

# Electromagnetics and Differential Forms

GEORGES A. DESCHAMPS, FELLOW, IEEE

*Invited Paper*

**Abstract**—Differential forms of various degrees go hand in hand with multiple integrals. They obviously constitute an essential tool in expressing the laws of physics. Some of their structures, however (exterior algebra, exterior differential operators, and others), are not widely known or used. This article concentrates on the relevance of the “exterior calculus” to electromagnetics. It is shown that the association of differential forms with electromagnetic quantities is quite natural. The basic relations between these quantities, displayed in flow diagrams, make use of a single operator “d” (exterior differential) in place of the familiar curl, grad and div operators of vector calculus. Their covariance properties (behavior under change of variables) are discussed. These formulas in space-time have a strikingly concise and elegant expression. Furthermore, they are also invariant under any change of coordinates involving both space and time. Physical dimensions of the electromagnetic forms are such that only two units (coulomb and weber, or  $e$  and  $g$ ) are needed.

A few sample applications of the exterior calculus are discussed, mostly to familiarize the reader with the aspect of equations written in this manner. The transition from differential to integral formulas is uniformly performed by means of Stokes’ theorem (concisely expressed in terms of forms). When integrations over moving domains are involved, the concepts of flow and Lie derivative come into play. The relation of the topology of a region to the existence of potentials valid in that region is illustrated by two examples: the magnetic field due to a steady electric current and the vector potential of the  $B$ -field due to a Dirac magnetic monopole.

An extensive appendix reviews most results needed in the main text.

## I. INTRODUCTION

**D**IFFERENTIAL forms are expressions on which integration operates. They obviously constitute an essential tool in expressing the laws of physics. Some of their properties, however, came to light only after the work of E. Cartan at the beginning of the century. They were applied mostly to differential geometry and they are not yet widely known or appreciated. A key property is that differential forms possess a natural algebraic structure which was first considered by H. Grassmann to deal with the calculus of “extensions” (Ausdehnung) and is now known as *exterior algebra*. Furthermore, an operation designated by “d” and called *exterior derivative* (or exterior differential) operates on differential forms to produce forms of one degree higher. The operation d replaces, and does generalize, the familiar curl, grad, and div operations of vector calculus. It obeys simple rules that are easy to memorize and leads to a most elegant formulation of Stokes’ theorem.

The applications of the calculus of differential forms (also called *exterior differential calculus* to emphasize the role of exterior algebra and exterior derivation) go far beyond differ-

ential geometry. Exterior differential forms are considered an essential tool for classical mechanics [1], [2], geometrical optics, semiclassical quantum mechanics, and ~~more recently~~ the theory of gauge fields.

The main purpose of this article is to introduce the applications of exterior differential forms to electromagnetics. This formulation is not new; it does appear in a few recent texts in mathematics and in [3]–[5] addressed to physicists, and in a short exposé by the present author [6, ch. 3]. The differential form approach has not yet had any impact on engineering in spite of its convenience, compactness, and many other qualities. The main reason for this is, of course, the lack of exposure in engineering publications: the entire literature on the subject of electromagnetics is written in vector calculus notation. It is hoped that this article will help remove this obstacle to a wider use of these techniques, and demonstrate some of the real advantages of this new notation.

To make the article self-contained, a short review of the main properties of exterior calculus is presented in the appendices. Those familiar with this topic will need only a cursory glance at this section, mostly to get acquainted with the particular notations used in the text. Others will find there all the results necessary to follow the main text. References should be consulted for proofs and complementary information.

The main body of this article is organized as follows. First, the representation of electromagnetic quantities by differential forms (Section I) is introduced by relying on the familiarity of most readers with the notation of multiple integrals: a differential form is the complete integrand appearing under an integral sign. For less-known properties of these forms, such as their algebraic structure and the definition of exterior differential, as well as for a precise definition of a differential form as a field of multivectors, the appendix must be consulted.

Next in Section II the equations relating electromagnetic quantities are presented in the form of flow diagrams. The single operator “d” (exterior derivative, or differential) replaces the usual curl, grad and div operations. Its properties, discussed in Appendix H, show, in particular, the following covariance property: the form of the equations is independent of the variables used. In space-time the field quantities pair up to produce new differential forms whose relations are strikingly simpler. Those relations also enjoy the covariance property. Consideration of the physical dimensions of the differential forms and their invariance under the operator d show that only two basic units are needed: that of electric charge and that of magnetic charge. These may be coulomb and weber in SI units, or  $e$  (the electron charge) and  $g$  (twice the magnetic pole charge). In Section IV, some familiar results about reciprocity (Section IV-A), Huygens’ principle

Manuscript received July 26, 1980; revised December 17, 1980. This work was supported in part by the National Science Foundation under Grant ENG-77-20820.

The author is with the Electromagnetics Laboratory, Department of Electrical Engineering, University of Illinois, Urbana, IL 61801.

(Section IV-B) and the Kirchhoff approximation (Section IV-C) are presented in terms of differential forms to acquaint the reader with the aspect of differential form equations. The integral formulations of electromagnetics are deduced from the differential equations by means of Stokes' theorem (Section IV-D) and are generalized to moving domains of integration (Section IV-D) by using Lie derivatives (Appendix L). Finally (Section IV-E), the existence of potentials is discussed for simple examples and is shown to be related to Poincaré's lemma.

## II. REPRESENTATION OF ELECTROMAGNETIC QUANTITIES BY DIFFERENTIAL FORMS

Differential forms are expressions on which integration operates. Forms of degree  $p$ , or  $p$ -forms, are expressions that occur in  $p$ -fold integrals, i.e., integrals over domains (or chains) of dimension  $p$ . It is not surprising that these forms occur widely in physics and in particular in electromagnetics. We shall introduce forms of various degrees (1, 2, 3, and 0) by describing some electromagnetic quantities that they represent quite naturally.

Consider first the electric field, conventionally represented, in rectangular coordinates, by a vector

$$\vec{E}(r) = X\hat{x} + Y\hat{y} + Z\hat{z} \tag{1}$$

where  $\hat{x}, \hat{y}, \hat{z}$  are unit vectors, and  $(X, Y, Z)$  are functions of  $r = (x, y, z)$ . (In rectangular coordinates the unit vectors  $\hat{x}, \hat{y}, \hat{z}$  are also represented by  $\partial_x, \partial_y, \partial_z$  as described in Appendix G.) The vector  $\vec{E}$  is interpreted as the force acting on a unit electric charge, at rest at point  $r$ . It serves to compute the work done on this test charge when it is moved along a path<sup>1</sup>

$$\gamma: [a, b] \subset R \rightarrow R^3: t \mapsto r = \gamma(t). \tag{2}$$

This work is represented by the line integral

$$w = \int_{\gamma} X dx + Y dy + Z dz. \tag{3}$$

The quantity under the integral sign is precisely a differential form of degree one, a one-form, that we shall designate by  $E$  without the arrow. At some point  $r$ , the form  $E(r)$  may be considered as a linear operator which applied to a displacement vector  $\delta r$  gives the work done by  $E$  on a unit charge. Thus,  $E(r)$  is a covector, element of the dual of the space of vectors with origin at  $r$ . The work is the duality product  $E|\delta r$  (see Appendices B and G). The integral (3) may be regarded as the limit of a Riemann sum:

