag{20}$$

COORDINATE SYSTEMS

1.14. Unitary and Reciprocal Vectors.—It is one of the principal advantages of vector calculus that the equations defining properties common to all electromagnetic fields may be formulated without reference to any particular system of coordinates. To determine the peculiarities that distinguish a given field from all other possible fields, it becomes necessary, unfortunately, to resolve each vector equation into an equivalent scalar system in appropriate coordinates.

In a given region let

$$u^1 = f_1(x, y, z), \qquad u^2 = f_2(x, y, z), \qquad u^3 = f_3(x, y, z), \tag{1}$$

be three independent, continuous, single-valued functions of the rectangular coordinates x, y, z. These equations may be solved with respect to x, y, z, and give

$$x = \varphi_1(u^1, u^2, u^3), \qquad y = \varphi_2(u^1, u^2, u^3), \qquad z = \varphi_3(u^1, u^2, u^3), \tag{2}$$

three functions which are also independent and continuous, and which are single-valued within certain limits. In general the functions φ_i as well as the functions f_i are continuously differentiable, but at certain singular points this property may fail and due care must be exercised in the application of general formulas.

With each point $P(x, y, z)$ in the region there is associated by means of (1) a triplet of values u^1, u^2, u^3; inversely (within limits depending on the boundaries of the region) there corresponds to each triplet u^1, u^2, u^3 a definite point. The functions u^1, u^2, u^3 are called *general* or *curvilinear coordinates.* Through each point P there pass three surfaces

$$u^1 = \text{constant}, \qquad u^2 = \text{constant}, \qquad u^3 = \text{constant}, \tag{3}$$

called the coordinate surfaces. On each coordinate surface one coordinate is constant and two are variable. A surface will be designated by the coordinate which is constant. Two surfaces intersect in a curve, called a coordinate curve, along which two coordinates are constant and one is variable. A coordinate curve will be designated by the variable coordinate.

Let **r** denote the vector from an arbitrary origin to a variable point $P(x, y, z)$. The point, and consequently also its position vector **r**, may be considered functions of the curvilinear coordinates u^1, u^2, u^3.

$$\mathbf{r} = \mathbf{r}(u^1, u^2, u^3). \tag{4}$$

A differential change in **r** due to small displacements along the coordinate curves is expressed by

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u^1} du^1 + \frac{\partial \mathbf{r}}{\partial u^2} du^2 + \frac{\partial \mathbf{r}}{\partial u^3} du^3. \tag{5}$$

Now if one moves unit distance along the u^1-curve, the change in **r** is directed tangentially to this curve and is equal to $\partial\mathbf{r}/\partial u^1$. The vectors

$$\mathbf{a}_1 = \frac{\partial \mathbf{r}}{\partial u^1}, \qquad \mathbf{a}_2 = \frac{\partial \mathbf{r}}{\partial u^2}, \qquad \mathbf{a}_3 = \frac{\partial \mathbf{r}}{\partial u^3}, \tag{6}$$

are known as the *unitary vectors* associated with the point P. They constitute a base system of reference for all other vectors associated with that particular point.

$$d\mathbf{r} = \mathbf{a}_1\, du^1 + \mathbf{a}_2\, du^2 + \mathbf{a}_3\, du^3. \tag{7}$$

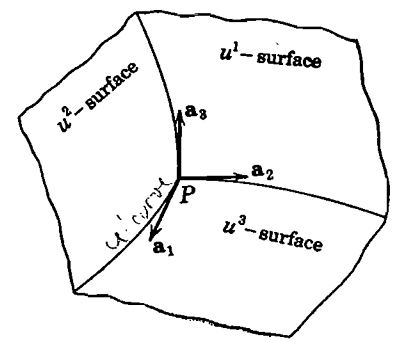

FIG. 3.—Base vectors for a curvilinear coordinate system.

It must be carefully noted that *the unitary vectors are not necessarily of unit length,* and their dimensions will depend on the nature of the general coordinates.

The three base vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ define a parallelepiped whose volume is

$$V = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) = \mathbf{a}_2 \cdot (\mathbf{a}_3 \times \mathbf{a}_1) = \mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2). \tag{8}$$

The three vectors of a new triplet defined by

$$\mathbf{a}^1 = \frac{1}{V}(\mathbf{a}_2 \times \mathbf{a}_3), \qquad \mathbf{a}^2 = \frac{1}{V}(\mathbf{a}_3 \times \mathbf{a}_1), \qquad \mathbf{a}^3 = \frac{1}{V}(\mathbf{a}_1 \times \mathbf{a}_2), \tag{9}$$

are respectively perpendicular to the planes determined by the pairs $(\mathbf{a}_2, \mathbf{a}_3)$, $(\mathbf{a}_3, \mathbf{a}_1)$, $(\mathbf{a}_1, \mathbf{a}_2)$. Upon forming all possible scalar products of the form $\mathbf{a}^i \cdot \mathbf{a}_j$, it is easy to see that they satisfy the condition

$$\mathbf{a}^i \cdot \mathbf{a}_j = \delta_{ij}, \tag{10}$$

where δ_{ij} is a commonly used symbol denoting unity when $i = j$, and zero when $i \neq j$. The unitary vectors can be expressed in terms of the system $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3$ by relations identical in form.

$$\mathbf{a}_1 = \frac{1}{V}(\mathbf{a}^2 \times \mathbf{a}^3), \qquad \mathbf{a}_2 = \frac{1}{V}(\mathbf{a}^3 \times \mathbf{a}^1), \qquad \mathbf{a}_3 = \frac{1}{V}(\mathbf{a}^1 \times \mathbf{a}^2). \tag{11}$$

Any two sets of noncoplanar vectors related by the Eqs. (8) to (11) are said to constitute *reciprocal systems.* The triplets $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3$ are called *reciprocal unitary vectors* and they may serve as a base system quite as well as the unitary vectors themselves.

If the reciprocal unitary vectors are employed as a base system, the differential $d\mathbf{r}$ will be written

$$d\mathbf{r} = \mathbf{a}^1\,du_1 + \mathbf{a}^2\,du_2 + \mathbf{a}^3\,du_3. \tag{12}$$

The differentials du_1, du_2, du_3 are evidently components of $d\mathbf{r}$ in the directions defined by the new base vectors. The quantities u_1, u_2, u_3 are functions of the coordinates u^1, u^2, u^3, but the differentials du_1, du_2, du_3 are not necessarily perfect. On the contrary they are related to the differentials of the coordinates by a set of linear equations which in general are nonintegrable. Thus equating (7) and (12), we have

$$d\mathbf{r} = \sum_{i=1}^{3} \mathbf{a}_i\,du^i = \sum_{j=1}^{3} \mathbf{a}^j\,du_j. \tag{13}$$

Upon scalar multiplication of (13) by $\mathbf{a}^i$ and by $\mathbf{a}_j$ in turn, we find, thanks to (10):

$$du_j = \sum_{i=1}^{3} \mathbf{a}_j \cdot \mathbf{a}_i\,du^i, \qquad du^i = \sum_{j=1}^{3} \mathbf{a}^i \cdot \mathbf{a}^j\,du_j. \tag{14}$$

It is customary to represent the scalar products of the unitary vectors and those of the reciprocal unitary vectors by the symbols

$$g_{ij} = \mathbf{a}_i \cdot \mathbf{a}_j = g_{ji}, \tag{15}$$
$$g^{ij} = \mathbf{a}^i \cdot \mathbf{a}^j = g^{ji}. \tag{16}$$

The components of $d\mathbf{r}$ in the unitary and in the reciprocal base systems are then related by

$$du_j = \sum_{i=1}^{3} g_{ji}\,du^i, \qquad du^i = \sum_{j=1}^{3} g^{ij}\,du_j. \tag{17}$$

A fixed vector $\mathbf{F}$ at the point P may be resolved into components either with respect to the base system $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, or with respect to the reciprocal system $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$.

$$\mathbf{F} = \sum_{i=1}^{3} f^i \mathbf{a}_i = \sum_{j=1}^{3} f_j \mathbf{a}^j. \tag{18}$$

The components of $\mathbf{F}$ in the unitary system are evidently related to those in its reciprocal system by

$$f_j = \sum_{i=1}^{3} g_{ji} f^i, \qquad f^i = \sum_{j=1}^{3} g^{ij} f_j, \tag{19}$$

and in virtue of the orthogonality of the base vectors $\mathbf{a}_j$ with respect to the reciprocal set $\mathbf{a}^i$ as expressed by (10), we may also write

$$f^i = \mathbf{F} \cdot \mathbf{a}^i, \qquad f_j = \mathbf{F} \cdot \mathbf{a}_j. \tag{20}$$

It follows from this that (18) is equivalent to

$$\mathbf{F} = \sum_{i=1}^{3} (\mathbf{F} \cdot \mathbf{a}^i)\mathbf{a}_i = \sum_{j=1}^{3} (\mathbf{F} \cdot \mathbf{a}_j)\mathbf{a}^j. \tag{21}$$

The quantities f^i are said to be the *contravariant* components of the vector $\mathbf{F}$, while the components f_j are called *covariant*. A small letter has been used to designate these components to avoid confusion with the components F_1, F_2, F_3 of $\mathbf{F}$ with respect to a base system coinciding with the $\mathbf{a}_i$ but of *unit* length. It has been noted previously that the length and dimensions of the unitary vectors depend on the nature of the curvilinear coordinates. An appropriate set of *unit* vectors which, like the unitary set $\mathbf{a}_i$, are tangent to the u^i-curves, is defined by

$$\mathbf{i}_1 = \frac{\mathbf{a}_1}{\sqrt{\mathbf{a}_1 \cdot \mathbf{a}_1}} = \frac{1}{\sqrt{g_{11}}}\mathbf{a}_1, \qquad \mathbf{i}_2 = \frac{1}{\sqrt{g_{22}}}\mathbf{a}_2, \qquad \mathbf{i}_3 = \frac{1}{\sqrt{g_{33}}}\mathbf{a}_3, \tag{22}$$

and, hence,

$$\mathbf{F} = F_1\mathbf{i}_1 + F_2\mathbf{i}_2 + F_3\mathbf{i}_3, \tag{23}$$

with

$$F_i = \sqrt{g_{ii}}\,f^i. \tag{24}$$

The F_i are of the same dimensions as the vector $\mathbf{F}$ itself.

The vector $d\mathbf{r}$ represents an infinitesimal displacement from the point $P(u^1, u^2, u^3)$ to a neighboring point whose coordinates are $u^1 + du^1$, $u^2 + du^2$, $u^3 + du^3$. The magnitude of this displacement, which constitutes a line element, we shall denote by ds. Then

$$ds^2 = d\mathbf{r} \cdot d\mathbf{r} = \sum_{i=1}^{3}\sum_{j=1}^{3} \mathbf{a}_i \cdot \mathbf{a}_j\,du^i\,du^j = \sum_{i=1}^{3}\sum_{j=1}^{3} \mathbf{a}^i \cdot \mathbf{a}^j\,du_i\,du_j; \tag{25}$$

or, in the notation of (15) and (16),

$$ds^2 = \sum_{i,j=1}^{3} g_{ij}\,du^i\,du^j = \sum_{i,j=1}^{3} g^{ij}\,du_i\,du_j. \tag{26}$$

The g_{ij} and g^{ij} appear here as coefficients of two differential quadratic forms expressing the length of a line element in the space of the general coordinates u^i or of its reciprocal set u_i. They are commonly called the *metrical coefficients.*

It is now a relatively simple matter to obtain expressions for elements of arc, surface, and volume in a system of curvilinear coordinates. Let ds_1 be an infinitesimal displacement at $P(u^1, u^2, u^3)$ along the u^1-curve.

(27) $$d\mathbf{s}_1 = \mathbf{a}_1\, du^1, \qquad ds_1 = |d\mathbf{s}_1| = \sqrt{g_{11}}\, du^1.$$

Similarly, for elements of length along the u^2- and u^3-curves, we have

(28) $$ds_2 = \sqrt{g_{22}}\, du^2, \qquad ds_3 = \sqrt{g_{33}}\, du^3.$$

Consider next an infinitesimal parallelogram in the u^1-surface bounded by intersecting u^2- and u^3-curves as indicated in Fig. 4. The area of such an element is equal in magnitude to

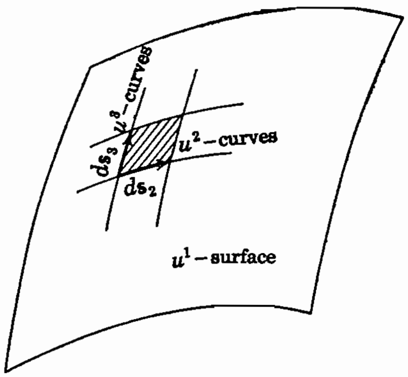

FIG. 4.—Element of area in the u^1-surface.

(29) $$da_1 = |d\mathbf{s}_2 \times d\mathbf{s}_3| = |\mathbf{a}_2 \times \mathbf{a}_3|\, du^2\, du^3 = \sqrt{(\mathbf{a}_2 \times \mathbf{a}_3)\cdot(\mathbf{a}_2 \times \mathbf{a}_3)}\, du^2\, du^3.$$

By a well-known vector identity

(30) $$(\mathbf{a} \times \mathbf{b})\cdot(\mathbf{c} \times \mathbf{d}) = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}),$$

where **a**, **b**, **c**, **d** are any four vectors, and hence

(31) $$(\mathbf{a}_2 \times \mathbf{a}_3)\cdot(\mathbf{a}_2 \times \mathbf{a}_3) = (\mathbf{a}_2\cdot\mathbf{a}_2)(\mathbf{a}_3\cdot\mathbf{a}_3) - (\mathbf{a}_2\cdot\mathbf{a}_3)(\mathbf{a}_3\cdot\mathbf{a}_2) = g_{22}g_{33} - g_{23}^2.$$

For the area of an element in the u^1-surface we have, therefore,

(32) $$da_1 = \sqrt{g_{22}g_{33} - g_{23}^2}\, du^2\, du^3,$$

and similarly for elements in the u^2- and u^3-surfaces,

(33) $$da_2 = \sqrt{g_{33}g_{11} - g_{31}^2}\, du^3\, du^1, \qquad da_3 = \sqrt{g_{11}g_{22} - g_{12}^2}\, du^1\, du^2.$$

Finally, a volume element bounded by coordinate surfaces is written as

(34) $$dv = d\mathbf{s}_1 \cdot d\mathbf{s}_2 \times d\mathbf{s}_3 = \mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3\, du^1\, du^2\, du^3.$$

If now in (21) we let $\mathbf{F} = \mathbf{a}_2 \times \mathbf{a}_3$, we have

(35) $$\mathbf{a}_2 \times \mathbf{a}_3 = (\mathbf{a}^1 \cdot \mathbf{a}_2 \times \mathbf{a}_3)\mathbf{a}_1 + (\mathbf{a}^2 \cdot \mathbf{a}_2 \times \mathbf{a}_3)\mathbf{a}_2 + (\mathbf{a}^3 \cdot \mathbf{a}_2 \times \mathbf{a}_3)\mathbf{a}_3,$$

or, on replacing the $\mathbf{a}^i$ by their values from (8) and (9),

(36) $$\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3 = \frac{\mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \cdot [(\mathbf{a}_2 \times \mathbf{a}_3 \cdot \mathbf{a}_2 \times \mathbf{a}_3)\mathbf{a}_1 + (\mathbf{a}_3 \times \mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3)\mathbf{a}_2 + (\mathbf{a}_1 \times \mathbf{a}_2 \cdot \mathbf{a}_2 \times \mathbf{a}_3)\mathbf{a}_3].$$

The quantities within parentheses can be expanded by (30) and the terms arranged in the form

(37) $$\begin{aligned}(\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3)^2 = \mathbf{a}_1 \cdot \mathbf{a}_1[(\mathbf{a}_2\cdot\mathbf{a}_2)(\mathbf{a}_3\cdot\mathbf{a}_3) - (\mathbf{a}_2\cdot\mathbf{a}_3)(\mathbf{a}_3\cdot\mathbf{a}_2)] \\ + \mathbf{a}_1 \cdot \mathbf{a}_2[(\mathbf{a}_2\cdot\mathbf{a}_3)(\mathbf{a}_3\cdot\mathbf{a}_1) - (\mathbf{a}_2\cdot\mathbf{a}_1)(\mathbf{a}_3\cdot\mathbf{a}_3)] \\ + \mathbf{a}_1 \cdot \mathbf{a}_3[(\mathbf{a}_2\cdot\mathbf{a}_1)(\mathbf{a}_3\cdot\mathbf{a}_2) - (\mathbf{a}_2\cdot\mathbf{a}_2)(\mathbf{a}_3\cdot\mathbf{a}_1)].\end{aligned}$$

If finally the scalar products in (37) are replaced by their g_{ij}, we obtain as an expression for a volume element

(38) $$dv = \sqrt{g}\, du^1\, du^2\, du^3,$$

in which

(39) $$g = \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix}.$$

A corresponding set of expressions for the elements of arc, area, and volume in the reciprocal base system may be obtained by replacing the g_{ij} by the g^{ij}, but they will not be needed in what follows.

Clearly the coefficients g_{ij} are sufficient to characterize completely the geometrical properties of space with respect to any curvilinear system of coordinates; it is therefore essential that we know how these coefficients may be determined. To unify our notation we shall represent the rectangular coordinates x, y, z of a point P by the letters x^1, x^2, x^3 respectively. Then

(40) $$ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2.$$

In this most elementary of all systems the metrical coefficients are

(41) $$g_{ij} = \delta_{ij}, \qquad (\delta_{ii} = 1, \qquad \delta_{ij} = 0 \text{ when } i \neq j).$$

From the orthogonality of the coordinate planes and the definition (9), it is evident that the unitary and the reciprocal unitary vectors are identical, are of unit length, and are the base vectors customarily represented by the letters **i**, **j**, **k**.

Suppose now that the rectangular coordinates are related functionally to a set of general coordinates as in (2) by the equations

(42) $$x^1 = x^1(u^1, u^2, u^3), \qquad x^2 = x^2(u^1, u^2, u^3), \qquad x^3 = x^3(u^1, u^2, u^3).$$

The differentials of the rectangular coordinates are linear functions of the differentials of the general coordinates, as we see upon differentiating Eqs. (42).

$$
\begin{aligned}
dx^1 &= \frac{\partial x^1}{\partial u^1} du^1 + \frac{\partial x^1}{\partial u^2} du^2 + \frac{\partial x^1}{\partial u^3} du^3, \\
dx^2 &= \frac{\partial x^2}{\partial u^1} du^1 + \frac{\partial x^2}{\partial u^2} du^2 + \frac{\partial x^2}{\partial u^3} du^3, \\
dx^3 &= \frac{\partial x^3}{\partial u^1} du^1 + \frac{\partial x^3}{\partial u^2} du^2 + \frac{\partial x^3}{\partial u^3} du^3.
\end{aligned} \tag{43}
$$

According to (26) and (40)

$$ds^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} g_{ij}\, du^i\, du^j = \sum_{k=1}^{3} (dx^k)^2, \tag{44}$$

whence on squaring the differentials in (43) and equating coefficients of like terms we obtain

$$g_{ij} = \frac{\partial x^1}{\partial u^i} \frac{\partial x^1}{\partial u^j} + \frac{\partial x^2}{\partial u^i} \frac{\partial x^2}{\partial u^j} + \frac{\partial x^3}{\partial u^i} \frac{\partial x^3}{\partial u^j}. \tag{45}$$

1.15. The Differential Operators.—The gradient of a scalar function $\phi(u^1, u^2, u^3)$ is a fixed vector defined in direction and magnitude as the maximum rate of change of ϕ with respect to the cordinates. The variation in ϕ incurred during a displacement $d\mathbf{r}$ is, therefore,

$$d\phi = \nabla\phi \cdot d\mathbf{r} = \sum_{i=1}^{3} \frac{\partial \phi}{\partial u^i} du^i. \tag{46}$$

Now the du^i are the contravariant components of the displacement vector $d\mathbf{r}$, and hence by (20),

$$du^i = \mathbf{a}^i \cdot d\mathbf{r}. \tag{47}$$

This value for du^i introduced into (46) leads to

$$\left(\nabla\phi - \sum_{i=1}^{3} \mathbf{a}^i \frac{\partial \phi}{\partial u^i}\right) \cdot d\mathbf{r} = 0 \tag{48}$$

and, since the displacement $d\mathbf{r}$ is arbitrary, we find for the gradient of a scalar function in any system of curvilinear coordinates:

$$\nabla\phi = \sum_{i=1}^{3} \mathbf{a}^i \frac{\partial \phi}{\partial u^i}. \tag{49}$$

In this expression the reciprocal unitary vectors constitute the base

system, but these may be replaced by the unitary vectors through the transformation

$$\mathbf{a}^i = \sum_{j=1}^{3} g^{ij} \mathbf{a}_j. \tag{50}$$

The divergence of a vector function $\mathbf{F}(u^1, u^2, u^3)$ at the point P may be deduced most easily from its definition in Eq. (9), page 4, as the limit of a surface integral of the normal component of $\mathbf{F}$ over a closed surface, per unit of enclosed volume. Consider those two ends of the volume element illustrated in Fig. 5 which lie in u^2-surfaces. The left end is located at u^2, the right at $u^2 + du^2$. The area of the face at u^2 is $(\mathbf{a}_1 \times \mathbf{a}_3) du^1\, du^3$, the order of the vectors being such that the normal is directed outward, *i.e.*, to the left. The net contribution of these two ends to the outward flux is, therefore,

$$[\mathbf{F} \cdot (\mathbf{a}_3 \times \mathbf{a}_1)\, du^1\, du^3]_{u^2+du^2} + [\mathbf{F} \cdot (\mathbf{a}_1 \times \mathbf{a}_3)\, du^1\, du^3]_{u^2}, \tag{51}$$

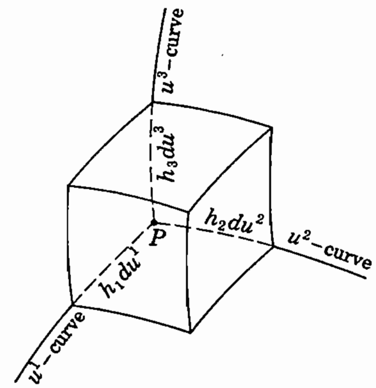

FIG. 5.—Element of volume in a curvilinear coordinate system.

the subscripts to the brackets indicating that the enclosed quantities are to be evaluated at $u^2 + du^2$ and u^2 respectively. For sufficiently small values of du^2, (51) may be approximated by the linear term of a Taylor expansion,

$$\frac{\partial}{\partial u^2} (\mathbf{F} \cdot \mathbf{a}_3 \times \mathbf{a}_1\, du^1\, du^2\, du^3), \tag{52}$$

$\mathbf{a}_1 \times \mathbf{a}_3$ having been replaced by $-\mathbf{a}_3 \times \mathbf{a}_1$. Now by (21), (20), and (37) we have

$$\mathbf{F} \cdot \mathbf{a}_3 \times \mathbf{a}_1 = \mathbf{F} \cdot \mathbf{a}^2(\mathbf{a}_2 \cdot \mathbf{a}_3 \times \mathbf{a}_1) = f^2 \sqrt{g}; \tag{53}$$

hence the contribution of the two ends to the surface integral is

$$\frac{\partial}{\partial u^2} (f^2 \sqrt{g})\, du^1\, du^2\, du^3. \tag{54}$$

Analogous contributions result from the two remaining pairs of faces. These are to be measured per unit volume; hence we divide by $dv = \sqrt{g}\, du^1\, du^2\, du^3$ and pass to the limit $du^1 \to 0$, $du^2 \to 0$, $du^3 \to 0$, ensuring thereby the vanishing of all but the linear terms in the Taylor expansion. The divergence of a vector $\mathbf{F}$ referred to a system of curvilinear coordinates is, therefore,

$$\nabla \cdot \mathbf{F} = \frac{1}{\sqrt{g}} \sum_{i=1}^{3} \frac{\partial}{\partial u^i} (f^i \sqrt{g}). \tag{55}$$

The curl of the vector $\mathbf{F}$ is found in the same manner by calculating the line integral of $\mathbf{F}$ around an infinitesimal closed path. According to Eq. (21), page 7, the component of the curl in a direction defined by a unit normal $\mathbf{n}$ is

$$(\nabla \times \mathbf{F}) \cdot \mathbf{n} = \lim_{C \to 0} \frac{1}{S} \int_C \mathbf{F} \cdot d\mathbf{s}. \tag{56}$$

Let us take the line integral of $\mathbf{F}$ about the contour of a rectangular element of area located in the u^1-surface, as indicated in Fig. 6. The sides of the rectangle are $\mathbf{a}_2\, du^2$ and $\mathbf{a}_3\, du^3$. The direction of circulation is such that the positive normal is in the sense of the positive u^1-curve. The contribution from the sides parallel to u^3-curves is

$$(\mathbf{F} \cdot \mathbf{a}_3\, du^3)_{u^2+du^2} - (\mathbf{F} \cdot \mathbf{a}_3\, du^3)_{u^2};$$

from the bottom and top parallel to u^2-curves, we obtain

$$-(\mathbf{F} \cdot \mathbf{a}_2\, du^2)_{u^3+du^3} + (\mathbf{F} \cdot \mathbf{a}_2\, du^2)_{u^3}.$$

Approximating these differences by the linear terms of a Taylor expansion, we obtain for the line integral

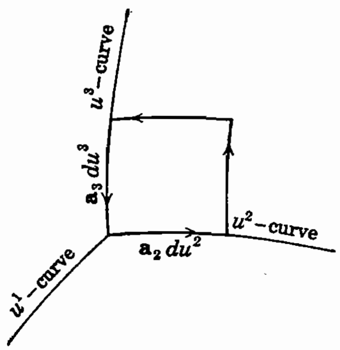

FIG. 6.—Calculation of the curl in curvilinear coordinates.

$$\left[\frac{\partial}{\partial u^2} (\mathbf{F} \cdot \mathbf{a}_3) - \frac{\partial}{\partial u^3} (\mathbf{F} \cdot \mathbf{a}_2)\right] du^2\, du^3. \tag{57}$$

This quantity must now be divided by the area of the rectangle, or $\sqrt{(\mathbf{a}_2 \times \mathbf{a}_3) \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}\, du^2\, du^3$. As for the unit normal $\mathbf{n}$, we note that the reciprocal vector $\mathbf{a}^1$, not the unitary vector $\mathbf{a}_1$, is always normal to the u^1-surface. Its magnitude must be unity; hence

$$\mathbf{n} = \frac{\mathbf{a}^1}{\sqrt{\mathbf{a}^1 \cdot \mathbf{a}^1}}. \tag{58}$$

These values introduced into (56) now lead to

$$(\nabla \times \mathbf{F}) \cdot \frac{\mathbf{a}^1}{\sqrt{\mathbf{a}^1 \cdot \mathbf{a}^1}} = \frac{1}{\sqrt{(\mathbf{a}_2 \times \mathbf{a}_3) \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}} \left[\frac{\partial}{\partial u^2} (\mathbf{F} \cdot \mathbf{a}_3) - \frac{\partial}{\partial u^3} (\mathbf{F} \cdot \mathbf{a}_2)\right]. \tag{59}$$

By (9) and (37)

$$\mathbf{a}_2 \times \mathbf{a}_3 = [\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)]\mathbf{a}^1 = \sqrt{g}\, a^1; \tag{60}$$

hence (59) reduces to

$$(\nabla \times \mathbf{F}) \cdot \mathbf{a}^1 = \frac{1}{\sqrt{g}}\left[\frac{\partial}{\partial u^2} (\mathbf{F} \cdot \mathbf{a}_3) - \frac{\partial}{\partial u^3} (\mathbf{F} \cdot \mathbf{a}_2)\right]. \tag{61}$$

The two remaining components of $\nabla \times \mathbf{F}$ are obtained from (61) by permutation of indices. Then by (21)

$$\nabla \times \mathbf{F} = \sum_{i=1}^{3} (\nabla \times \mathbf{F} \cdot \mathbf{a}^i)\mathbf{a}_i. \tag{62}$$

Remembering that $\mathbf{F} \cdot \mathbf{a}_i$ is the covariant component f_i, we have for the curl of a vector with respect to a set of general coordinates

$$\nabla \times \mathbf{F} = \frac{1}{\sqrt{g}}\left[\left(\frac{\partial f_3}{\partial u^2} - \frac{\partial f_2}{\partial u^3}\right)\mathbf{a}_1 + \left(\frac{\partial f_1}{\partial u^3} - \frac{\partial f_3}{\partial u^1}\right)\mathbf{a}_2 + \left(\frac{\partial f_2}{\partial u^1} - \frac{\partial f_1}{\partial u^2}\right)\mathbf{a}_3\right]. \tag{63}$$

Finally, we consider the operation $\nabla^2\phi$, by which we must understand $\nabla \cdot \nabla\phi$. We need only let $\mathbf{F} = \nabla\phi$ in (55). The contravariant components of the gradient are

$$f^i = \mathbf{F} \cdot \mathbf{a}^i = \sum_{j=1}^{3} \mathbf{a}^i \cdot \mathbf{a}^j \frac{\partial \phi}{\partial u^j} = \sum_{j=1}^{3} g^{ij} \frac{\partial \phi}{\partial u^j}. \tag{64}$$

Then

$$\nabla \cdot \nabla\phi = \nabla^2\phi = \frac{1}{\sqrt{g}} \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial}{\partial u^i}\left(g^{ij}\sqrt{g}\, \frac{\partial \phi}{\partial u^j}\right). \tag{65}$$

1.16. Orthogonal Systems.—Thus far no restriction has been imposed on the base vectors other than that they shall be noncoplanar. Now it happens that in almost all cases only the orthogonal systems can be usefully applied, and these allow a considerable simplification of the formulas derived above. Oblique systems might well be of the greatest practical importance; but they lead, unfortunately, to partial differential equations which cannot be mastered by present-day analysis.