$$\Sigma(E_i|\delta_i\gamma) = \Sigma(X_i\delta_ix + Y_i\delta_iz + Z_i\delta_iz) \tag{4}$$

where the arc of curve  $\gamma$  is approximated by a polygon defined by the points  $(r_1, r_2, \dots, r_N)$ ,  $\delta_i\gamma = r_{i+1} - r_i$  (with components  $\delta_ix, \delta_iz, \delta_iz$ ), and  $E_i = E(r_i)$ . The limit is taken as the largest  $|\delta_i\gamma|$  approaches zero. The integral (3), called the *circulation* of  $E$  along  $\gamma$ , will be written  $E|\gamma$ . The vertical bar is unnecessary if one pays attention to the nature of the factors: (one-form|curve) = scalar.

The replacement of the vector field  $\vec{E}$  by the one-form  $E$  and writing  $E$  instead of  $E \cdot \delta r$  may seem insignificant at this point.

However, it must be noted that in the second expression one needs to give meaning to the scalar product, which requires a metric, such as the Euclidean metric implied by rectangular coordinates. In contrast, the one-form  $E$  can be written in the same manner in any system of coordinates. For example, in spherical coordinates  $(r, \theta, \phi)$

$$E = R dr + \Theta d\theta + \Phi d\phi \tag{5}$$

where  $(R, \Theta, \Phi)$  are functions of  $(r, \theta, \phi)$ . The integral  $E|\gamma$  in that case is calculated in exactly the same manner as previously, by Riemann sum, or better by pullback (see Appendix I). The vector  $\vec{E}$  can still be deduced from the form  $E$ , but not anymore by replacing  $(dr, d\theta, d\phi)$  by  $(\hat{r}, \hat{\theta}, \hat{\phi})$ . Metrical coefficients become involved, which should, in fact, be irrelevant since they occur also in forming the scalar product  $\vec{E} \cdot \delta r$  in such a manner that they cancel.

An example of a two-form and its interpretation is provided by the electric current. Conventionally represented by a current density vector

$$\vec{J} = U\hat{x} + V\hat{y} + W\hat{z} \tag{6}$$

it serves to express the current through an oriented surface  $S$  by means of the integral

$$\int_S \vec{J} \cdot \vec{n} dS \tag{7}$$

where  $\vec{n}$  is a unit vector normal to the surface  $S$ . The quantity under the two-fold integral (7) is a two-form, that can be written

$$J = U dydz + V dzdx + W dx dy. \tag{8}$$

The current through  $S$  is then  $\int_S J$ , which we shall also write  $J|S$ , as it can be interpreted as the limit of a Riemann sum

$$J|S = \lim \Sigma J_i|\delta_iS \tag{9}$$

where each term is the duality product of  $J_i$  (two-form  $J$  at point  $r_i$  on the surface) by a two-vector  $\delta_iS$  which represents an element of a polyhedral surface that approximates  $S$ . The actual computation of (7) is better carried out by means of the pullback of some parameterization of  $S$  (see Appendix I).

In rectangular coordinates,  $J$  is deduced from  $\vec{J}$  by replacing  $(\hat{x}, \hat{y}, \hat{z})$  by  $(dydz, dzdx, dx dy)$ . In curvilinear coordinates, the expression  $\vec{J}$  involves the metric (in integral (7) one needs a scalar product and the notions of unit vector and of vector normal to a surface), while  $J$  does not. This can be of advantage when evaluating  $J|S$  the *flux* of  $J$  through some surface  $S$ .

An example of a three-form is provided by the electric charge. Conventionally represented by a density  $q$ , a scalar function of position, it serves to express the charge inside a volume  $V$  by the integral

$$\int_V q dV. \tag{10}$$

The quantity integrated is a three-form

$$\rho = q dx dy dz. \tag{11}$$

The *content* of  $V$  in electric charge is  $\int_V \rho$ , denoted  $\rho|V$ . It is the limit of a Riemann sum of duality products  $\rho_i|\delta_iV$ .

Finally, scalar functions of position, such as the potential  $\phi(r)$ , are considered as zero-forms. They "integrate" over

<sup>1</sup> The notation  $f:A \rightarrow B:a \mapsto b$  means that the function  $f$  maps set  $A$  into set  $B$ , carrying element  $a$  of  $A$  into element  $b$  of  $B$ .

regions of dimension zero, e.g., a point or a set of points. At point  $A$  the integral

$$\int_A \phi = \phi|_A \tag{12}$$

is taken to be the *value* of  $\phi$  at point  $A$ . The points may be "oriented," i.e., given a sign. Then,

$$\phi|(-A) = -\phi|A = -\phi(A). \tag{13}$$

Other quantities occurring in the equation of electromagnetics,  $H, D, A, B$ , etc., can all be made into differential forms of appropriate degrees. This is done by looking at the dimension of the domain over which they are integrated: the degree of the form equals the dimension of the domain.

Conventional notations have been preserved by using the same letters for the corresponding forms. Their degree can be read from the Tables I, II, and III.

The correspondence with conventional representations can all be expressed by means of the overbar and the star operators defined in Appendices E and F. Thus,  $\vec{E} = \bar{E}, \vec{J} = *\bar{J}, q = *\rho$ .

### III. EQUATIONS OF ELECTROMAGNETICS

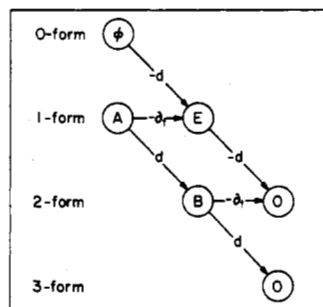
The equations that relate the electromagnetics quantities will now be presented. Their proper introduction, in a textbook manner, should start with a description of the basic experiments (Coulomb, Ampère, Faraday, etc.): leading step by step to the final result using differential forms all along. This article does not allow enough space to do this properly. Therefore, we shall state the equations without other justification than their internal consistency and their agreement with the familiar vector calculus expressions. The task of translating these equations back and forth between the two formalisms and checking their agreement will be left to the reader.

#### A. Flow Diagrams

The equations of electromagnetics are displayed in Tables I, II, and III. All quantities are represented by forms of various degrees, and they are designated by the letters conventionally used in the vector representation. The vector corresponding to a one-form is obtained by means of the overbar operator (e.g.,  $E \rightarrow \bar{E} = \vec{E}$ ), while a vector corresponding to a two-form results from the star operator composed with the overbar (e.g.,  $J \rightarrow \bar{J} = *\vec{J}$ ). Note that vectors corresponding to one-forms and two-forms are sometimes called *polar* and *axial*, respectively. This indicates different behavior under reflection which are obvious for differential forms submitted to a pull-back under this operation.