The unitary vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ of an orthogonal system are by definition mutually perpendicular, whence it follows that $\mathbf{a}^i$ is parallel to $\mathbf{a}_i$ and is its reciprocal in magnitude.

$$\mathbf{a}^i = \frac{1}{\mathbf{a}_i \cdot \mathbf{a}_i}\, \mathbf{a}_i = \frac{1}{g_{ii}}\, \mathbf{a}_i. \tag{66}$$

Furthermore

$$\mathbf{a}_1 \cdot \mathbf{a}_2 = \mathbf{a}_2 \cdot \mathbf{a}_3 = \mathbf{a}_3 \cdot \mathbf{a}_1 = 0; \tag{67}$$

hence $g_{ij} = 0$, when $i \neq j$. It is customary in this orthogonal case to introduce the abbreviations

$$(68) \qquad h_1 = \sqrt{g_{11}}, \qquad h_2 = \sqrt{g_{22}}, \qquad h_3 = \sqrt{g_{33}},$$

$$(69) \qquad g^{ii} = \frac{1}{g_{ii}} = \frac{1}{h_i^2}.$$

The h_i may be calculated from the formula

$$(70) \qquad h_i^2 = \left(\frac{\partial x^1}{\partial u^i}\right)^2 + \left(\frac{\partial x^2}{\partial u^i}\right)^2 + \left(\frac{\partial x^3}{\partial u^i}\right)^2,$$

although their value is usually obvious from the geometry of the system. The elementary cell bounded by coordinate surfaces is now a rectangular box whose edges are

$$(71) \qquad ds_1 = h_1\,du^1, \qquad ds_2 = h_2\,du^2, \qquad ds_3 = h_3\,du^3,$$

and whose volume is

$$(72) \qquad dv = h_1h_2h_3\,du^1\,du^2\,du^3.$$

All off-diagonal terms of the determinant for g vanish and hence

$$(73) \qquad \sqrt{g} = h_1h_2h_3.$$

The distinction between the contravariant and covariant components of a vector with respect to a unitary or reciprocal unitary base system is essential to an understanding of the invariant properties of the differential operators and of scalar and vector products. However, in a fixed reference system this distinction may usually be ignored. It is then convenient to express the vector **F** in terms of its components, or projections, F_1, F_2, F_3 on an orthogonal base system of *unit* vectors $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$. By (22) and (66)

$$(74) \qquad \mathbf{a}_i = h_i\mathbf{i}_i, \qquad \mathbf{a}^i = \frac{1}{h_i}\mathbf{i}_i.$$

In terms of the components F_i the contravariant and covariant components are

$$(75) \qquad f^i = \frac{1}{h_i}F_i, \qquad f_i = h_iF_i.$$

Also

$$(76) \qquad F = F_1\mathbf{i}_1 + F_2\mathbf{i}_2 + F_3\mathbf{i}_3,$$

$$(77) \qquad \mathbf{i}_j \cdot \mathbf{i}_k = \delta_{jk}.$$

The gradient, divergence, curl, and Laplacian in an orthogonal system of curvilinear coordinates can now be written down directly from the results of the previous section.

From (49) we have for the gradient

$$(78) \qquad \nabla\phi = \sum_{j=1}^{3} \frac{1}{h_j}\frac{\partial\phi}{\partial u^j}\mathbf{i}_j.$$

According to (55) the divergence of a vector **F** is

$$(79) \qquad \nabla\cdot\mathbf{F} = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial u^1}(h_2h_3F_1) + \frac{\partial}{\partial u^2}(h_3h_1F_2) + \frac{\partial}{\partial u^3}(h_1h_2F_3)\right].$$

For the curl of **F** we have by (63)

$$(80) \qquad \nabla\times\mathbf{F} = \frac{1}{h_2h_3}\left[\frac{\partial}{\partial u^2}(h_3F_3) - \frac{\partial}{\partial u^3}(h_2F_2)\right]\mathbf{i}_1 + \frac{1}{h_3h_1}\left[\frac{\partial}{\partial u^3}(h_1F_1) - \frac{\partial}{\partial u^1}(h_3F_3)\right]\mathbf{i}_2 + \frac{1}{h_1h_2}\left[\frac{\partial}{\partial u^1}(h_2F_2) - \frac{\partial}{\partial u^2}(h_1F_1)\right]\mathbf{i}_3.$$

It may be remarked that (80) is the expansion of the determinant

$$(81) \qquad \nabla\times\mathbf{F} = \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\mathbf{i}_1 & h_2\mathbf{i}_2 & h_3\mathbf{i}_3 \\ \frac{\partial}{\partial u^1} & \frac{\partial}{\partial u^2} & \frac{\partial}{\partial u^3} \\ h_1F_1 & h_2F_2 & h_3F_3 \end{vmatrix}.$$

Finally, the Laplacian of an *invariant scalar* ϕ is

$$(82) \qquad \nabla^2\phi = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial u^1}\left(\frac{h_2h_3}{h_1}\frac{\partial\phi}{\partial u^1}\right) + \frac{\partial}{\partial u^2}\left(\frac{h_3h_1}{h_2}\frac{\partial\phi}{\partial u^2}\right) + \frac{\partial}{\partial u^3}\left(\frac{h_1h_2}{h_3}\frac{\partial\phi}{\partial u^3}\right)\right].$$

By an invariant scalar is meant a quantity such as temperature or energy which is invariant to a rotation of the coordinate system. The components, or measure numbers, F_i of a vector **F** are scalars, but they transform with a transformation of the base vectors. Now in the analysis of the field we encounter frequently the operation

$$(83) \qquad \nabla\times\nabla\times\mathbf{F} = \nabla\nabla\cdot\mathbf{F} - \nabla\cdot\nabla\mathbf{F}.$$

No meaning has been attributed as yet to $\nabla\cdot\nabla\mathbf{F}$. In a rectangular, Cartesian system of coordinates x^1, x^2, x^3, it is clear that this operation is equivalent to

$$(84) \qquad \nabla\cdot\nabla\mathbf{F} = \nabla^2\mathbf{F} = \sum_{j=1}^{3}\left(\frac{\partial^2F_j}{\partial(x^1)^2} + \frac{\partial^2F_j}{\partial(x^2)^2} + \frac{\partial^2F_j}{\partial(x^3)^2}\right)\mathbf{i}_j,$$

i.e., the Laplacian acting on the rectangular components of **F**. In generalized coordinates $\nabla \times \nabla \times \mathbf{F}$ is represented by the determinant

$$(85)\qquad \nabla \times \nabla \times \mathbf{F} = \begin{vmatrix} \dfrac{1}{h_2h_3}\,\mathbf{i}_1 & \dfrac{1}{h_3h_1}\,\mathbf{i}_2 & \dfrac{1}{h_1h_2}\,\mathbf{i}_3 \\ \dfrac{\partial}{\partial u^1} & \dfrac{\partial}{\partial u^2} & \dfrac{\partial}{\partial u^3} \\ \dfrac{h_1}{h_2h_3}\left[\dfrac{\partial}{\partial u^2}(h_3F_3) - \dfrac{\partial}{\partial u^3}(h_2F_2)\right] & \dfrac{h_2}{h_3h_1}\left[\dfrac{\partial}{\partial u^3}(h_1F_1) - \dfrac{\partial}{\partial u^1}(h_3F_3)\right] & \dfrac{h_3}{h_1h_2}\left[\dfrac{\partial}{\partial u^1}(h_2F_2) - \dfrac{\partial}{\partial u^2}(h_1F_1)\right] \end{vmatrix}.$$

The vector $\nabla \cdot \nabla\mathbf{F}$ may now be obtained by subtraction of (85) from the expansion of $\nabla\nabla \cdot \mathbf{F}$, and the result differs from that which follows a direct application of the Laplacian operator to the curvilinear components of $\mathbf{F}$.

1.17. The Field Equations in General Orthogonal Coordinates.—In any orthogonal system of curvilinear coordinates characterized by the coefficients h_1, h_2, h_3, the Maxwell equations can be resolved into a set of eight partial differential equations relating the scalar components of the field vectors.

$$(\mathrm{I})\qquad \begin{aligned} &\frac{1}{h_2h_3}\left[\frac{\partial}{\partial u^2}(h_3E_3) - \frac{\partial}{\partial u^3}(h_2E_2)\right] + \frac{\partial B_1}{\partial t} = 0. \\ &\frac{1}{h_3h_1}\left[\frac{\partial}{\partial u^3}(h_1E_1) - \frac{\partial}{\partial u^1}(h_3E_3)\right] + \frac{\partial B_2}{\partial t} = 0. \\ &\frac{1}{h_1h_2}\left[\frac{\partial}{\partial u^1}(h_2E_2) - \frac{\partial}{\partial u^2}(h_1E_1)\right] + \frac{\partial B_3}{\partial t} = 0. \end{aligned}$$

$$(\mathrm{II})\qquad \begin{aligned} &\frac{1}{h_2h_3}\left[\frac{\partial}{\partial u^2}(h_3H_3) - \frac{\partial}{\partial u^3}(h_2H_2)\right] - \frac{\partial D_1}{\partial t} = J_1. \\ &\frac{1}{h_3h_1}\left[\frac{\partial}{\partial u^3}(h_1H_1) - \frac{\partial}{\partial u^1}(h_3H_3)\right] - \frac{\partial D_2}{\partial t} = J_2. \\ &\frac{1}{h_1h_2}\left[\frac{\partial}{\partial u^1}(h_2H_2) - \frac{\partial}{\partial u^2}(h_1H_1)\right] - \frac{\partial D_3}{\partial t} = J_3. \end{aligned}$$

$$(\mathrm{III})\qquad \frac{\partial}{\partial u^1}(h_2h_3B_1) + \frac{\partial}{\partial u^2}(h_3h_1B_2) + \frac{\partial}{\partial u^3}(h_1h_2B_3) = 0.$$

$$(\mathrm{IV})\qquad \frac{\partial}{\partial u^1}(h_2h_3D_1) + \frac{\partial}{\partial u^2}(h_3h_1D_2) + \frac{\partial}{\partial u^3}(h_1h_2D_3) = h_1h_2h_3\rho.$$

It is not feasible to solve this system simultaneously in such a manner as to separate the components of the field vectors and to obtain equations satisfied by each individually. In any given problem one must make the most of whatever advantages and peculiarities the various coordinate systems have to offer.

1.18. Properties of Some Elementary Systems.—An orthogonal coordinate system has been shown to be completely characterized by the three metrical coefficients, h_1, h_2, h_3. These parameters will now be determined for certain elementary systems and in a few cases the differential operators set down for convenient reference.

1. *Cylindrical Coordinates.*—Let P' be the projection of a point $P(x, y, z)$ on the z-plane and r, θ be the polar coordinates of P' in this plane (Fig. 7). The variables

$$(86)\qquad u^1 = r, \qquad u^2 = \theta, \qquad u^3 = z,$$

are called circular cylindrical coordinates. They are related to the rectangular coordinates by the equations

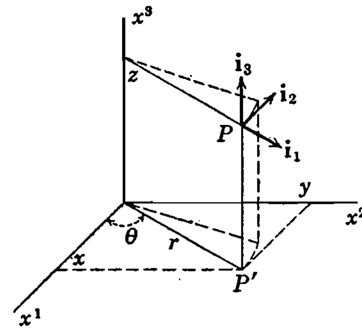

FIG. 7.—Coordinates of the circular cylinder.

$$(87)\qquad x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z.$$

The coordinate surfaces are coaxial cylinders of circular cross section intersected orthogonally by the planes θ = constant and z = constant. The infinitesimal line element is

$$(88)\qquad ds^2 = dr^2 + r^2\,d\theta^2 + dz^2,$$

whence it is apparent that the metrical coefficients are

$$(89)\qquad h_1 = 1, \qquad h_2 = r, \qquad h_3 = 1.$$

If ψ is any scalar and $\mathbf{F}$ a vector function we find:

$$(90)\qquad \begin{aligned} \nabla\psi &= \frac{\partial\psi}{\partial r}\,\mathbf{i}_1 + \frac{1}{r}\frac{\partial\psi}{\partial\theta}\,\mathbf{i}_2 + \frac{\partial\psi}{\partial z}\,\mathbf{i}_3, \\ \nabla \cdot \mathbf{F} &= \frac{1}{r}\frac{\partial}{\partial r}(rF_1) + \frac{1}{r}\frac{\partial F_2}{\partial\theta} + \frac{\partial F_3}{\partial z}, \\ \nabla \times \mathbf{F} &= \left(\frac{1}{r}\frac{\partial F_3}{\partial\theta} - \frac{\partial F_2}{\partial z}\right)\mathbf{i}_1 + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial r}\right)\mathbf{i}_2 + \left[\frac{1}{r}\frac{\partial}{\partial r}(rF_2) - \frac{1}{r}\frac{\partial F_1}{\partial\theta}\right]\mathbf{i}_3, \\ \nabla^2\psi &= \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2} + \frac{\partial^2\psi}{\partial z^2}. \end{aligned}$$

2. *Spherical Coordinates.*—The variables

$$u^1 = r, \qquad u^2 = \theta, \qquad u^3 = \phi, \tag{91}$$

related to the rectangular coordinates by the transformation

$$x = r \sin\theta \cos\phi, \qquad y = r \sin\theta \sin\phi, \qquad z = r\cos\theta, \tag{92}$$

are called the spherical coordinates of the point P. The coordinate surfaces, r = constant, are concentric spheres intersected by meridian planes, ϕ = constant, and a family of cones, θ = constant. The unit vectors $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ are drawn in the direction of *increasing* r, θ, and ϕ such as to constitute a right-hand base system, as indicated in Fig. 8. The line element is

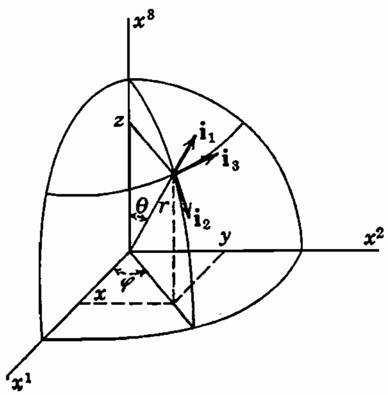

FIG. 8.—Spherical coordinates.

$$ds^2 = dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\phi^2, \tag{93}$$

whence for the metrical coefficients we obtain

$$h_1 = 1, \qquad h_2 = r, \qquad h_3 = r\sin\theta. \tag{94}$$

These values lead to

$$\begin{aligned}
\nabla\psi &= \frac{\partial\psi}{\partial r}\mathbf{i}_1 + \frac{1}{r}\frac{\partial\psi}{\partial\theta}\mathbf{i}_2 + \frac{1}{r\sin\theta}\frac{\partial\psi}{\partial\phi}\mathbf{i}_3,\\
\nabla\cdot\mathbf{F} &= \frac{1}{r^2}\frac{\partial}{\partial r}(r^2F_1) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta F_2) + \frac{1}{r\sin\theta}\frac{\partial F_3}{\partial\phi},\\
\nabla\times\mathbf{F} &= \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta F_3) - \frac{\partial F_2}{\partial\phi}\right]\mathbf{i}_1 + \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial F_1}{\partial\phi} - \frac{\partial}{\partial r}(rF_3)\right]\mathbf{i}_2 + \frac{1}{r}\left[\frac{\partial}{\partial r}(rF_2) - \frac{\partial F_1}{\partial\theta}\right]\mathbf{i}_3,\\
\nabla^2\psi &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}.
\end{aligned} \tag{95}$$

3. *Elliptic Coordinates.*—Let two fixed points P_1 and P_2 be located at $x = c$ and $x = -c$ on the x-axis and let r_1 and r_2 be the distances of a variable point P in the z-plane from P_1 and P_2. Then the variables

$$u^1 = \xi, \qquad u^2 = \eta, \qquad u^3 = z, \tag{96}$$

defined by equations

$$\xi = \frac{r_1 + r_2}{2c}, \qquad \eta = \frac{r_1 - r_2}{2c}, \tag{97}$$

are called elliptic coordinates. From these relations it is evident that

$$\xi \geqq 1, \qquad -1 \leqq \eta \leqq 1. \tag{98}$$

The coordinate surface, ξ = constant, is a cylinder of elliptic cross section, whose foci are P_1 and P_2. The semimajor and semiminor axes of an ellipse ξ are given by

$$a = c\xi, \qquad b = c\sqrt{\xi^2 - 1}, \tag{99}$$

and the eccentricity is

$$e = \frac{c}{a} = \frac{1}{\xi}. \tag{100}$$

The surfaces, η = constant, represent a family of confocal hyperbolic

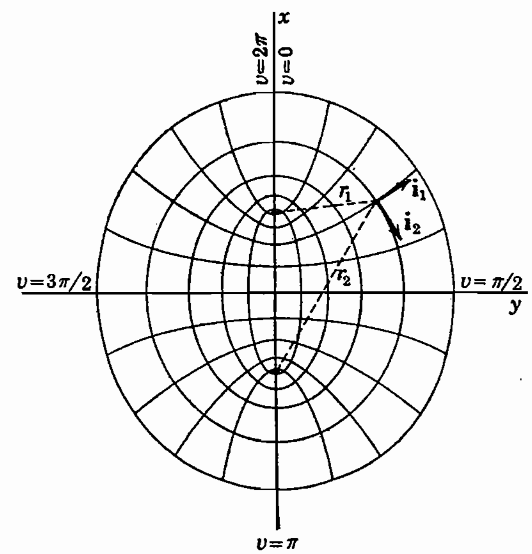

FIG. 9.—Coordinates of the elliptic cylinder. Ambiguity of sign is avoided by placing $\xi = \cosh u$, $\eta = \cos v$.

cylinders of two sheets as illustrated in Fig. 9. The equations of these two confocal systems are

$$\frac{x^2}{\xi^2} + \frac{y^2}{\xi^2 - 1} = c^2, \qquad \frac{x^2}{\eta^2} - \frac{y^2}{1 - \eta^2} = c^2, \tag{101}$$

from which we deduce the transformation

$$x = c\xi\eta, \qquad y = c\sqrt{(\xi^2 - 1)(1 - \eta^2)}, \qquad z = z. \tag{102}$$

The variable η corresponds to the cosine of an angle measured from the x-axis and the unit vectors $\mathbf{i}_1, \mathbf{i}_2$ of a right-hand base system are therefore drawn as indicated in Fig. 9, with $\mathbf{i}_3$ normal to the page and directed from the reader.

The metrical coefficients are calculated from (102) and (70), giving

$$h_1 = c\sqrt{\frac{\xi^2 - \eta^2}{\xi^2 - 1}}, \qquad h_2 = c\sqrt{\frac{\xi^2 - \eta^2}{1 - \eta^2}}, \qquad h_3 = 1. \tag{103}$$

4. *Parabolic Coordinates.*—If r, θ are polar coordinates of a variable point in the z-plane, one may define two mutually orthogonal families of parabolas by the equations

$$\xi = \sqrt{2r}\sin\frac{\theta}{2}, \qquad \eta = \sqrt{2r}\cos\frac{\theta}{2}. \tag{104}$$

The surfaces, ξ = constant and η = constant, are intersecting parabolic

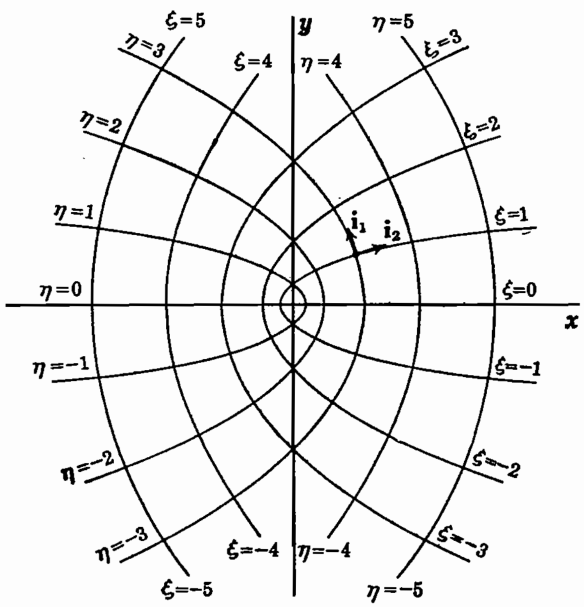

FIG. 10.—Parabolic coordinates.

cylinders whose elements are parallel to the z-axis as shown in Fig. 10. The parameters

$$u^1 = \xi, \qquad u^2 = \eta, \qquad u^3 = -z \tag{105}$$

are called parabolic coordinates. Upon replacing r and θ in (104) by rectangular coordinates we find

$$\xi^2 = \sqrt{x^2 + y^2} - x, \qquad \eta^2 = \sqrt{x^2 + y^2} + x, \tag{106}$$

whence for the transformation from rectangular to parabolic coordinates we have

$$x = \tfrac{1}{2}(\eta^2 - \xi^2), \qquad y = \xi\eta, \qquad z = -z. \tag{107}$$

The unit vectors $\mathbf{i}_1$ and $\mathbf{i}_2$ are directed as shown in Fig. 10, with $\mathbf{i}_3$ normal to the page and away from the reader. The calculation of the metrical coefficients from (107) and (70) leads to

$$h_1 = h_2 = \sqrt{\xi^2 + \eta^2}, \qquad h_3 = 1. \tag{108}$$

5. *Bipolar Coordinates.*—Let P_1 and P_2 be two fixed points in any z-plane with the coordinates $(a, 0)$, $(-a, 0)$ respectively. If ξ is a parameter, the equation

$$(x - a\coth\xi)^2 + y^2 = a^2\operatorname{csch}^2\xi, \tag{109}$$

describes two families of circles whose centers lie on the x-axis. These two families are symmetrical with respect to the y-axis as shown in

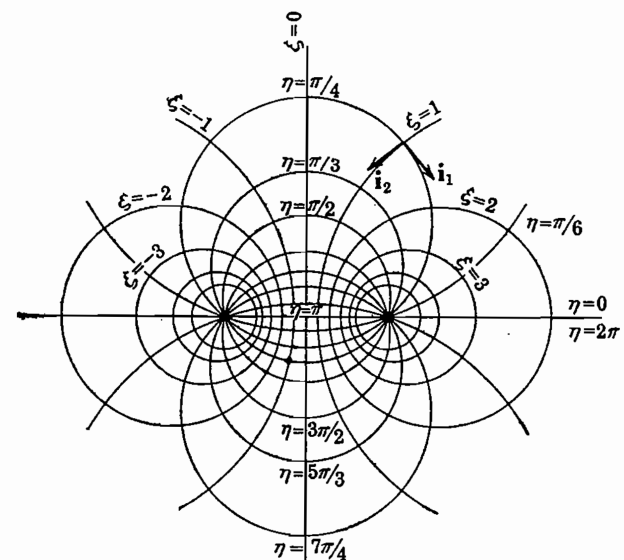

FIG. 11.—Bipolar coordinates.

Fig. 11. The point P_1 at $(a, 0)$ corresponds to $\xi = +\infty$, whereas its image P_2 at $(-a, 0)$ is approached when $\xi = -\infty$. The locus of (109), when $\xi = 0$, coincides with the y-axis. The orthogonal set is likewise a family of circles whose centers all lie on the y-axis and all of which pass through the fixed points P_1 and P_2. They are defined by the equation,

$$x^2 + (y - a\cot\eta)^2 = a^2\csc^2\eta, \tag{110}$$

wherein the parameter η is confined to the range $0 \leqq \eta \leqq 2\pi$. In order that the coordinates of a point P in a given quadrant shall be single-valued, each circle of this family is separated into two segments by the points P_1 and P_2. A value less than π is assigned to the arc above the x-axis, while the lower arc is denoted by a value of η equal to π plus the value of η assigned to the upper segment of the same circle.

The variables

$$(111) \qquad u^1 = \xi, \qquad u^2 = \eta, \qquad u^3 = z,$$

are called bipolar coordinates. From (109) and (110) the transformation to rectangular coordinates is found to be

$$(112) \qquad x = \frac{a \sinh \xi}{\cosh \xi - \cos \eta}, \qquad y = \frac{a \sin \eta}{\cosh \xi - \cos \eta}, \qquad z = z.$$

The unit vectors $\mathbf{i}_1$ and $\mathbf{i}_2$ are in the direction of increasing ξ and η as indicated in Fig. 11, while $\mathbf{i}_3$ is directed away from the reader along the z-axis. The calculation of the metrical coefficients yields

$$(113) \qquad h_1 = h_2 = \frac{a}{\cosh \xi - \cos \eta}, \qquad h_3 = 1.$$

6. *Spheroidal Coordinates.*—The coordinates of the elliptic cylinder were generated by translating a system of confocal ellipses along the z-axis. The spheroidal coordinates are obtained by rotation of the ellipses about an axis of symmetry. Two cases are to be distinguished, according to whether the rotation takes place about the major or about the minor axis. In Fig. 9 the major axes are oriented along the x-axis of a rectangular system. If the figure is rotated about this axis, a set of confocal prolate spheroids is generated whose orthogonal surfaces are hyperboloids of two sheets. If ϕ measures the angle of rotation from the y-axis in the x-plane and r the perpendicular distance of a point from the x-axis, so that

$$(114) \qquad y = r \cos \phi, \qquad z = r \sin \phi,$$

then the variables

$$(115) \qquad u^1 = \xi, \qquad u^2 = \eta, \qquad u^3 = \phi,$$

defined by (97) and (114) are called prolate spheroidal coordinates. In place of (101) we have for the equations of the two confocal systems

$$(116) \qquad \frac{x^2}{\xi^2} + \frac{r^2}{\xi^2 - 1} = c^2, \qquad \frac{x^2}{\eta^2} - \frac{r^2}{1 - \eta^2} = c^2,$$

from which we deduce

$$(117) \qquad x = c\xi\eta, \qquad y = c\sqrt{(\xi^2 - 1)(1 - \eta^2)} \cos \phi, \qquad z = c\sqrt{(\xi^2 - 1)(1 - \eta^2)} \sin \phi,$$

$$(118) \qquad \xi \geqq 1, \qquad -1 \leqq \eta \leqq 1, \qquad 0 \leqq \phi \leqq 2\pi.$$

A calculation of the metrical coefficients gives

$$(119) \qquad h_1 = c\sqrt{\frac{\xi^2 - \eta^2}{\xi^2 - 1}}, \quad h_2 = c\sqrt{\frac{\xi^2 - \eta^2}{1 - \eta^2}}, \quad h_3 = c\sqrt{(\xi^2 - 1)(1 - \eta^2)}.$$

When the ellipses of Fig. 9 are rotated about the y-axis, the spheroids are oblate and the focal points P_1, P_2 describe a circle in the plane $y = 0$. Let r, ϕ, y, be cylindrical coordinates about the y-axis,

$$(120) \qquad z = r \cos \phi, \qquad x = r \sin \phi.$$

If by P_1 and P_2 we now understand the points where the focal ring of radius c intercepts the plane ϕ = constant, the variables ξ and η are still defined by (97); but for the equations of the coordinate surfaces we have

$$(121) \qquad \frac{r^2}{\xi^2} + \frac{y^2}{\xi^2 - 1} = c^2, \qquad \frac{r^2}{\eta^2} - \frac{y^2}{1 - \eta^2} = c^2,$$

from which we deduce the transformation from oblate spheroidal coordinates

$$(122) \qquad u^1 = \xi, \qquad u^2 = \eta, \qquad u^3 = \phi,$$

to rectangular coordinates

$$(123) \qquad x = c\xi\eta \sin \phi, \qquad y = c\sqrt{(\xi^2 - 1)(1 - \eta^2)}, \qquad z = c\xi\eta \cos \phi.$$

The surfaces, ξ = constant, are oblate spheroids, whereas the orthogonal family, η = constant, are hyperboloids of one sheet. The metrical coefficients are

$$(124) \qquad h_1 = c\sqrt{\frac{\xi^2 - \eta^2}{\xi^2 - 1}}, \qquad h_2 = c\sqrt{\frac{\xi^2 - \eta^2}{1 - \eta^2}}, \qquad h_3 = c\xi\eta.$$

The practical utility of spheroidal coordinates may be surmised from the fact that as the eccentricity approaches unity the prolate spheroids become rod-shaped, whereas the oblate spheroids degenerate into flat, elliptic disks. In the limit, as the focal distance $2c$ and the eccentricity approach zero, the spheroidal coordinates go over into spherical coordinates, with $\xi \rightarrow r$, $\eta \rightarrow \cos \theta$.

7. *Paraboloidal Coordinates.*—Another set of rotational coordinates may be obtained by rotating the parabolas of Fig. 10 about their axis

of symmetry. The variables

$$u^1 = \xi, \qquad u^2 = \eta, \qquad u^3 = \phi, \tag{125}$$

defined by

$$x = \xi\eta \cos\phi, \qquad y = \xi\eta \sin\phi, \qquad z = \tfrac{1}{2}(\xi^2 - \eta^2), \tag{126}$$

are called paraboloidal coordinates. The surfaces, ξ = constant, η = constant, are paraboloids of revolution about an axis of symmetry which in this case has been taken coincident with the z-axis. The plane, $y = 0$, is cut by these surfaces along the curves

$$x^2 = 2\xi^2\left(\frac{\xi^2}{2} - z\right), \qquad x^2 = 2\eta^2\left(\frac{\eta^2}{2} + z\right), \tag{127}$$

which are evidently parabolas whose foci are located at the origin and whose parameters are ξ^2 and η^2. The metrical coefficients are

$$h_1 = h_2 = \sqrt{\xi^2 + \eta^2}, \qquad h_3 = \xi\eta. \tag{128}$$

8. *Ellipsoidal Coordinates.*—The equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \qquad (a > b > c), \tag{129}$$

is that of an ellipsoid whose semiprincipal axes are of length a, b, c. Then

$$\begin{aligned} \frac{x^2}{a^2+\xi} + \frac{y^2}{b^2+\xi} + \frac{z^2}{c^2+\xi} &= 1, \qquad (\xi > -c^2),\\ \frac{x^2}{a^2+\eta} + \frac{y^2}{b^2+\eta} + \frac{z^2}{c^2+\eta} &= 1, \qquad (-c^2 > \eta > -b^2),\\ \frac{x^2}{a^2+\zeta} + \frac{y^2}{b^2+\zeta} + \frac{z^2}{c^2+\zeta} &= 1, \qquad (-b^2 > \zeta > -a^2), \end{aligned} \tag{130}$$

are the equations respectively of an ellipsoid, a hyperboloid of one sheet, and a hyperboloid of two sheets, all confocal with the ellipsoid (129). Through each point of space there will pass just one surface of each kind, and to each point there will correspond a unique set of values for ξ, η, ζ. The variables

$$u^1 = \xi, \qquad u^2 = \eta, \qquad u^3 = \zeta, \tag{131}$$

are called ellipsoidal coordinates. The surface, ξ = constant, is a hyperboloid of one sheet and η = constant, a hyperboloid of two sheets. The transformation to rectangular coordinates is obtained by solving (130) simultaneously for x, y, z. This gives

$$\begin{aligned} x &= \pm\left[\frac{(\xi+a^2)(\eta+a^2)(\zeta+a^2)}{(b^2-a^2)(c^2-a^2)}\right]^{\frac{1}{2}},\\ y &= \pm\left[\frac{(\xi+b^2)(\eta+b^2)(\zeta+b^2)}{(c^2-b^2)(a^2-b^2)}\right]^{\frac{1}{2}},\\ z &= \pm\left[\frac{(\xi+c^2)(\eta+c^2)(\zeta+c^2)}{(a^2-c^2)(b^2-c^2)}\right]^{\frac{1}{2}}. \end{aligned} \tag{132}$$

The mutual orthogonality of the three families of surfaces may be verified by calculating the coefficients g_{ij} from (132) by means of (45). They are zero when $i \neq j$; for the diagonal terms we find

$$\begin{aligned} h_1 &= \frac{1}{2}\left[\frac{(\xi-\eta)(\xi-\zeta)}{(\xi+a^2)(\xi+b^2)(\xi+c^2)}\right]^{\frac{1}{2}},\\ h_2 &= \frac{1}{2}\left[\frac{(\eta-\zeta)(\eta-\xi)}{(\eta+a^2)(\eta+b^2)(\eta+c^2)}\right]^{\frac{1}{2}},\\ h_3 &= \frac{1}{2}\left[\frac{(\zeta-\xi)(\zeta-\eta)}{(\zeta+a^2)(\zeta+b^2)(\zeta+c^2)}\right]^{\frac{1}{2}}, \end{aligned} \tag{133}$$