The equations decompose into two sets displayed in Table I (Maxwell-Faraday) and Table II (Maxwell-Ampère). In the upper part of these tables, and also in Table III, diagonal arrows represent the operator  $d$  (with respect to space variables) and horizontal arrows the time derivative  $\partial_t$ , or, for fields at frequency  $\omega$ , the product by  $-i\omega$  ( $e^{-i\omega t}$  convention). (The arrows for  $d$  go down in Table I, up in Table II, only to facilitate putting the two tables together in Table III.) A bar across an arrow means a negative sign. The quantity in any circle is the sum of those contributed by the arrows leading to it. Thus,  $B = \alpha A, 0 = dE + \partial_t B, E = -d\phi - dA$ , etc. Since  $d \cdot d = 0, dB = 0$  follows from  $B = dA$ . Conversely (Poincaré lemma), if  $dB = 0$  within a ball or a domain homeomorphic to

TABLE I  
MAXWELL-FARADAY EQUATIONS



SPACE-TIME FORMULATION

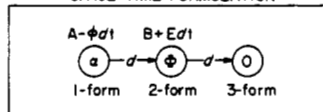
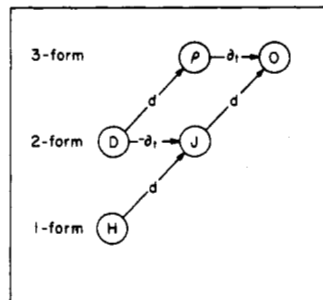
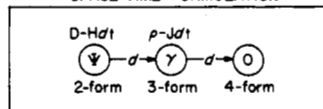


TABLE II  
MAXWELL-AMPÈRE EQUATIONS



SPACE-TIME FORMULATION



it, there exists an  $A$  in that domain such that  $dA = B$  (see Appendix K).

The equations in Tables I and II have the same expression in *any system of coordinates* (see Appendix I). This ceases to be true when the relations between the two tables are considered. With vector notations, this relation is expressed by

$$\vec{D} = \epsilon_0 \vec{E} \quad \vec{B} = \mu_0 \vec{H} \tag{14}$$

where  $\epsilon_0, \mu_0$  are the material constants, permittivity, and permeability.

In these equations,  $\vec{E}$  and  $\vec{H}$  correspond to one-forms  $E$  and  $H, \vec{D}$  and  $\vec{B}$  to two-forms  $D$  and  $B$ . The relation between the forms becomes

$$D = \epsilon_0 * E \quad B = \mu_0 * H \tag{15}$$

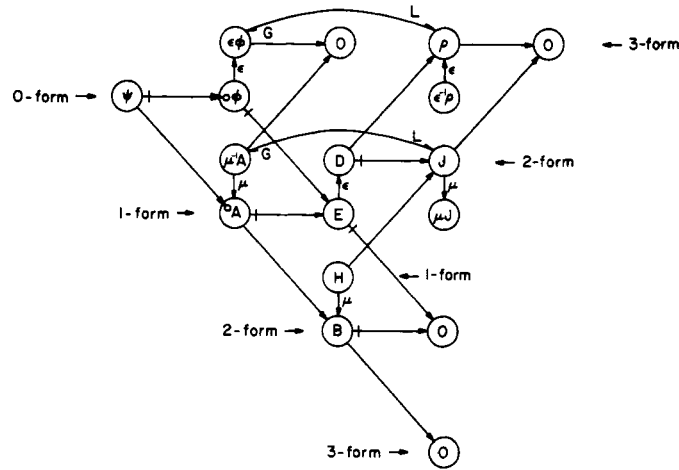
where  $*$  is the star operator in  $R^3$  defined in Appendix F, equation (F.8). We shall write

$$D = \epsilon E \quad B = \mu H \tag{16}$$

making  $\epsilon$  and  $\mu$  into operators  $\epsilon_0 *, \mu_0 *$ , that include the star.

In Table III, the vertical arrows represent  $\epsilon$  when they point upward,  $\mu$  when they point downward. They indicate a relation only between the two elements that they connect, and their

TABLE III  
ELECTROMAGNETICS FLOW DIAGRAM



contribution is not to be added to others. Thus,  $B$  equals to  $dA$  and to  $\mu H$ , but not to their sum.

On the left of the Table III, the small circles inside  $\phi$  and  $A$  represent quantities  $-\partial_t \phi$  and  $dA$  that can be added to  $\phi$  and  $A$  without changing the fields (Appendix K). This is called a *gauge transformation of the potential*. A gauge transformation exists which ensures the condition

$$\partial_t(\epsilon\phi) + d(\mu^{-1}A) = 0 \tag{17}$$

called the Lorentz gauge condition. This, in turn, implies the relations indicated by curved arrows:  $G$  and  $L$  are inverse operators. For harmonic fields  $L = -(\Delta + k^2)$  and  $G$  is convolution by  $e^{ikr}/4\pi r$ .

To any differential equation (in three or four dimensions) in Tables I, II, and III, there corresponds an integral relation that results from Stokes' theorem. For instance, Faraday's law

$$dE + \partial_t B = 0 \tag{18}$$

implies for a surface  $S$  bounded by the area  $\Gamma = \partial S$ ,

$$(dE + \partial_t B)|_S = E|\Gamma + (\partial_t B)|_S = 0. \tag{19}$$

The flow diagrams may also be used as they stand to discuss the equations in vector notation. The diagonal arrows then represent grad, curl, or div in an obvious manner.

**B. Space-Time Representation**

The lower parts of Tables I and II show how quantities in each column of the upper parts can be combined into a differential form in the four-dimensional space  $R^4$ . If  $d$  now indicates a differential with respect to space and time,

$$d = d + dt \partial_t. \tag{20}$$

The quantities

$$\begin{aligned} \alpha &= A - \phi dt \\ \Phi &= B + E dt \end{aligned} \tag{21}$$

are related by

$$\alpha \xrightarrow{d} \Phi \xrightarrow{d} 0 \text{ (Table I)} \tag{22}$$

while

$$\begin{aligned} \Psi &= D - H dt \\ \gamma &= \rho - J dt \end{aligned} \tag{23}$$

are related by

$$\Psi \xrightarrow{d} \gamma \xrightarrow{d} 0 \text{ (Table II)}. \tag{24}$$

Equations (21) and (23) illustrate  $d \cdot d = 0$  and the Poincaré lemma. Each triplet of arrows that relates one column to the next in the upper parts of Tables I and II represent the operator  $d$ . Material constants represented by operators  $\epsilon, \mu$  can be combined into a single star operator  $\star_4$  such that

$$\Psi = \star \Phi. \tag{25}$$

In vacuum this star operator is precisely the star operator characteristic of the Minkowski metric (see Appendix F, equation (F.15)).

The correspondence between three- and four-dimensional formalism is summarized in Table IV.