It is convenient to introduce the abbreviation

$$R_s = \sqrt{(s+a^2)(s+b^2)(s+c^2)}, \qquad (s = \xi, \eta, \zeta). \tag{134}$$

For the Laplacian of a scalar ψ we then have

$$\begin{aligned} \nabla^2\psi = \frac{4}{(\xi-\eta)(\xi-\zeta)(\eta-\zeta)}\Big[(\eta-\zeta)R_\xi\frac{\partial}{\partial\xi}\left(R_\xi\frac{\partial\psi}{\partial\xi}\right)\\ + (\zeta-\xi)R_\eta\frac{\partial}{\partial\eta}\left(R_\eta\frac{\partial\psi}{\partial\eta}\right) + (\xi-\eta)R_\zeta\frac{\partial}{\partial\zeta}\left(R_\zeta\frac{\partial\psi}{\partial\zeta}\right)\Big]. \end{aligned} \tag{135}$$

THE FIELD TENSORS

1.19. Orthogonal Transformations and Their Invariants.—In the theory of relativity one undertakes the formulation of the laws of physics, and in particular the equations of the electromagnetic field, such that they are invariant to transformations of the system of reference. Although in the present volume we shall have no occasion to examine the foundations of the relativity theory, it will nevertheless prove occasionally advantageous to employ the symmetrical, four-dimensional notation introduced by Minkowski and Sommerfeld and to deduce the Lorentz transformation with respect to which the field equations are invariant. To discover quantities which are invariant to a transformation from one system of general curvilinear coordinates to another, it is essential that one distinguish between the covariant and contravariant components of vectors and between unitary and reciprocal unitary base systems. For our present purposes it will be sufficient, however, to confine the discussion to systems of rectangular, Cartesian coordinates in which, as we have seen, covariant and contravariant components are identical.[1]

Let $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ be three orthogonal, unit base vectors defining a rectangular coordinate system X whose origin is located at the fixed point O, and let $\mathbf{r}$ be the position vector of any point P with respect to O.

$$\mathbf{r} = x_1\mathbf{i}_1 + x_2\mathbf{i}_2 + x_3\mathbf{i}_3, \tag{1}$$

and since

$$\mathbf{i}_j \cdot \mathbf{i}_k = \delta_{jk}, \tag{2}$$

the coordinates of P in the system X are

$$x_k = \mathbf{r} \cdot \mathbf{i}_k. \tag{3}$$

Suppose now that $\mathbf{i}'_1, \mathbf{i}'_2, \mathbf{i}'_3$ are the base vectors of a second rectangular system X' whose origin coincides with O and which, therefore, differs from X only by a rotation of the coordinate axes. Since

$$\mathbf{r} = x'_1\mathbf{i}'_1 + x'_2\mathbf{i}'_2 + x'_3\mathbf{i}'_3, \tag{4}$$

the coordinates of P with respect to X' are

$$x'_j = \mathbf{r} \cdot \mathbf{i}'_j = x_1\mathbf{i}_1 \cdot \mathbf{i}'_j + x_2\mathbf{i}_2 \cdot \mathbf{i}'_j + x_3\mathbf{i}_3 \cdot \mathbf{i}'_j; \tag{5}$$

each coordinate of P in X' is a linear function of its coordinates in X, whereby the coefficients

$$a_{jk} = \mathbf{i}'_j \cdot \mathbf{i}_k \tag{6}$$

of the linear form are clearly the direction cosines of the coordinate axes of X' with respect to the axes of X. A rotation of a rectangular coordinate system effects a change in the coordinates of a point which may be represented by the linear transformation

$$x'_j = \sum_{k=1}^{3} a_{jk}x_k, \qquad (j = 1, 2, 3). \tag{7}$$

The coefficients a_{jk} are subject to certain conditions which are a consequence of the fact that the distance from O to P, that is to say, the

[1] This section is based essentially on the following papers: MINKOWSKI, *Ann. Physik*, **47**, 927, 1915; SOMMERFELD, *Ann. Physik*, **32**, 749, 1910 and **33**, 649, 1910; MIE, *Ann. Physik*, **37**, 511, 1912; PAULI, Relativitätstheorie, in the Encyklopädie der mathematischen Wissenschaften, Vol. V, part 2, p. 539, 1920.

magnitude of $\mathbf{r}$, is independent of the orientation of the coordinate system.

$$\sum_{j=1}^{3} (x_j)^2 = \sum_{j=1}^{3} (x'_j)^2. \tag{8}$$

$$\sum_{j=1}^{3} (x'_j)^2 = \sum_{j=1}^{3}\left(\sum_{i=1}^{3} a_{ji}x_i\right)\left(\sum_{k=1}^{3} a_{jk}x_k\right) = \sum_{i=1}^{3}\sum_{k=1}^{3} x_ix_k\left(\sum_{j=1}^{3} a_{ji}a_{jk}\right), \tag{9}$$

whence it follows that

$$\sum_{j=1}^{3} a_{ji}a_{jk} = \delta_{ik} = \begin{matrix} 1, & \text{when } i = k, \\ 0, & \text{when } i \neq k. \end{matrix} \tag{10}$$

Equation (10) expresses in fact the relations which must prevail among the cosines of the angles between coordinate axes in order that they be rectangular and which are, therefore, known as *conditions of orthogonality*. The system (7), when subject to (10), is likewise called an *orthogonal transformation*. As a direct consequence of (10), it may be shown that the square of the determinant $|a_{jk}|$ is equal to unity and hence $|a_{jk}| = \pm 1$. Any set of coefficients a_{jk} which satisfy (10) define an orthogonal transformation in the sense that the relation (8) is preserved. Geometrically the transformation (7) represents a rotation only when the determinant $|a_{jk}| = +1$. The orthogonal transformation whose determinant is equal to -1 corresponds to an inversion followed by a rotation.

Since the determinant of an orthogonal transformation does not vanish, the x_k may be expressed as linear functions of the x'_j. These relations are obtained most simply by writing as in (5):

$$x_k = \mathbf{r} \cdot \mathbf{i}_k = x'_1\mathbf{i}'_1 \cdot \mathbf{i}_k + x'_2\mathbf{i}'_2 \cdot \mathbf{i}_k + x'_3\mathbf{i}'_3 \cdot \mathbf{i}_k, \tag{11}$$

or

$$x_k = \sum_{j=1}^{3} a_{jk}x'_j, \qquad (k = 1, 2, 3), \tag{12}$$

whence it follows from (8) that

$$\sum_{j=1}^{3} a_{ij}a_{kj} = \delta_{ik}. \tag{13}$$

Let $\mathbf{A}$ be any fixed vector in space, so that

$$\mathbf{A} = \sum_{k=1}^{3} A_k\mathbf{i}_k = \sum_{k=1}^{3} A'_k\mathbf{i}'_k. \tag{14}$$

The component A'_j of this vector with respect to the system X' is given by

$$A'_j = \mathbf{A} \cdot \mathbf{i}'_j = \sum_{k=1}^{3} A_k\mathbf{i}_k \cdot \mathbf{i}'_j = \sum_{k=1}^{3} a_{jk}A_k; \tag{15}$$

thus the rectangular components of a fixed vector upon rotation of the coordinate system transform like the coordinates of a point. Now while every vector has in general three scalar components, it does not follow that any three scalar quantities constitute the components of a vector. *In order that three scalars A_1, A_2, A_3 may be interpreted as the components of a vector, it is necessary that they transform like the coordinates of a point.*

Among scalar quantities one must distinguish the *variant* from the *invariant*. Quantities such as temperature, pressure, work, and the like are independent of the orientation of the coordinate system and are, therefore, called invariant scalars. On the other hand the coordinates of a point, and the measure numbers, or components, of a vector have only magnitude, but they transform with the coordinate system itself. We know that the product $\mathbf{A} \cdot \mathbf{B}$ of two vectors $\mathbf{A}$ and $\mathbf{B}$ is a scalar, but a scalar of what kind? In virtue of (12) and (13) we have

$$\text{(16)} \qquad \mathbf{A} \cdot \mathbf{B} = \sum_{k=1}^{3} A_k B_k = \sum_{k=1}^{3} \Big(\sum_{j=1}^{3} a_{jk} A'_j \Big) \Big(\sum_{i=1}^{3} a_{ik} B'_i \Big) = \sum_{j=1}^{3} A'_j B'_j;$$

the scalar product of two vectors is invariant to an orthogonal transformation of the coordinate system.

Let ϕ be an invariant scalar and consider the triplet of quantities

$$\text{(17)} \qquad B_i = \frac{\partial \phi}{\partial x_i}, \qquad (i = 1, 2, 3).$$

Now by (12),

$$\text{(18)} \qquad \frac{\partial x_k}{\partial x'_i} = a_{ik},$$

and hence

$$\text{(19)} \qquad B'_i = \frac{\partial \phi}{\partial x'_i} = \sum_{k=1}^{3} \frac{\partial \phi}{\partial x_k} \frac{\partial x_k}{\partial x'_i} = \sum_{k=1}^{3} a_{ik} B_k;$$

the B_i transform like the components of a vector and therefore the gradient of ϕ,

$$\text{(20)} \qquad \nabla \phi = \sum_{k=1}^{3} \frac{\partial \phi}{\partial x_k} \mathbf{i}_k,$$

calculated at a point P, is a fixed vector associated with that point.

Let A_i be a rectangular component of a vector $\mathbf{A}$, and

$$\text{(21)} \qquad B_i = \frac{\partial A_i}{\partial x_i}.$$

Then by (18) and (15),

$$\text{(22)} \qquad B'_i = \frac{\partial A'_i}{\partial x'_i} = \sum_{k=1}^{3} a_{ik} \frac{\partial A'_i}{\partial x_k} = \sum_{k=1}^{3} \sum_{j=1}^{3} a_{ik} a_{ij} \frac{\partial A_j}{\partial x_k},$$

whence, from (10), it follows that

$$\text{(23)} \qquad \sum_{i=1}^{3} B'_i = \sum_{k=1}^{3} \sum_{j=1}^{3} \Big(\sum_{i=1}^{3} a_{ik} a_{ij} \Big) \frac{\partial A_j}{\partial x_k} = \sum_{j=1}^{3} B_j;$$

the divergence of a vector is invariant to an orthogonal transformation of the coordinate system.

Lastly, since the gradient of an invariant scalar is a vector and since the divergence of a vector is invariant, it follows that *the Laplacian*

$$\text{(24)} \qquad \nabla^2 \phi = \nabla \cdot \nabla \phi = \sum_{i=1}^{3} \frac{\partial^2 \phi}{\partial x_i^2}$$

is invariant to an orthogonal transformation.

The transformation properties of vectors may be extended to manifolds of more than three dimensions. Let x_1, x_2, x_3, x_4 be the rectangular coordinates of a point P with respect to a reference system X in a four-dimensional continuum. The location of P with respect to a fixed origin O is determined by the vector

$$\text{(25)} \qquad \mathbf{r} = \sum_{j=1}^{4} x_j \mathbf{i}_j, \qquad \mathbf{i}_j \cdot \mathbf{i}_k = \delta_{jk}.$$

The linear transformation

$$\text{(26)} \qquad x'_j = \sum_{k=1}^{4} a_{jk} x_k, \qquad (j = 1, 2, 3, 4),$$

will be called orthogonal if the coefficients satisfy the conditions

$$\text{(27)} \qquad \sum_{j=1}^{4} a_{ji} a_{jk} = \delta_{ik}.$$

The characteristic property of an orthogonal transformation is that it leaves the sum of the squares of the coordinates invariant:

$$\text{(28)} \qquad \sum_{j=1}^{4} (x_j)^2 = \sum_{j=1}^{4} (x'_j)^2.$$

The square of the determinant formed from the a_{jk} is readily shown to be positive and equal to unity, and hence the determinant itself may equal ± 1. However, if (26) is to include the identical transformation

$$\text{(29)} \qquad x'_j = x_j, \qquad (j = 1, 2, 3, 4),$$

it is obvious that the determinant must be positive. Henceforth we shall confine ourselves to the subgroup of orthogonal transformations characterized by (27) and the condition

$$|a_{jk}| = +1. \tag{30}$$

The transformation then corresponds geometrically to a rotation of the coordinate axes.

A *four-vector* is now defined as any set of four variant scalars A_i $(i = 1, 2, 3, 4)$ which transform with a rotation of the coordinate system like the coordinates of a point.

$$A'_j = \sum_{k=1}^{4} a_{jk}A_k, \qquad A_k = \sum_{j=1}^{4} a_{jk}A'_j. \tag{31}$$

It is then easy to show, as above, that the scalar product of two four-vectors and the four-dimensional divergence of a four-vector are invariant to a rotation of the coordinate system.

$$\mathbf{A} \cdot \mathbf{B} = \sum_{k=1}^{4} A_k B_k = \sum_{j=1}^{4} A'_j B'_j, \tag{32}$$

$$\Box \cdot \mathbf{A} = \sum_{i=1}^{4} \frac{\partial A_i}{\partial x_i} = \sum_{j=1}^{4} \frac{\partial A'_j}{\partial x'_j}. \tag{33}$$

Furthermore the derivatives of a scalar,

$$B_i = \frac{\partial \phi}{\partial x_i} \qquad (i = 1, 2, 3, 4) \tag{34}$$

transform like the components of a four-vector and hence the four-dimensional Laplacian of an invariant scalar,

$$\Box^2 \phi = \sum_{i=1}^{4} \frac{\partial^2 \phi}{\partial x_i^2} = \sum_{j=1}^{4} \frac{\partial^2 \phi}{\partial x_j'^2}, \tag{35}$$

is also invariant to an orthogonal transformation.

1.20. Elements of Tensor Analysis.—Although most physical quantities may be classified either as scalars, having only magnitude, or as vectors, characterized by magnitude and direction, there are certain entities which cannot be properly represented by either of these terms. The displacement of the center of gravity of a metal rod, for example, may be defined by a vector; but the rod may also be stretched along the axis by application of a tension at the two ends without displacing the center at all. The quantity employed to represent this stretching must thus indicate a *double* direction. The inadequacy of the vector concept becomes all the more apparent when one attempts the description of a volume deformation, taking into account the lateral contraction of the rod. In the present section we shall deal only with the simpler aspects of tensor calculus, which is the appropriate tool for the treatment of such problems.

In a three-dimensional continuum let each rectangular component of a vector **B** be a linear function of the components of a vector **A.**

$$\begin{aligned} B_1 &= T_{11}A_1 + T_{12}A_2 + T_{13}A_3, \\ B_2 &= T_{21}A_1 + T_{22}A_2 + T_{23}A_3, \\ B_3 &= T_{31}A_1 + T_{32}A_2 + T_{33}A_3. \end{aligned} \tag{36}$$

In order that this association of the components of **B** with the components of **A** in the system X be preserved as the coordinates are rotated, it is necessary that the coefficients T_{jk} transform in a specific manner. The T_{jk} are therefore *variant scalars*. A tensor—or more properly, a tensor of rank two—will now be defined as a linear transformation of the components of a vector **A** into the components of a vector **B** which is invariant to rotations of the coordinate system. The nine coefficients T_{jk} of the linear transformation are called the tensor components.

To determine the manner in which a tensor component must transform we write first (36) in the abbreviated form

$$B_j = \sum_{k=1}^{3} T_{jk}A_k, \qquad (j = 1, 2, 3). \tag{37}$$

If (37) is to be invariant to the transformation defined by

$$x'_j = \sum_{k=1}^{3} a_{jk}x_k, \qquad \sum_{j=1}^{3} a_{ji}a_{jk} = \delta_{ik}, \tag{38}$$

then the T_{jk} must transform to T'_{il} such that

$$B'_i = \sum_{l=1}^{3} T'_{il}A'_l, \qquad (i = 1, 2, 3). \tag{39}$$

Multiply (37) by a_{ij} and sum over the index j.

$$\sum_{j=1}^{3} a_{ij}B_j = \sum_{j=1}^{3}\sum_{k=1}^{3} a_{ij}T_{jk}A_k. \tag{40}$$

But

$$B'_i = \sum_{=1}^{3} a_{ij}B_j, \qquad A_k = \sum_{l=1}^{3} a_{lk}A'_l, \tag{41}$$

and, hence,

$$B'_i = \sum_{l=1}^{3} \left(\sum_{j=1}^{3}\sum_{k=1}^{3} a_{ij}a_{lk}T_{jk} \right) A'_l = \sum_{l=1}^{3} T'_{il}A'_l. \tag{42}$$

The components of a tensor of rank two transform according to the law

$$T'_{il} = \sum_{j=1}^{3}\sum_{k=1}^{3} a_{ij}a_{lk}T_{jk}, \qquad (i, l = 1, 2, 3); \tag{43}$$

inversely, *any set of nine quantities which transform according to* (43) *constitutes a tensor*.

By an analogous procedure one can show that the reciprocal transformation is

$$T_{jk} = \sum_{i=1}^{3}\sum_{l=1}^{3} a_{ij}a_{lk}T'_{il}. \tag{44}$$

If the order of the indices in all the components of a tensor may be changed with no resulting change in the tensor itself, so that $T_{jk} = T_{kj}$, the tensor is said to be completely symmetric. A tensor is completely antisymmetric if an interchange of the indices in each component results in a change in sign of the tensor. The diagonal terms T_{jj} of an antisymmetric tensor evidently vanish, while for the off-diagonal terms, $T_{jk} = -T_{kj}$. It is clear from (43) that if $T_{jk} = T_{kj}$, then also $T'_{il} = T'_{li}$. Likewise if $T_{jk} = -T_{kj}$, it follows that $T'_{il} = -T'_{li}$. The symmetric or antisymmetric character of a tensor is invariant to a rotation of the coordinate system.

The sum or difference of two tensors is constructed from the sums or differences of their corresponding components. If ${}^2\mathbf{R}$ is the sum of the tensors ${}^2\mathbf{S}$ and ${}^2\mathbf{T}$,[1] its components are by definition

$$R_{jk} = S_{jk} + T_{jk}, \qquad (j, k = 1, 2, 3). \tag{45}$$

In virtue of the linear character of (43) the quantities R_{jk} transform like the S_{jk} and T_{jk} and, therefore, constitute the components of a tensor ${}^2\mathbf{R}$. From this rule it follows that any asymmetric tensor may be represented as the sum of a symmetric and an antisymmetric tensor. Assuming ${}^2\mathbf{R}$ to be the given asymmetric tensor, we construct a symmetric tensor ${}^2\mathbf{S}$ from the components

$$S_{jk} = \tfrac{1}{2}(R_{jk} + R_{kj}) = S_{kj} \tag{46}$$

and an antisymmetric tensor ${}^2\mathbf{T}$ from the components

$$T_{jk} = \tfrac{1}{2}(R_{jk} - R_{kj}) = -T_{kj}. \tag{47}$$

Then by (45) the sum of ${}^2\mathbf{S}$ and ${}^2\mathbf{T}$ so constructed is equal to ${}^2\mathbf{R}$.

In a three-dimensional manifold an antisymmetric tensor reduces to three independent components and in this sense resembles a vector. The

[1] Tensors of second rank will be indicated by a superscript as shown.

tensor (36), for example, reduces in this case to

$$\begin{aligned} B_1 &= 0 - T_{21}A_2 + T_{13}A_3, \\ B_2 &= T_{21}A_1 + 0 - T_{32}A_3, \\ B_3 &= -T_{13}A_1 + T_{32}A_2 + 0. \end{aligned} \tag{48}$$

These, however, are the components of a vector,

$$\mathbf{B} = \mathbf{T} \times \mathbf{A}, \tag{49}$$

wherein the vector $\mathbf{T}$ has the components

$$T_1 = T_{32}, \qquad T_2 = T_{13}, \qquad T_3 = T_{21}. \tag{50}$$

Now it will be recalled that in vector analysis it is customary to distinguish *polar vectors*, such as are employed to represent translations and mechanical forces, from *axial vectors* with which there are associated directions of rotation. Geometrically, a polar vector is represented by a displacement or line, whereas an axial vector corresponds to an area. A typical axial vector is that which results from the vector or cross product of two polar vectors, and we must conclude from the above that an axial vector is in fact an antisymmetric tensor and its components should properly be denoted by two indices rather than one. Thus for the components of $\mathbf{T} = \mathbf{A} \times \mathbf{B}$ we write

$$T_{jk} = A_jB_k - A_kB_j = -T_{kj}, \qquad (j, k = 1, 2, 3). \tag{51}$$

If the coordinate system is rotated, the components of $\mathbf{A}$ and $\mathbf{B}$ are transformed according to

$$A_j = \sum_{l=1}^{3} a_{lj}A'_l, \qquad B_k = \sum_{i=1}^{3} a_{ik}B'_i. \tag{52}$$

Upon introducing these values into (51) we find

$$A_jB_k - A_kB_j = \sum_{l=1}^{3}\sum_{i=1}^{3} a_{lj}a_{ik}(A'_lB'_i - A'_iB'_l), \tag{53}$$

a relation which is identical with (44) and which demonstrates that the components of a "vector product" of two vectors transform like the components of a tensor. The essential differences in the properties of polar vectors and the properties of those axial vectors by means of which one represents angular velocities, moments, and the like, are now clear: axial vectors are vectors only in their manner of composition, not in their law of transformation. It is important to add that an antisymmetric tensor can be represented by an axial or pseudo-vector only in a three-dimensional space, and then only in rectangular components.

Since the cross product of two vectors is in fact an antisymmetric tensor, one should anticipate that the same is true of the curl of a vector. That the quantity $\partial A_j/\partial x_k$, where A_j is a component of a vector $\mathbf{A}$, is the component of a tensor is at once evident from Eq. (22).

$$\frac{\partial A'_i}{\partial x'_l} = \sum_{j=1}^{3}\sum_{k=1}^{3} a_{ij}a_{lk}\frac{\partial A_j}{\partial x_k}. \tag{54}$$

The components of $\nabla \times \mathbf{A}$,

$$T_{jk} = \frac{\partial A_k}{\partial x_j} - \frac{\partial A_j}{\partial x_k} = -T_{kj}, \qquad (j, k = 1, 2, 3), \tag{55}$$

therefore, transform like the components of an antisymmetric tensor.

The divergence of a tensor is defined as the operation

$$(\operatorname{div} {}^2\mathbf{T})_j = \sum_{k=1}^{3}\frac{\partial T_{jk}}{\partial x_k} = B_j, \qquad (j = 1, 2, 3). \tag{56}$$

The quantities B_j are easily shown to transform like the components of a vector.

$$\frac{\partial T'_{il}}{\partial x'_l} = \sum_{k=1}^{3} a_{lk}\frac{\partial T'_{il}}{\partial x_k} = \sum_{k=1}^{3}\sum_{j=1}^{3} a_{lk}a_{ij}a_{lk}\frac{\partial T_{jk}}{\partial x_k}, \tag{57}$$

or, on summing over l and applying the conditions of orthogonality,

$$B'_i = \sum_{l=1}^{3}\frac{\partial T'_{il}}{\partial x'_l} = \sum_{j=1}^{3} a_{ij}\left(\sum_{k=1}^{3}\frac{\partial T_{jk}}{\partial x_k}\right) = \sum_{j=1}^{3} a_{ij}B_j. \tag{58}$$

The divergence of a tensor of second rank is a vector, or tensor of first rank. The divergence of a vector is an invariant scalar, or tensor of zero rank. These are examples of a process known in tensor analysis as contraction.

As in the case of vectors, the tensor concept may be extended to manifolds of four dimensions. Any set of 16 quantities which transform according to the law,

$$T'_{il} = \sum_{j=1}^{4}\sum_{k=1}^{4} a_{ij}a_{lk}T_{jk}, \qquad (i, l = 1, 2, 3, 4), \tag{59}$$

or its reciprocal,

$$T_{jk} = \sum_{i=1}^{4}\sum_{l=1}^{4} a_{ij}a_{lk}T'_{il}, \qquad (j, k = 1, 2, 3, 4), \tag{60}$$

will be called a tensor of second rank in a four-dimensional manifold. As in the three-dimensional case, the tensor is said to be completely symmetric if $T_{jk} = T_{kj}$, and completely antisymmetric if $T_{jk} = -T_{kj}$, with $T_{jj} = 0$. In virtue of its definition it is evident that an antisymmetric four-tensor contains only six independent components. Upon expanding (59) and replacing T_{kj} by $-T_{jk}$, then re-collecting terms, we obtain as the transformation formula of an antisymmetric tensor the relation

$$T'_{il} = \sum_{j=1}^{4}\sum_{k=1}^{4}(a_{ij}a_{lk} - a_{ik}a_{lj})T_{jk} = \sum_{j=1}^{4}\sum_{k=1}^{4}\begin{vmatrix} a_{ij} & a_{ik} \\ a_{lj} & a_{lk}\end{vmatrix} T_{jk}, \quad (k > j). \tag{61}$$

Any six quantities that transform according to this rule constitute an antisymmetric four-tensor or, as it is frequently called, a *six-vector*.

In three-space the vector product is represented geometrically by the area of a parallelogram whose sides are defined by two vectors drawn from a common origin. The components of this product are then the projections of the area on the three coordinate planes. By analogy, the vector product in four-space is *defined* as the "area" of a parallelogram formed by two four-vectors, A and B, drawn from a common origin. The components of this extended product are now the projections of the parallelogram on the *six* coordinate planes, whose areas are

$$T_{jk} = A_jB_k - A_kB_j = -T_{kj}, \qquad (j, k = 1, 2, 3, 4). \tag{62}$$

The vector product of two four-vectors is therefore an antisymmetric four-tensor, or six-vector.

If again $\mathbf{A}$ is a four-vector, the quantities

$$T_{jk} = \frac{\partial A_k}{\partial x_j} - \frac{\partial A_j}{\partial x_k} = -T_{kj}, \qquad (j, k = 1, 2, 3, 4), \tag{63}$$

can be shown as in (54) to transform like the components of an antisymmetric tensor. The T_{jk} may be interpreted as the components of the curl of a four-vector.

As in the three-dimensional case the divergence of a four-tensor is defined by

$$(\operatorname{div} {}^2\mathbf{T})_j = \sum_{k=1}^{4}\frac{\partial T_{jk}}{\partial x_k}, \qquad (j = 1, 2, 3, 4), \tag{64}$$

a set of quantities which are evidently the components of a four-vector.

1.21. The Space-time Symmetry of the Field Equations.—A remarkable symmetry of form is apparent in the equations of the electromagnetic field when one introduces as independent variables the four lengths

$$x_1 = x, \qquad x_2 = y, \qquad x_3 = z, \qquad x_4 = ict, \tag{65}$$

where c is the velocity of light in free space. When expanded in rectangular coordinates, the equations

$$\text{(II)}\quad \nabla\times\mathbf{H}-\frac{\partial\mathbf{D}}{\partial t}=\mathbf{J},\qquad \text{(IV)}\quad \nabla\cdot\mathbf{D}=\rho,$$

are represented by the system

$$\begin{aligned}
0+\frac{\partial H_3}{\partial x_2}-\frac{\partial H_2}{\partial x_3}-ic\frac{\partial D_1}{\partial x_4}&=J_1,\\
-\frac{\partial H_3}{\partial x_1}+0+\frac{\partial H_1}{\partial x_3}-ic\frac{\partial D_2}{\partial x_4}&=J_2,\\
\frac{\partial H_2}{\partial x_1}-\frac{\partial H_1}{\partial x_2}+0-ic\frac{\partial D_3}{\partial x_4}&=J_3,\\
ic\frac{\partial D_1}{\partial x_1}+ic\frac{\partial D_2}{\partial x_2}+ic\frac{\partial D_3}{\partial x_3}+0&=ic\rho.
\end{aligned}\tag{66}$$

We shall treat the right-hand members of this system as the components of a "four-current" density,

$$J_1=J_x,\qquad J_2=J_y,\qquad J_3=J_z,\qquad J_4=ic\rho,\tag{67}$$

and introduce in the left-hand members a set of dependent variables defined by

$$\begin{array}{llll}
G_{11}=0 & G_{12}=H_3 & G_{13}=-H_2 & G_{14}=-icD_1\\
G_{21}=-H_3 & G_{22}=0 & G_{23}=H_1 & G_{24}=-icD_2\\
G_{31}=H_2 & G_{32}=-H_1 & G_{33}=0 & G_{34}=-icD_3\\
G_{41}=icD_1 & G_{42}=icD_2 & G_{43}=icD_3 & G_{44}=0.
\end{array}\tag{68}$$

Then in the reference system X, Eqs. (II) and (IV) reduce to

$$\sum_{k=1}^{4}\frac{\partial G_{jk}}{\partial x_k}=J_j,\qquad (j=1,2,3,4).\tag{69}$$

Only six of the G_{jk} are independent, and the resemblance of this set of quantities to the components of an antisymmetric four-tensor is obvious. Since the divergence of a four-tensor is a four-vector, it follows from (69) that *if* the G_{jk} constitute a tensor, then the J_k are the components of a four-vector; inversely, if we can show that $\mathbf{J}$ is indeed a four-vector, we may then infer the tensor character of ${}^2\mathbf{G}$. However, we have as yet offered no evidence to justify such an assumption. In the preceding sections it was shown that the vector or tensor properties of sets of scalar quantities are determined by the manner in which they transform on passing from one reference system to another. Evidently an orthogonal transformation of the coordinates x_k corresponds to a simultaneous change in both the space coordinates x, y, z *and* the time t, and only recourse to experiment will tell us how the field intensities may

be expected to transform under such circumstances. In Sec. 1.22 we shall set forth briefly the experimental facts which lead one to conclude that the J_k *are* components of a four-vector, and the G_{jk} the components of a field tensor; in the interim we regard (69) and the deductions that follow below merely as concise and symmetrical expressions of the field equations in a fixed system of coordinates.

The two homogeneous equations

$$\text{(I)}\quad \nabla\times\mathbf{E}+\frac{\partial\mathbf{B}}{\partial t}=0,\qquad \text{(III)}\quad \nabla\cdot\mathbf{B}=0,$$

are represented by the system

$$\begin{aligned}
0+\frac{\partial E_3}{\partial x_2}-\frac{\partial E_2}{\partial x_3}+ic\frac{\partial B_1}{\partial x_4}&=0,\\
-\frac{\partial E_3}{\partial x_1}+0+\frac{\partial E_1}{\partial x_3}+ic\frac{\partial B_2}{\partial x_4}&=0,\\
\frac{\partial E_2}{\partial x_1}-\frac{\partial E_1}{\partial x_2}+0+ic\frac{\partial B_3}{\partial x_4}&=0,\\
-\frac{\partial B_1}{\partial x_1}-\frac{\partial B_2}{\partial x_2}-\frac{\partial B_3}{\partial x_3}+0&=0.
\end{aligned}\tag{70}$$

After division of the first three of these equations by ic, an antisymmetrical array of components is defined as follows:

$$\begin{array}{llll}
F_{11}=0 & F_{12}=B_3 & F_{13}=-B_2 & F_{14}=-\frac{i}{c}E_1\\
F_{21}=-B_3 & F_{22}=0 & F_{23}=B_1 & F_{24}=-\frac{i}{c}E_2\\
F_{31}=B_2 & F_{32}=-B_1 & F_{33}=0 & F_{34}=-\frac{i}{c}E_3\\
F_{41}=\frac{i}{c}E_1 & F_{42}=\frac{i}{c}E_2 & F_{43}=\frac{i}{c}E_3 & F_{44}=0.
\end{array}\tag{71}$$

Then all the equations of (70) are contained in the system

$$\frac{\partial F_{ij}}{\partial x_k}+\frac{\partial F_{ki}}{\partial x_j}+\frac{\partial F_{jk}}{\partial x_i}=0,\tag{72}$$

where i, j, k are any three of the four numbers 1, 2, 3, 4.