The simplification that results from passing to a four-dimensional formalism led Sommerfeld [7, p. 212] to exclaim, "I wish to create the impression in my readers that the true mathematical structure of these entities will appear only now, as in a mountain landscape when the fog lifts." With differential forms, this landscape is even more striking: all the equations are consequences of

$$d\Phi = 0 \quad d\star\Phi = \gamma \tag{26}$$

where  $\star$  is the star operator for the Minkowski metric.

**C. Lorentz Force Equation**

The equations discussed in Sections III-A and III-B permit the determination of fields due to known sources (currents and charges). They have to be completed by equations that tell the effect of a field on charges at rest or in motion. For a point charge  $q$ , this is given by the Lorentz force equation.

Its relativistic expression involves the field

$$\Phi = B + E dt \tag{27}$$

which for this reason could be called the *force-field*, distinguishing it from the *source-field*  $\Psi = D - H dt$ . (The fields  $\Phi$

TABLE IV  
RELATIONS BETWEEN THE THREE- AND FOUR-DIMENSIONAL FORMULATIONS  
OF THE EQUATIONS OF ELECTROMAGNETICS

Formulation in Space-Time $R^{3+1}$	Correspondence	Formulation in Space $R^3$
First set of equations: Maxwell-Faraday $\alpha \xrightarrow{d} \Phi \xrightarrow{d} 0$		
Potential One-Form $\alpha$	$\alpha = A - \phi dt$ $A = A_1 dx + A_2 dy + A_3 dz$	$(A, \phi)$
EM Force-Field $\Phi$ (two-form)	$\Phi = E dt + B$	$(E, B)$
$\Phi = d\alpha$	$\begin{cases} E = E_1 dx + E_2 dy + E_3 dz \\ B = B_1 dz dx + B_2 dz dx + B_3 dx dy \end{cases}$	$\begin{cases} E = -d\phi - \dot{A} \\ B = dA \end{cases}$
(Unit: Weber or magnetic charge $g = 137e$ )	$d\alpha = dA - (\dot{A} + d\phi) dt$	
$d\Phi = 0$	$d\Phi = (dE + \dot{B}) dt + dB$	$\begin{cases} dE + \dot{B} = 0 \\ dB = 0 \end{cases}$
Gauge Transformation $\alpha \rightarrow \alpha + df$	$df = d\dot{f} + \dot{f} dt$	$A \rightarrow A + d\dot{f}$ $\phi \rightarrow \phi - \dot{f}$
Second set of equations: Maxwell-Ampère $\Psi \xrightarrow{d} \gamma \xrightarrow{d} 0$		
EM Source-Field $\Psi$ (two-form)	$\Psi = D - H dt$	$(D, H)$
$d\Psi = \gamma$	$\begin{cases} D = D_1 dy dz + D_2 dz dx + D_3 dx dy \\ H = H_1 dx + H_2 dy + H_3 dz \end{cases}$	$\begin{cases} dD = \rho \\ dH - \dot{D} = J \end{cases}$
Charge Current $\gamma$ (three-form)	$\gamma = \rho - J dt$	$(\rho, J)$
(Unit: Coulomb or electron charge $e$ )	$\begin{cases} \rho = q dx dy dz \\ J = J_1 dy dz + J_2 dz dx + J_3 dx dy \end{cases}$	
Continuity eq. $d\gamma = 0$	$d\gamma = -(\dot{\rho} + dJ) dt$	$dJ + \dot{\rho} = 0$

The differential  $d$  applies only to space coordinates, while  $d$  applies also to time. The dot means a time derivative.

and  $\Psi$  have been called *FARADAY* and *MAXWELL* by Misner *et al.* [3]-[5].

The motion of the charge is a trajectory in  $R^4$ . The arc along this trajectory, measured with the Minkowski metric, is the *proper time*  $s$ . The velocity  $dr/ds = u$  is a unit vector tangent to the trajectory. If  $m$  is the rest mass, the momentum is the one-form  $p = m\bar{u}$  (overbar relative to Minkowski metric), and the equation of motion is

$$\partial_s p = -q\Phi|u. \quad (28)$$

The product  $\Phi|u$  is a one-form duality product (contraction) of the two-form  $\Phi$  and the one-vector  $u$  (see Appendix E). Equation (28) may be deduced from an action principle which makes the trajectory an extremal for the integral of the action one-form  $p + e\alpha$  [6, p. 51], where  $\alpha = A - \phi dt$  is the potential one-form in  $R^4$ .

#### D. Covariance of the Equations of Electromagnetics

One advantage, among others, of using differential forms is that the laws of electromagnetics, expressed by the two sets of equations, Maxwell-Ampère and Maxwell-Faraday, have the same form in any system of coordinates. They are valid without modification whether  $(x, y, z)$  represent rectangular or curvilinear coordinates.

To emphasize and illustrate this point, consider the equation

$$dA = B \quad (29)$$

which relates the potential one-form  $A$  and the magnetic field two-form  $B$  in a system of Cartesian coordinates  $(x, y, z)$ . If we express these coordinates in terms of curvilinear coordinates, say the spherical coordinates  $(r, \theta, \phi)$ , by a function

$$f: (r, \theta, \phi) \rightarrow (x, y, z) = (r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta)$$

the potential one-form  $A$  and the magnetic field  $B$  can be pulled back to

$$A' = f^*A \quad B' = f^*B. \quad (30)$$

Since the pullback  $f^*$  commutes with  $d$ , we have

$$dA' = B' \tag{31}$$

which is exactly the same equation as before the transformation. One may dispense with the prime (see Appendix I) and consider (29) valid in *any system of coordinates*. Explicitly, in spherical coordinates, if

$$A = R dr + \Theta d\theta + \Phi d\phi \tag{32}$$

where  $(R, \Theta, \Phi)$  are functions of  $(r, \theta, \phi)$ , then

$$B = dA = (\Phi_\theta - \Theta_\phi) d\theta d\phi + (R_\phi - \Phi_r) d\phi dr + (\Theta_r - R_\theta) dr d\theta \tag{33}$$

exactly as if  $(r, \theta, \phi)$  were rectangular coordinates. One must, of course, make sure that the variables used are true coordinates in the region of interest.

The space-time formulations in Tables I and II are also expressed by exterior differentials; hence, they are also independent of the coordinate system, including both space and time variables. Of course, the relation between  $\Phi$  and  $\Psi$ , since it is metric dependent, does not enjoy this property. It can be shown (see problem in Appendix I, equation (I.13)) that, in vacuum, the Lorentz transformation leaves both the Minkowski metric and the relation  $\Psi \rightarrow \Phi$  invariant. This is the basis of special relativity.

The metric independence of Maxwell's equations in three dimensions (Tables I and II) and more generally the covariance in four dimensions under any space-time transformation was recognized early by H. Weyl (1918) and by E. Cartan (1926). It was later rediscovered by D. Van Dantzig [8] (1934) and still later by the present author!

### E. Physical Dimensions of Electromagnetic Quantities

When electromagnetic quantities are represented by differential forms, an important simplification occurs in the discussion of their physical dimensions. This is due to the fact that a differential form carries with it the differentials of the variables and that consequently exterior differentiation preserves the physical dimension.