The arrays (68) and (71) are congruent in the sense that in each the real components pertain to the magnetic field, while the imaginary components are associated with the electric field. To indicate this partition it is convenient to represent the sets of components by the symbols

$${}^2\mathbf{F}=\left(\mathbf{B},\,-\frac{i}{c}\mathbf{E}\right),\qquad {}^2\mathbf{G}=(\mathbf{H},\,-ic\mathbf{D}).\tag{73}$$

Now the field equations may be defined equally well in terms of the "dual" systems:

$$(74)\qquad {}^2\mathbf{F}^* = \left(-\frac{i}{c}\mathbf{E}, \mathbf{B}\right), \qquad {}^2\mathbf{G}^* = (-ic\mathbf{D}, \mathbf{H}),$$

or

$$(75)\qquad \begin{array}{llll} F^*_{11} = 0 & F^*_{12} = -\frac{i}{c}E_3 & F^*_{13} = \frac{i}{c}E_2 & F^*_{14} = B_1 \\ F^*_{21} = \frac{i}{c}E_3 & F^*_{22} = 0 & F^*_{23} = -\frac{i}{c}E_1 & F^*_{24} = B_2 \\ F^*_{31} = -\frac{i}{c}E_2 & F^*_{32} = \frac{i}{c}E_1 & F^*_{33} = 0 & F^*_{34} = B_3 \\ F^*_{41} = -B_1 & F^*_{42} = -B_2 & F^*_{43} = -B_3 & F^*_{44} = 0, \end{array}$$

and a corresponding system for the components of ${}^2\mathbf{G}^*$. Upon introducing these values into (I) and (III), and (II) and (IV), respectively, we obtain

$$(76)\qquad \sum_{k=1}^{4}\frac{\partial F^*_{jk}}{\partial x_k} = 0, \qquad (j = 1, 2, 3, 4),$$

$$(77)\qquad \frac{\partial G^*_{ij}}{\partial x_k} + \frac{\partial G^*_{ki}}{\partial x_j} + \frac{\partial G^*_{jk}}{\partial x_i} = J_l, \quad (i, j, k, l = 1, 2, 3, 4).$$

It has been pointed out by various writers that this last representation is artificial, in that (74) implies that **E** is an axial vector in three-space and **B** a polar vector, whereas the contrary is known to be true. The representation

$$(72)\qquad \frac{\partial F_{ij}}{\partial x_k} + \frac{\partial F_{ki}}{\partial x_j} + \frac{\partial F_{jk}}{\partial x_i} = 0, \qquad (i, j, k = 1, 2, 3, 4),$$

$$(69)\qquad \sum_{k=1}^{4}\frac{\partial G_{jk}}{\partial x_k} = J_j, \qquad (j = 1, 2, 3, 4),$$

must in this sense be considered the "natural" form of the field equations. To these we add the equation of continuity,

$$(\mathrm{V})\qquad \nabla\cdot\mathbf{J} + \frac{\partial\rho}{\partial t} = 0,$$

which in four-dimensional notation becomes

$$(78)\qquad \sum_{k=1}^{4}\frac{\partial J_k}{\partial x_k} = 0.$$

If the components F_{jk} are defined in terms of the components of a "four-potential" $\mathbf{\Phi}$ by

$$(79)\qquad F_{jk} = \frac{\partial\Phi_k}{\partial x_j} - \frac{\partial\Phi_j}{\partial x_k}, \qquad (j, k = 1, 2, 3, 4),$$

one may readily verify that Eq. (72) is satisfied identically. Now in three-space the vectors **E** and **B** are derived from a vector and scalar potential.

$$(80)\qquad E = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}, \qquad \mathbf{B} = \nabla\times\mathbf{A};$$

or, in component form,

$$(81)\qquad -\frac{i}{c}E_j = \frac{\partial}{\partial x_j}\left(\frac{i}{c}\phi\right) - \frac{\partial A_j}{\partial x_4}, \qquad B_i = \frac{\partial A_k}{\partial x_j} - \frac{\partial A_j}{\partial x_k},$$

where the indices i, j, k are to be taken in cyclical order. Clearly all these equations are comprised in the system (79) if we define the components of the four-potential by

$$(82)\qquad \Phi_1 = A_x, \qquad \Phi_2 = A_y, \qquad \Phi_3 = A_z, \qquad \Phi_4 = \frac{i}{c}\phi.$$

As in three dimensions, the four-potential is useful only if we can determine from it the field ${}^2\mathbf{G}(\mathbf{H}, -ic\mathbf{D})$ as well as the field ${}^2\mathbf{F}(\mathbf{B}, -\frac{i}{c}\mathbf{E})$. Some supplementary condition must, therefore, be imposed upon $\mathbf{\Phi}$ in order that it satisfy (69) as well as (72); thus it is necessary that the components G_{jk} be related functionally to the F_{jk}. We shall confine the discussion here to the usual case of a *homogeneous*, *isotropic medium* and assume the relations to be *linear*. To preserve symmetry of notation it will be convenient to write the proportionality factor which characterizes the medium as γ_{jk}, so that

$$(83)\qquad G_{jk} = \gamma_{jk}F_{jk};$$

but it is clear from (68) and (71) that[1]

$$(84)\qquad \gamma_{jk} = \frac{1}{\mu} \text{ when } j, k = 1, 2, 3, \qquad \gamma_{jk} = \epsilon c^2 \text{ when } j \text{ or } k = 4.$$

These coefficients are in fact components of a symmetrical tensor, and with a view to subsequent needs the diagonal terms are given the values

$$(85)\qquad \gamma_{jj} = \frac{1}{\mu}, \ (j = 1, 2, 3), \qquad \gamma_{44} = \mu\epsilon^2 c^4.$$

Equation (69) is now to be replaced by

$$(86)\qquad \sum_{k=1}^{4}\gamma_{jk}\frac{\partial F_{jk}}{\partial x_k} = J_j, \qquad (j = 1, 2, 3, 4).$$

[1] A medium which is anisotropic in either its electrical properties or its magnetic properties may be represented as in (83) provided the coordinate system is chosen to coincide with the principal axes. This also is the case if its principal axes of electrical anisotropy coincide with those of magnetic anisotropy.

Upon introducing (79) we find that (86) is satisfied, provided Φ is a solution of

$$\sum_{k=1}^{4} \frac{\partial^2}{\partial x_k^2} (\gamma_{jk}\Phi_j) = -J_j, \qquad (j = 1, 2, 3, 4), \tag{87}$$

subject to the condition

$$\sum_{k=1}^{4} \frac{\partial}{\partial x_k} (\gamma_{jk}\Phi_k) = 0, \qquad (j = 1, 2, 3, 4). \tag{88}$$

This last relation is evidently equivalent to

$$\nabla \cdot \mathbf{A} + \mu\epsilon \frac{\partial \phi}{\partial t} = 0, \tag{89}$$

and (87) comprises the two equations

$$\nabla^2 A_j - \mu\epsilon \frac{\partial^2 A_j}{\partial t^2} = -\mu J_j, \qquad (j = 1, 2, 3),$$

$$\nabla^2 \phi - \mu\epsilon \frac{\partial^2 \phi}{\partial t^2} = -\frac{1}{\epsilon}\rho. \tag{90}$$

In free space $\mu_0\epsilon_0 = c^{-2}$, $\gamma_{jk} = \mu_0^{-1}$, for all values of the indices. Equations (87) and (88) then reduce to the simple form:

$$\sum_{k=1}^{4} \frac{\partial^2 \Phi_j}{\partial x_k^2} = -\mu_0 J_j, \qquad \sum_{k=1}^{4} \frac{\partial \Phi_k}{\partial x_k} = 0, \quad (j = 1, 2, 3, 4). \tag{91}$$

1.22. The Lorentz Transformation.—The physical significance of these results is of vastly greater importance than their purely formal elegance. A series of experiments, the most decisive being the celebrated investigation of Michelson and Morley,[1] have led to the establishment of two fundamental postulates as highly probable, if not absolutely certain. According to the first of these, called the *relativity postulate*, it is impossible to detect by means of physical measurements made within a reference system X a uniform translation relative to a second system X'. That the earth is moving in an orbit about the sun we know from observations on distant stars; but if the earth were enveloped in clouds, no measurement on its surface would disclose a uniform translational motion in space. The course of natural phenomena must therefore be unaffected by a nonaccelerated motion of the coordinate systems to which they are referred, and all reference systems moving linearly and uniformly relative to each other are equivalent. For our present needs we shall state the relativity postulate as follows: *When properly formulated, the laws of*

[1] MICHELSON and MORLEY, *Am. J. Sci.*, **3**, 34, 1887.

physics are invariant to a transformation from one reference system to another moving with a linear, uniform relative velocity. A direct consequence of this postulate is that the components of all vectors or tensors entering into an equation must transform in the same way, or *covariantly*. The existence of such a principle restricted to uniform translations was established for classical mechanics by Newton, but we are indebted to Einstein for its extension to electrodynamics.

The second postulate of Einstein is more remarkable: *The velocity of propagation of an electromagnetic disturbance in free space is a universal constant c which is independent of the reference system.* This proposition is evidently quite contrary to our experience with mechanical or acoustical waves in a material medium, where the velocity is known to depend on the relative motion of medium and observer. Many attempts have been made to interpret the experimental evidence without recourse to this radical assumption, the most noteworthy being the electrodynamic theory of Ritz.[1] The results of all these labors indicate that although a constant velocity of light is *not* necessary to account for the negative results of the Michelson-Morley experiment, this postulate alone is consistent with that experiment *and* other optical phenomena.[2]

Let us suppose, then, that a source of light is fixed at the origin O of a system of coordinates $X(x, y, z)$. At the instant $t = 0$, a spherical wave is emitted from this source. An observer located at the point x, y, z in X will first note the passage of the wave at the instant ct, and the equation of a point on the wave front is therefore

$$x^2 + y^2 + z^2 - c^2t^2 = 0. \tag{92}$$

The observer, however, is free to measure position and time with respect to a second reference frame $X'(x', y', z')$ which is moving along a straight line with a uniform velocity relative to O. For simplicity we shall assume the origin O' to coincide with O at the instant $t = 0$. According to the second postulate the light wave is propagated in X' with the same velocity as in X, and the equation of the wave front in X' is

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0. \tag{93}$$

By t' we must understand the time as measured by an observer in X' with instruments located in that system. Here, then, is the key to the transformation that connects the coordinates x, y, z, t of an observation or event in X with the coordinates x', y', z', t' of the same event in X': it must be linear and must leave the quadratic form (92) invariant. The linearity follows from the requirement that a uniform, linear motion of a particle in X should also be linear in X'.

[1] RITZ, *Ann. chim. phys.*, **13**, 145, 1908.
[2] An account of these investigations will be found in Pauli's article, *loc. cit.*, p. 549.

Let

(94) $x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ict,$

be the components of a vector $\mathbf{R}$ in a four-dimensional manifold $X(x_1, x_2, x_3, x_4)$.

(95) $$R^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

The postulate on the constancy of the velocity c will be satisfied by the group of transformations which leaves this length invariant. But in Sec. 1.19 it was shown that (95) is invariant to the group of rotations in four-space and we conclude, therefore, that the transformations which take one from the coordinates of an event in X to the coordinates of that event in X' are of the form

(26) $$x_j' = \sum_{k=1}^{4} a_{jk}x_k, \qquad (j = 1, 2, 3, 4),$$

where

(27) $$\sum_{j=1}^{4} a_{ji}a_{jk} = \delta_{ik}, \qquad (i, k = 1, 2, 3, 4),$$

the determinant $|a_{jk}|$ being equal to unity.

We have now to find these coefficients. The calculation will be simplified if we assume that the rotation involves only the axes x_3 and x_4, and the resultant lack of generality is inconsequential. We take, therefore, $x_1' = x_1$, $x_2' = x_2$, and write down the coefficient matrix as follows:

(96)

	x_1	x_2	x_3	x_4
x_1'	1	0	0	0
x_2'	0	1	0	0
x_3'	0	0	a_{33}	a_{34}
x_4'	0	0	a_{43}	a_{44}

The conditions of orthogonality reduce to

(97) $$a_{33}^2 + a_{43}^2 = 1, \qquad a_{34}^2 + a_{44}^2 = 1, \qquad a_{33}a_{34} + a_{43}a_{44} = 0.$$

If we put $a_{33} = \alpha$, $a_{34} = i\alpha\beta$, we find from (97) that $a_{44} = \pm\alpha$, $a_{43} = \mp i\alpha\beta$, $\alpha\sqrt{1-\beta^2} = \pm 1$. Only the upper sign is consistent with the requirement that the determinant of the coefficients be positive unity, and this in turn is the necessary condition that the group shall contain the identical transformation. In terms of the single parameter β the coefficients are

(98) $$a_{33} = a_{44} = \frac{1}{\sqrt{1-\beta^2}}, \qquad a_{34} = -a_{43} = \frac{i\beta}{\sqrt{1-\beta^2}};$$

for the transformation itself, we have

(99) $$x_1' = x_1, \qquad x_2' = x_2,$$
$$x_3' = \frac{1}{\sqrt{1-\beta^2}}(x_3 + i\beta x_4), \qquad x_4' = \frac{1}{\sqrt{1-\beta^2}}(x_4 - i\beta x_3).$$

Reverting to the original space-time manifold this is equivalent to

(100) $$x' = x, \qquad y' = y,$$
$$z' = \frac{1}{\sqrt{1-\beta^2}}(z - \beta ct), \qquad t' = \frac{1}{\sqrt{1-\beta^2}}\left(t - \frac{\beta}{c}z\right).$$

The parameter β may be determined by considering x', y', z' to be the coordinates of a fixed point in X'. The coordinates of this point with respect to X are x, y, z. Since $dz' = 0$, it follows that

(101) $$\frac{dz}{dt} = v = \beta c, \qquad \beta = \frac{v}{c},$$

and hence the rotation defined in (96) and (97) is equivalent to a translation of the system X' along the z-axis with the constant velocity v relative to the unprimed system X.

The transformation

(102) $$x' = x, \qquad y' = y,$$
$$z' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(z - vt), \qquad t' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(t - \frac{v}{c^2}z\right),$$

obtained from (100) by substitution of the value for β, or its inversion,

(103) $$z = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(z' + vt'), \qquad t = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(t' + \frac{v}{c^2}z'\right),$$

has been named for Lorentz, who was the first to show that Maxwell's equations are invariant with respect to the change of variables defined by (102), but *not* invariant under the "Galilean transformation:"

(104) $$z' = z - vt, \qquad t' = t.$$

All known electromagnetic phenomena may be properly accounted for if the position *and time* coordinates of an event in a moving system X' be related to the coordinates of that event in an arbitrarily fixed system X by a Lorentz transformation. The Galilean transformation of classical mechanics represents the limit approached by (102) when $v \ll c$, and may be interpreted as the relativity principle appropriate to a world in which electromagnetic forces are propagated with infinite velocity.