The latter property is easily verified: the exterior differential  $d\alpha$  of the  $p$ -form  $\alpha = \sum a_J dx^J$  is  $\sum da_J dx^J$ , but  $da_J = \sum a_{J,k} dx^k$ , where  $a_{J,k}$  is the partial derivation of  $a_J$  with respect to  $x^k$ . Thus,  $\dim(da_J) = \dim(a_J)$ , and  $\dim(d\alpha) = \dim \alpha$ .

In Tables I, II, and III, all quantities occurring on the same diagonal have the same physical dimension:  $(\rho, D)$  in coulombs,  $(J, H)$  in amperes,  $(\Phi, E)$  in volts, and  $(A, B)$  in webers. For the space-time descriptions:  $(\gamma, \Psi)$  are in coulombs and  $(\alpha, \Phi)$  are in webers.

It is interesting to note that both coulombs and webers can be naturally quantized. Electric charges occur in integral multiples of  $e$ , the negative of the electronic charge. Webers measure magnetic charges. It was shown by Dirac (1931) that if these exist in nature, they would occur in half-integer multiples of a charge  $g = 137 e$ . (See also Section IV-E.) The charges  $e$  and  $g$  in this relation are expressed in Gaussian units, and therefore have the same dimension. If they are used as units instead of the usual Gaussian units, the material constants in vacuum reduce to dimensionless quantities  $\mu_0 = \epsilon_0^{-1} = \alpha$ . If one uses units  $e$  and  $g$ , the material constants in a

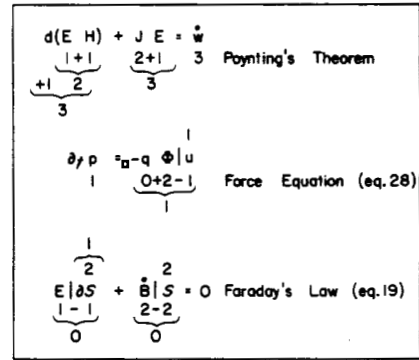


Fig. 1. Balancing the degrees of formulas.

vacuum reduce to the dimensionless quantities

$$\epsilon_0 = \mu_0^{-1} = \alpha \tag{34}$$

where  $\alpha = 1/137$  is the fine structure constant  $e^2/hc$ .

A common check on the validity of a formula, which may help detect some errors, is to verify that terms that are equated or added have the same physical dimension. When using differential forms, these terms must also have the same degree. It may be useful when learning to compute with differential forms to indicate by an underscript the degree of the various terms being considered. At the same time, one would use an overscript for the dimension of chains and the degree of multivectors. In balancing the degree of a formula, the overscripts are counted as negative degree. Fig. 1 shows examples of formulas balanced in this manner. With a little practice, one becomes conscious of the degrees of the forms involved and performs these checks mentally.

## IV. SELECTED APPLICATIONS

We shall in this section discuss applications to a few well-known problems. They should illustrate sufficiently the new formulation and some of its advantages. In engineering problems it is convenient to describe the electromagnetic field by the pair of one-forms  $(E, H)$  rather than  $(E, B)$  or  $(D, H)$ . In the following, we shall designate this pair by  $F$ .

### A. Reciprocity Relations

Consider two electromagnetic fields  $(E_1, H_1)$  and  $(E_2, H_2)$ , at frequency  $\omega$ , that satisfy Maxwell's equations in a medium  $m$  characterized by the functions  $\epsilon_0(r)$  and  $\mu_0(r)$ :

$$\begin{pmatrix} i\omega\epsilon & d \\ -d & i\omega\mu \end{pmatrix} \begin{pmatrix} E \\ H \end{pmatrix}_i = \begin{pmatrix} J \\ K \end{pmatrix}_i, \quad i = (1, 2) \tag{35}$$

where  $(J, K)$  are the electric and magnetic current two-forms.

The two-form

$$\beta_{12} = E_1 H_2 - E_2 H_1 = [E_1 H_2] \tag{36}$$

is called the *crossflux* of fields 1 and 2. The square bracket in the third term is used as a shorthand expression for the second term. It has the properties of a product: 1) linearity with respect to both factors, and 2) differentiation by the Leibnitz rule appropriate to differential forms, i.e.,

$$d[E_1 H_2] = [(dE_1) H_2 - E_1 (dH_2)]. \tag{37}$$

In the right-hand side, the square bracket still means subtrac-

tion of the expression obtained by exchanging 1 and 2 according to the general scheme  $f(1, 2) - f(2, 1) = [f(1, 2)]$ . Substituting  $dE_i$  and  $dH_i$  from (35) gives

$$d[E_1 H_2] = [J_1 E_2 - K_1 H_2]. \quad (38)$$

The terms  $[(i\omega\mu H_1) H_2]$  and  $[E_1 (i\omega\epsilon E_2)]$  equal zero because the products  $E_1 * E_2$  and  $H_1 * H_2$ , which equal  $*(E_1 \cdot E_2)$  and  $*(H_1 \cdot H_2)$  are symmetric (see Appendix F). The expression

$$\gamma_{12} = J_1 E_2 - K_1 H_2 \quad (39)$$

is the three-form reaction of the current  $(J_1, K_1)$  with the fields  $(E_2, H_2)$ .

Equation (37) can be written

$$d\beta_{12} = \gamma_{12} - \gamma_{21}. \quad (40)$$

This is the differential expression of the Lorentz reciprocity relation. The corresponding integral expression results immediately from Stokes' theorem. If the volume  $V$  has boundary  $S = \partial V$ ,

$$\beta_{12} | S = (\gamma_{12} - \gamma_{21}) | V. \quad (41)$$

Note that the two fields involved in this relation may exist in different media provided those media coincide within the volume  $V$ . The functions  $\epsilon$  and  $\mu$  may differ outside of  $V$ . Note also that the right-hand side depends only on those sources of the two fields that are inside  $V$ . If the surface  $S$  surrounds a region free of sources,  $\beta_{12} | S = 0$ . Hence, inside the region  $d\beta_{12} = 0$ , i.e.,  $\beta_{12}$  is a closed two-form.

If the sources of both fields are contained within a finite volume  $V$  with boundary  $S$ , it can be shown that  $\beta_{12} | S = 0$  obtains if both fields satisfy the Sommerfeld radiation condition. This condition can be written

$$\begin{aligned} r(Y E + dr H) &\rightarrow 0 \\ r(Z H - dr E) &\rightarrow 0 \end{aligned} \quad \text{as } r \rightarrow \infty \quad (42)$$

where  $Y, Z$  are the free-space admittance and impedance operators  $Y = (\epsilon_0/\mu_0)^{1/2} *$ ,  $Z = (\mu_0/\epsilon_0)^{1/2} *$ ; and  $r$  is the radial distance to a point in the source region. Sommerfeld's condition means that far away from the sources, the field behaves locally as a plane wave propagating in the radial direction.

Under these conditions

$$\gamma_{12} | V = \gamma_{21} | V \quad (43)$$

or

$$\gamma_{12} | R^3 = \gamma_{21} | R^3 \quad (44)$$

since  $V$  contains the sources of 1 and 2. This is the Rayleigh-Carson form of reciprocity.

### B. Huygens' Principle

According to Huygens' principle, the field  $(E_1, H_1)$  at any point  $r$  of a region  $V$ , free from sources, occupied by a medium

$m_1$  (characterized by functions  $\epsilon_1, \mu_1$ ; eventually complex to account for conduction), and bounded by the surface  $S = \partial V$ , can be expressed in terms of the components  $\underline{E}_1$  and  $\underline{H}_1$  of  $E_1$  and  $H_1$  tangent to that surface. An actual expression for that field at point  $r$  within  $V$  can be obtained if one knows the field on  $S$  due to an electric dipole  $p$  at point  $r$  in a medium  $m_p$  (functions  $\epsilon_p, \mu_p$ ) which coincides with the initial medium inside  $V$  but which may be different outside. Advantage may be taken of this freedom by modifying the medium outside of  $V$  for the purpose of simplifying the evaluation of the field  $(E_p, H_p)$  of the electric dipole  $p$ . The dipole is represented by an electric current distribution localized at point  $r$ :  $\delta_r(x, y, z) dx dy dz | p$ .

The reaction of the dipole  $p$  and the field  $(E_1, H_1)$  is simply the duality product

$$\gamma_{p1} | V = E_1(r) | p \quad (45)$$

which is also the scalar product  $\vec{E}_1 \cdot \vec{p}$ .

Applying Lorentz reciprocity to the volume  $V$ , where the media  $m_1$  and  $m_p$  coincide,

$$\gamma_{p1} | V = \beta_{p1} | S. \quad (46)$$

The right-hand side is computable since on  $S(E_1, H_1)$  are known as well as  $(E_p, H_p)$  and

$$\beta_{p1} = [E_p H_1]. \quad (47)$$

If the medium  $m_p$  consists of a conducting surface placed along  $S$  immediately outside of  $V$ , the field  $H_p$  becomes tangent to  $S$  while  $E_p$  is normal to it. The integral reduces to

$$-E_1 H_p | S \quad (48)$$

which depends only on the tangential field  $\underline{E}_1$ . Similarly, if  $m_p$  consists of a magnetic wall along  $S$ , the integral reduces to

$$E_p H_1 | S \quad (49)$$

which depends only on the tangential field  $\underline{H}_1$ .

To evaluate the field  $E_1(r)$  completely, one can use test dipoles such as  $p$ , in three orthogonal directions. This gives three components of  $E_1$ .

The magnetic field  $H_1(r)$  is evaluated similarly by means of magnetic test dipoles  $q$  in three orthogonal directions. Of course, for an actual computation of the field, in both cases one must be able to compute the field of dipoles on the surface  $S$ .

### C. Kirchhoff Approximation

Let the space  $R^3$  be divided into two regions  $V_1$  and  $V_2$  separated by a surface  $S = \partial V_1$  (its normal points toward  $V_2$ ). The surface  $S$  is composed of a screen with surface  $\Sigma$  and of an aperture  $A$  such that  $S = \Sigma \cup A$ . The sources are given in  $V_1$  and radiate a field  $(E_1, H_1)$  when the screen is absent (incident field  $F_1$ ). The problem is to find the actual field  $F_2 = (E_2, H_2)$  that results when the screen is in place. In particular, we shall look for the field diffracted in region  $V_2$ . This field would be known and computable as in Section IV-B if it was known on the (mathematical) surface  $S$ , which is assumed to lie on side  $V_2$  of the screen (see Fig. 2). Indeed, if an electric dipole  $p$  is placed at point  $r$  in  $V_2$ .

$$E_2(r) | p = \beta_{p2} | S \quad (50)$$

according to (44) and (45). This would allow one to compute

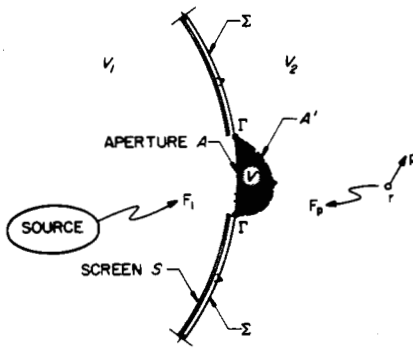


Fig. 2. Kirchhoff approximation.

$E_2(\mathbf{r})$  by using three dipoles forming a reference frame at point  $\mathbf{r}$ .

The Kirchhoff approximation consists of assuming that the field  $(\underline{E}_2, \underline{H}_2)$  tangent to  $S$  is zero on the screen  $\Sigma$  and equal to the incident field  $(\underline{E}_1, \underline{H}_1)$  tangent to  $S$  in the aperture. Then (50) becomes

$$E_2(\mathbf{r})|_{\mathbf{p}} = \beta_{p1} |A. \quad (51)$$

We note that the dipole field  $(E_p, H_p)$  used in this formula can be computed in a modified environment provided the modification occurs outside of the volume  $V_2$  in which the Lorentz reciprocity formula is applied. Thus one may place on the negative side of  $S$  either an electric or a magnetic wall making  $\underline{E}_p = 0$  or  $\underline{H}_p = 0$ . This leads to variants of the Kirchhoff approximation where only the tangential  $\underline{E}_1$  or the tangential  $\underline{H}_1$  have to be known in the aperture.

Note that the validity of the assumptions on which the approximation is based must be examined critically in any application, as there are cases where they are grossly incorrect! Also, note that the nature of the screen is not taken into account. It is sometimes considered as "absorbing." Furthermore, one must make sure that the surface  $S$  over the screen is not illuminated by the incident field. These considerations, of course, do not depend on whether vector or exterior calculus is used; therefore, we shall not discuss them further.

Consider the original problem where only one medium is involved in computing  $(E_1, H_1)$  and  $(E_p, H_p)$ . The aperture surface  $A$  may be distorted continuously into a surface  $A'$ , without changing the value of the crossflux integral  $\beta_{p1} |A$ , provided the following conditions hold:

$$\partial A' = \partial A = \Gamma \quad (52)$$

$$A' - A = \partial V \quad (53)$$

and the volume  $V$  (swept out during this deformation) does not contain the sources of  $F_1$  or the point  $\mathbf{r}$ . Thus since  $\beta_{p1}$  is closed over  $V$ , i.e.,  $d\beta_{p1} = 0$ ,

$$\beta_{p1} |(A' - A) = \beta_{p1} |\partial V = d\beta_{p1} |V = 0. \quad (54)$$

Another consequence of  $d\beta_{p1} = 0$  is that *locally* there will exist one-forms  $\alpha_{p1}$  such that

$$d\alpha_{p1} = \beta_{p1}. \quad (55)$$

If this is satisfied at all points of the aperture  $A$  for some regular one-form  $\alpha_{p1}$ , the surface integral  $\beta_{p1} |A$  can be replaced by

$$\beta_{p1} |A = d\alpha_{p1} |A = \alpha_{p1} |\partial A = \alpha_{p1} |\Gamma \quad (56)$$

which is a line integral over the contour (boundary) of  $A$ . This is essentially the Maggi-Rabinowicz integral.