1.23. Transformation of the Field Vectors to Moving Systems.—We shall not dwell upon the manifold consequences of the Lorentz transformation; the Fitzgerald-Lorentz contraction, the modified concept of simultaneity, the variation in apparent mass, the upper limit c which is imposed upon the velocity of matter, belong properly to the theory of relativity. The application of the principles of relativity to the equations of the electromagnetic field is essential, however, to an understanding of the four-dimensional formulation of Sec. 1.21.

The Lorentz transformation has been deduced from the postulate on the constancy of the velocity of light and has been shown to be equivalent to a rotation in a space $x_1, x_2, x_3, x_4 = ict$. Now according to the relativity postulate, the laws of physics, when properly stated, must have the same form in all systems moving with a relative, uniform motion; otherwise, it would obviously be possible to detect such a motion. In Secs. 1.19 and 1.20 it was shown that the curl, divergence, and Laplacian of vectors and tensors in a four-dimensional manifold are invariant to a rotation of the coordinate system. Therefore, to ensure the invariance of the field equations under a Lorentz transformation it is only necessary to assume that the four-current $\mathbf{J}$ and the four-potential $\mathbf{\Phi}$ do indeed transform like vectors, and that the quantities ${}^2\mathbf{F}$, ${}^2\mathbf{G}$ transform like tensors. In other words, we base the vector and tensor character of these four-dimensional quantities directly on the two postulates.

The four-current $\mathbf{J}$ satisfies the equation

$$\square \cdot \mathbf{J} = \sum_{k=1}^{4} \frac{\partial J_k}{\partial x_k} = 0. \tag{78}$$

Under a rotation of the coordinate system the components transform as

$$J'_j = \sum_{k=1}^{4} a_{jk} J_k, \tag{105}$$

or, upon introducing the values for a_{jk} from (98),

$$\begin{gathered} J'_x = J_x, \qquad J'_y = J_y, \\ J'_z = \frac{1}{\sqrt{1-\beta^2}}(J_z - v\rho), \qquad \rho' = \frac{1}{\sqrt{1-\beta^2}}\left(\rho - \frac{v}{c^2}J_z\right), \end{gathered} \tag{106}$$

with its inverse transformation

$$J_z = \frac{1}{\sqrt{1-\beta^2}}(J'_z + v\rho'), \qquad \rho = \frac{1}{\sqrt{1-\beta^2}}\left(\rho' + \frac{v}{c^2}J'_z\right). \tag{107}$$

We shall assume henceforth that the reference system X' is fixed within a material body which moves with the constant velocity v relative to the system X. This latter may usually be assumed at rest with respect to the earth. If the velocity v is very much less than the speed of light, Eq. (107) is approximately equal to

$$J_z = J'_z + v\rho', \qquad \rho = \rho'. \tag{108}$$

An observer on the moving body measures a charge density ρ' and a current density J'_z; his colleague at rest in X finds the current J'_z augmented by the *convection current* $v\rho'$.

In like manner the relations between the electric and magnetic vectors defining a given field in a fixed and in a moving system are obtained directly from the rule (61) for the transformation of the components of an antisymmetric tensor. Upon substitution of the appropriate values for the coefficients a_{jk}, one obtains for the components of ${}^2\mathbf{F}$:

$$\begin{aligned} F'_{12} &= a_{11}a_{22}F_{12} = F_{12}, \\ F'_{13} &= a_{11}a_{33}F_{13} + a_{11}a_{34}F_{14} = \frac{1}{\sqrt{1-\beta^2}}(F_{13} + i\beta F_{14}), \\ F'_{14} &= a_{11}a_{43}F_{13} + a_{11}a_{44}F_{14} = \frac{1}{\sqrt{1-\beta^2}}(F_{14} - i\beta F_{13}), \\ F'_{23} &= a_{22}a_{33}F_{23} + a_{22}a_{34}F_{24} = \frac{1}{\sqrt{1-\beta^2}}(F_{23} + i\beta F_{24}), \\ F'_{24} &= a_{22}a_{43}F_{23} + a_{22}a_{44}F_{24} = \frac{1}{\sqrt{1-\beta^2}}(F_{24} - i\beta F_{23}), \\ F'_{34} &= (a_{33}a_{44} - a_{34}a_{43})F_{34} = F_{34}, \end{aligned} \tag{109}$$

and, hence,

$$\begin{aligned} B'_x &= \frac{1}{\sqrt{1-\beta^2}}\left(B_x + \frac{v}{c^2}E_y\right), & E'_x &= \frac{1}{\sqrt{1-\beta^2}}(E_x - vB_y), \\ B'_y &= \frac{1}{\sqrt{1-\beta^2}}\left(B_y - \frac{v}{c^2}E_x\right), & E'_y &= \frac{1}{\sqrt{1-\beta^2}}(E_y + vB_x), \\ B'_z &= B_z, & E'_z &= E_z. \end{aligned} \tag{110}$$

The restriction to translations along the z-axis may be discarded by writing $\mathbf{v}$ as a vector representing the translational velocity of X' (the moving body) in any direction with respect to a fixed system X. Since in (110) the orientation of the z-axis was arbitrary, we have in general

$$\begin{aligned} B'_{\parallel} &= B_{\parallel}, & E'_{\parallel} &= E_{\parallel}, \\ B'_{\perp} &= \frac{1}{\sqrt{1-\beta^2}}\left(\mathbf{B} - \frac{1}{c^2}\mathbf{v}\times\mathbf{E}\right)_{\perp}, \\ & & E'_{\perp} &= \frac{1}{\sqrt{1-\beta^2}}(\mathbf{E} + \mathbf{v}\times\mathbf{B})_{\perp}, \end{aligned} \tag{111}$$

where $\|$ denotes components parallel, $\perp$ components perpendicular to the axis of translation. Dropping terms in c^{-2}, as is justifiable whenever the body is moving with a velocity $v \ll 3 \times 10^8$ meters/second, we obtain the approximate formulas

$$(112) \qquad \begin{aligned} B'_{\|} &= B_{\|}, & E'_{\|} &= E_{\|}, \\ B'_{\perp} &= B_{\perp}, & E'_{\perp} &= E_{\perp} + (\mathbf{v} \times \mathbf{B})_{\perp}. \end{aligned}$$

The implication of these results is striking indeed: the electric and magnetic fields **E** and **B** have no independent existence as separate entities. The fundamental complex is the field tensor ${}^2\mathbf{F} = (\mathbf{B}, -\frac{i}{c}\mathbf{E})$; the resolution into electric and magnetic components is wholly relative to the motion of the observer. When at rest with respect to permanent magnets or stationary currents, one measures a purely magnetic field **B**. An observer within a moving body or system X', on the other hand, notes approximately the same magnetic field, but in addition an electrostatic field of intensity $\mathbf{E}' = \mathbf{v} \times \mathbf{B}$. Or, inversely, the moving body may carry a fixed charge. To an observer on the body, moving with the charge, the field is purely electrostatic, whereas his colleague aground finds a magnetic field in company with the electric, identifying quite rightly the moving charge with a current.

From the tensor ${}^2\mathbf{G} = (\mathbf{H}, -ic\mathbf{D})$ are calculated in like fashion the transformations of the vectors **H** and **D** from a fixed to a moving system.

$$(113) \qquad \begin{aligned} H'_{\|} &= H_{\|}, & D'_{\|} &= D_{\|}, \\ H'_{\perp} &= \frac{1}{\sqrt{1-\beta^2}}(\mathbf{H} - \mathbf{v} \times \mathbf{D})_{\perp}, & D'_{\perp} &= \frac{1}{\sqrt{1-\beta^2}} \left(\mathbf{D} + \frac{1}{c^2}\mathbf{v} \times \mathbf{H}\right)_{\perp}. \end{aligned}$$

The invariance of Maxwell's equations to uniform translations amounts to this: if the vector functions **E, B, H, D** define an electromagnetic field in a system X, the equations

$$(114) \qquad \begin{aligned} \nabla' \times \mathbf{E}' + \frac{\partial \mathbf{B}'}{\partial t'} &= 0, & \nabla' \cdot \mathbf{B}' &= 0, \\ \nabla' \times \mathbf{H}' - \frac{\partial \mathbf{D}'}{\partial t'} &= \mathbf{J}', & \nabla' \cdot \mathbf{D}' &= \rho', \end{aligned}$$

are satisfied in a system X' which moves with the constant velocity **v** relative to X, the operator ∇' implying that differentiation is to be effected with respect to the variables x', y', z'. An observer at rest in X' interprets the vectors **E′, B′, H′, D′** as the intensities of an electromagnetic field satisfying Maxwell's equations. Clearly the ratios of **D** to **E** and **H** to **B** are not preserved in both systems. The macroscopic parameters ϵ, μ, σ are also subject to transformation, which may be ascribed to an actual change in the structure of matter in motion. In practice one is interested usually in *the mechanical and electromotive forces, measured in the fixed system X, which act on moving matter*, rather than in the transformed field intensities **E′, B′, H′, D′**. The determination of these forces and of the differential equations which they satisfy within the framework of the relativity theory was accomplished by Minkowski in the course of his investigation on the electrodynamics of moving bodies.

The vector character of the four-potential is demonstrated by Eq. (79) which expresses the field tensor ${}^2\mathbf{F}$ as the curl of $\mathbf{\Phi}$. Under a Lorentz transformation

$$(115) \qquad \Phi'_j = \sum_{k=1}^{4} a_{jk}\Phi_k \qquad (j = 1, 2, 3, 4),$$

or, in terms of vector and scalar potentials,

$$(116) \qquad \begin{aligned} A'_x &= A_x, & A'_y &= A_y \\ A'_z &= \frac{1}{\sqrt{1-\beta^2}}\left(A_z - \frac{v}{c^2}\phi\right), & \phi' &= \frac{1}{\sqrt{1-\beta^2}}(\phi - vA_z). \end{aligned}$$

As in the case of the field vectors, the resolution into vector and scalar potentials in three-space is determined by the relative motion of the observer.

In conclusion it may be remarked that a rotation of the coordinate system leaves invariant the scalar product of any two vectors. It was in fact from the required invariance of the quantity

$$(117) \qquad \mathbf{R} \cdot \mathbf{R} = R^2 = x^2 + y^2 + z^2 - c^2t^2,$$

that we deduced the Lorentz transformation. Since the current density **J** and the potential $\mathbf{\Phi}$ have been shown to be four-vectors, it follows that the quantities

$$(118) \qquad \begin{aligned} J^2 &= J_x^2 + J_y^2 + J_z^2 - c^2\rho^2, \\ \Phi^2 &= A_x^2 + A_y^2 + A_z^2 - \frac{\phi^2}{c^2}, \\ \mathbf{\Phi} \cdot \mathbf{J} &= A_xJ_x + A_yJ_y + A_zJ_z - \phi\rho, \end{aligned}$$

are true scalar invariants in a space-time continuum. There are, moreover, certain other scalar invariants of fundamental importance to the general theory of the electromagnetic field. From the transformation formula Eq. (59) the reader will verify that if ${}^2\mathbf{S}$ and ${}^2\mathbf{T}$ are two tensors of second rank, the sums

$$(119) \qquad \sum_{j=1}^{4}\sum_{k=1}^{4} S_{jk}T_{jk} = \text{invariant}, \qquad \sum_{j=1}^{4}\sum_{k=1}^{4} S_{jk}T_{kj} = \text{invariant},$$

are invariant to a rotation of the coordinate axes. These quantities may be interpreted as scalar products of the two tensors. Let us form first the scalar product of ${}^2\mathbf{F}$ with itself. According to (71) and (119) we find

$$\sum_{j=1}^{4}\sum_{k=1}^{4} F_{jk}^2 = 2\left(B^2 - \frac{1}{c^2}E^2\right) = \text{invariant}. \tag{120}$$

Next we construct the scalar product of ${}^2\mathbf{F}$ with its dual ${}^2\mathbf{F}^*$ defined in (75).

$$\sum_{j=1}^{4}\sum_{k=1}^{4} F_{jk}F_{jk}^* = -4\frac{i}{c}\mathbf{B}\cdot\mathbf{E} = \text{invariant}. \tag{121}$$

From the tensor ${}^2\mathbf{G}$ and its dual ${}^2\mathbf{G}^*$ may be constructed the invariants

$$\sum_{j=1}^{4}\sum_{k=1}^{4} G_{jk}^2 = 2(H^2 - c^2D^2) = \text{invariant}, \tag{122}$$

$$\sum_{j=1}^{4}\sum_{k=1}^{4} G_{jk}G_{jk}^* = -4ic\mathbf{H}\cdot\mathbf{D} = \text{invariant}. \tag{123}$$

Proceeding in the same fashion, we obtain

$$\sum_{j=1}^{4}\sum_{k=1}^{4} F_{jk}G_{jk} = \sum_{j=1}^{4}\sum_{k=1}^{4} F_{jk}^*G_{jk}^* = 2(\mathbf{B}\cdot\mathbf{H} - \mathbf{E}\cdot\mathbf{D}) = \text{invariant}, \tag{124}$$

$$\sum_{j=1}^{4}\sum_{k=1}^{4} F_{jk}G_{jk}^* = \sum_{j=1}^{4}\sum_{k=1}^{4} F_{jk}^*G_{jk} = -2i\left(c\mathbf{B}\cdot\mathbf{D} + \frac{1}{c}\mathbf{E}\cdot\mathbf{H}\right) = \text{invariant}. \tag{125}$$

The invariance of these quantities in configuration space is trivial; they are set apart from other scalar products by the fact that they preserve the same value in every system moving with a uniform relative velocity.

CHAPTER II

STRESS AND ENERGY

To translate the mathematical structure developed in the preceding pages into experiments which can be conducted in the laboratory, we must calculate the mechanical forces exerted in the field upon elements of charge and current or upon bodies of neutral matter. In the present chapter it will be shown how by an appropriate definition of the vectors **E** and **B** these forces may be deduced directly from the Maxwell equations. In the course of this investigation we shall have to take account of the elastic properties of material media. A brief digression on the analysis of elastic stress and strain will provide an adequate basis for the treatment of the body and surface forces exerted by electric or magnetic fields.

STRESS AND STRAIN IN ELASTIC MEDIA

2.1. The Elastic Stress Tensor.—Let us suppose that a given solid or fluid body of matter is in static equilibrium under the action of a specified system of applied forces. Within this body we isolate a finite volume V by means of a closed surface S, as indicated in Fig. 12.

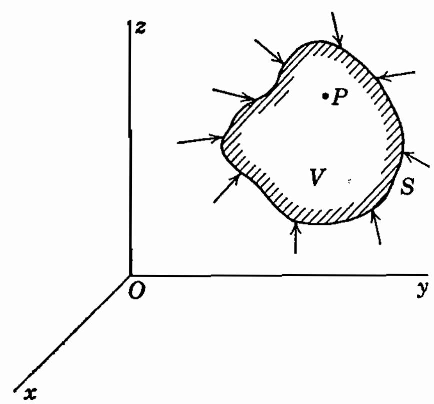

FIG. 12.—A region V bounded by a surface S in an elastic medium under stress.

Since equilibrium has been assumed for the body and all its parts, the resultant force **F** exerted on the matter within S must be zero. Contributing to this resultant are *volume* or *body forces*, such as gravity, and *surface forces* exerted by elements of matter just outside the enclosed region on contiguous elements within. Throughout V, therefore, we suppose force to be distributed with a density **f** per unit volume, while the force exerted by matter outside S on a unit area of S will be represented by the vector **t**. The components of **t** are evidently normal pressures or tensions and tangential shears. The condition of translational equilibrium is expressed by the equation

$$\int_V \mathbf{f}\,dv + \int_S \mathbf{t}\,da = 0. \tag{1}$$