Before applying (56), one must ascertain that the conditions for its validity are met: the forms  $\alpha_{p1}$  and  $\beta_{p1}$  must be regular in a domain that contains  $A$  and its boundary  $\Gamma$ , and (55) must hold in that domain.

To illustrate the importance of these conditions, let us discuss the case of scalar field solutions of the Helmholtz equation:

$$(\Delta + k^2)u = 0. \quad (57)$$

More specifically, take two fields

$$G_i(\mathbf{r}) = G(\mathbf{r} - \mathbf{s}_i), \quad i = 1, 2 \quad (58)$$

due to point sources at  $\mathbf{s}_1$  and  $\mathbf{s}_2$ . The Green's function  $G(\mathbf{r}) = e^{ik|\mathbf{r}|}/4\pi|\mathbf{r}|$ . The crossflux two-form  $\beta_{12}$  of fields  $G_1$  and  $G_2$  is

$$\begin{aligned} \beta_{12} &= *[G_1 dG_2] \\ &= *G_1 G_2 \left[ \left( ik - \frac{1}{r_2} \right) dr_2 - \left( ik - \frac{1}{r_1} \right) dr_1 \right]. \end{aligned} \quad (59)$$

One verifies easily that  $d\beta_{12} = 0$  in the domain  $D'$ , complement of the pair of points  $(\mathbf{s}_1, \mathbf{s}_2)$  where  $\beta_{12}$  is singular. If one excludes the segment  $\mathbf{s}_1\mathbf{s}_2$ , a one-form  $\alpha_{12}$  that satisfies  $d\alpha_{12} = \beta_{12}$  in the resulting domain  $D''$  is given by

$$\alpha_{12}(\mathbf{r}) = G_1 G_2 \tan \frac{\theta}{2} \rho d\phi \quad (60)$$

where  $\theta$  is the angle between the vectors  $\mathbf{r} - \mathbf{s}_1$  and  $\mathbf{r} - \mathbf{s}_2$ ,  $\rho$  is the distance from  $\mathbf{r}$  to the line  $\mathbf{s}_1\mathbf{s}_2$ , and  $\phi$  is the azimuthal angle about that line. This form is singular on the segment  $\mathbf{s}_1\mathbf{s}_2$ .

The equation  $d\alpha_{12} = \beta_{12}$  is conveniently verified in elliptical coordinates about foci  $\mathbf{s}_1, \mathbf{s}_2$ . The vector  $\vec{\alpha}_{12}$  associated with  $\alpha_{12}$  is found (for instance) in [14], where this problem is thoroughly discussed. This discussion would be appreciably simplified by the use of differential forms.

If the surface  $A$  of the aperture does not intersect the segment  $\mathbf{s}_1\mathbf{s}_2$ , (56) is valid and gives the Kirchhoff approximation to the field at  $\mathbf{s}_2$  due to the incident field  $G_1$  diffracted through  $A$  (Fig. 3(a)). If  $\mathbf{s}_1\mathbf{s}_2$  intersects  $A$  (Fig. 3(b)) it can be shown that the contribution of the singularity is precisely  $G_1(\mathbf{s}_2)$  and the field at  $\mathbf{s}_2$  is

$$u_1(\mathbf{s}_2) = G_1(\mathbf{s}_2) + \alpha_{12} |\Gamma. \quad (61)$$

The first term is the geometrical optic field, the second is the diffracted field. The expression of the latter shows that it may be considered as originating on the edge  $\Gamma$  of the aperture, an interpretation that agrees with the viewpoint of the geometrical theory of diffraction (GTD). The two terms are discontinuous on the shadow boundary, but their sum is continuous. This can also be seen by noting that  $\alpha_{12}$  is not unique. It can be modified by adding to it any closed one-form  $\gamma$ . This may be used to shift the singularities of  $\alpha_{12}$  from the segment  $\mathbf{s}_1\mathbf{s}_2$  to its complement on the line  $\mathbf{s}_1\mathbf{s}_2$  (Fig. 3(b)). The one-form

$$\alpha'_{12} = \alpha_{12} - \frac{1}{2\pi} G_1(\mathbf{s}_2) d\phi \quad (62)$$

is regular on the segment  $\mathbf{s}_1\mathbf{s}_2$  and the field  $u_1(\mathbf{s}_2)$  can be ob-



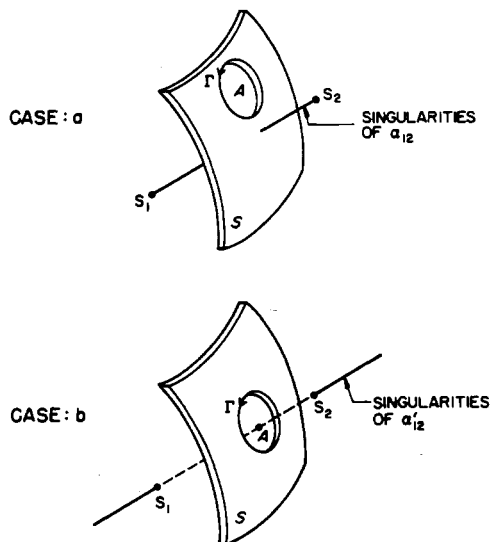


Fig. 3. Reduction of the Kirchoff approximation to a line integral.

tained by a single line integral:

$$u_1(s_2) = \alpha'_{12} | \Gamma. \tag{63}$$

*D. Applications of Stokes' Theorem and Lie Derivative*

Any of the formulas represented in Tables I and II can be converted to an integral relation by utilizing Stokes' theorem (see Appendix J). We have already invoked this theorem a number of times. For example,  $dD = \rho$  over a volume  $V$  with boundary  $S = \partial V$  implies

$$\rho | V = D | S \tag{64}$$

since  $dD | V = D | \partial V$ . It may be noted at this occasion that although Stokes' theorem is often stated only for smooth differential forms, it is valid more generally for weak forms (deRham's "currents," [9]), i.e., those forms whose coefficients are distributions. Consequently, if a unit point charge at the origin is represented by the weak form

$$\rho = \delta(x, y, z) \, dx \, dy \, dz \tag{65}$$

Equation (64) is still valid. If the origin is in  $V$ , the charge content of  $V$  is  $\rho | V = 1$ . Thus the integral  $D | S = 1$ . From this result, taking for  $S$  a sphere of radius  $r$  centered at 0 and invoking symmetry, one finds

$$D = \frac{*dr}{4\pi r^2}. \tag{66}$$

Hence,

$$E = \epsilon^{-1} D = \frac{dr}{4\pi\epsilon_0 r^2}. \tag{67}$$

(Note that the star in the operator  $\epsilon^{-1}$  cancels the one in (66).) Correspondingly,

$$\vec{E} = \frac{\hat{r}}{4\pi\epsilon_0 r^2} \tag{68}$$

a well-known result.

Another simple example is Poynting's theorem. The Poynting vector  $P$  is replaced by the Poynting two-form  $P = EH$ , which

integrates naturally over a surface  $S$ . Since

$$dP = d(EH) = (dE)H - E(dH) \tag{69}$$

(see (H.16)) and  $dE = -\dot{B}$  and  $dH = J + \dot{D}$ , we have (in a region free of currents)

$$d(EH) + (E\dot{D} + H\dot{B}) = 0. \tag{70}$$

Letting the energy three-form be

$$w = \frac{1}{2} (ED + HB) \tag{71}$$

we find

$$dP + \dot{w} = 0. \tag{72}$$

Integrating this three-form over a volume  $V$  and  $\partial V = S$ ,

$$P | S + \dot{w} | V = 0 \tag{73}$$

which has a well-known interpretation.

A variant of (72) results from considering the three-form  $\mathfrak{E} = P \, dt + w$  in the space  $R^4$ . It is easy to see that

$$d\mathfrak{E} = 0. \tag{74}$$

Hence, the integral of  $\mathfrak{E}$  over the boundary of any four-domain  $D$  is zero.

A multitude of formulas also result from computing the differential of various products. For example, in  $R^4$ , with the notations of Table IV,

$$d(\alpha\Psi) = \Phi\Psi - \alpha\gamma \tag{75}$$

so that

$$\alpha\Psi | \partial D = (\Phi\Psi - \alpha\gamma) | D \tag{76}$$

for any four-domain  $D$ . As early as 1908, Hargreaves wrote equations which, with our notation, are

$$\Phi | \partial D = 0 \tag{77}$$

and

$$\Psi | \partial D - \gamma | D = 0 \tag{78}$$

and claimed, justly, that if they held for any four-domain  $D$ , they expressed the laws of electromagnetics. This is a direct consequence of  $d\Phi = 0$  and  $d\Psi = \gamma$  (see bottoms of Tables I and II). Furthermore, equations similar to (77) and (78) are handled by Bateman in 1909 as differential forms obeying Grassmann's algebra.

Some relations between integrals, when their domains of integration are moving, are also easily written. These require the concept of the Lie derivative (L.11) and (L.12), in Appendix L, and Stokes' theorem. Consider, for example, Faraday's law

$$dE + \dot{B} = 0. \tag{79}$$

Integrating it over a surface  $S$  with boundary  $\Gamma = \partial S$  gives:

$$(dE + \dot{B}) | S = E | \Gamma + \dot{B} | S = 0. \tag{80}$$

If the surface is fixed, the last term

$$\dot{B} | S = \partial_t(B | S) \tag{81}$$

which gives the familiar integral form of Faraday's law. On the other hand, if the surface (and its boundary) are carried

along by a flow  $V^t$  defined by the vector field  $V$  (see Appendix L),

$$\partial_{t^0}(B|S_t) = (L_V B + \dot{B})|S \tag{82}$$

where  $S_t = V^t S$ . However, from (L.12)

$$L_V B = (dB)|V + d(B|V) \tag{83}$$

and since  $dB = 0$ , one has

$$(L_V B)|S = d(B|V)|S = (B|V)|\Gamma. \tag{84}$$

Thus (80) becomes

$$(E - (B|V))|\Gamma + \partial_{t^0}(B|S_t) = 0. \tag{85}$$

The one-form  $(E - (B|V)) = E'$  may be considered as a modified electric field. Reverting to vector notation, one obtains

$$\vec{E}' = \vec{E} + \vec{V} \times \vec{B} \tag{86}$$

therefore, equation (85) represents the induction law for a moving circuit  $\Gamma = \partial S$ . This equation is rigorously correct and is not, as is sometimes claimed, the first approximation to a relativistic formula. Relativity plays no role in its deduction.

Note that the domain of integration is arbitrary and that its "motion" does not necessarily coincide with the motion of the matter it contains. The variable  $t$  may not be the real time, but a parameter that characterizes the variation (flow) of the domain and, eventually, the variation of the differential forms involved.

On the other hand, if the vector field  $U$  represents the flow of electric charge and if the volume  $V$  is carried along with that flow to  $U^t V = V_t$ , the content of  $V_t$  is constant. This expresses the continuity of charge.

Using the Lie derivative,

$$\partial_{t^0}(\rho|V_t) = (\dot{\rho} + L_U \rho)|V = 0 \tag{87}$$

and

$$L_U \rho = d(\rho|U) + (d\rho)|U = d(\rho|U) \tag{88}$$

since  $d\rho = 0$ . Thus the continuity condition in integral form is

$$\dot{\rho}|V + (\rho|U)|\partial V = 0. \tag{89}$$

### E. Applications of Poincaré's Lemma

As shown in Appendix J, the Poincaré lemma gives the answer to the following question: given a  $p$ -form  $\alpha$ , is it possible to express it as the differential of a  $(p - 1)$ -form  $\beta$ , over some domain  $D$ ? Since  $\alpha = d\beta$  implies  $d\alpha = d(d\beta) = 0$ , it is necessary that  $\alpha$  be closed over  $D$ . This is also sufficient locally, i.e., in some spherical neighborhood of any point in  $D$ ; or globally over  $D$  if that domain has appropriate topological properties (see Appendix J). The proof of the lemma actually provides a construction for the form  $\beta$ . (This construction, however, is not always the most convenient.) Having obtained  $\beta$ , the integration of  $\alpha$  over some  $p$ -domain (or  $p$ -chain) in  $D$  is reduced to that of  $\beta$  over the boundary of that domain (or chain).

If  $\alpha$  describes a field,  $\beta$  is in general called its "potential." Thus, a potential's existence depends critically upon the topology of the region over which it is to be used. A common feature of a potential is its lack of uniqueness:  $\beta$  may be increased by any closed form or by an exact differential  $d\gamma$ ,

without ceasing to be a solution. Perhaps this is what led Heaviside to call the potentials "treacherous and useless." The "useless" meant that they could be dispensed with. It turns out, however, that when quantum mechanics is involved [10], potentials are more meaningful than Heaviside suspected.

As an example of Poincaré's lemma, consider Faraday's equations:

$$\begin{cases} dE + \dot{B} = 0 \\ dB = 0. \end{cases} \tag{90}$$

Letting  $\Phi = B + E dt$ , they can be expressed, in space-time, by the single equation

$$d\Phi = 0. \tag{91}$$

This equation holds throughout  $R^4$  and asserts the non-existence of free magnetic charges. Poincaré's lemma applies globally to the space  $R^4$ ; hence, there exists a one-form  $\alpha$  such that

$$\Phi = d\alpha. \tag{92}$$

Writing

$$\alpha = A - \phi dt \tag{93}$$

where  $A$  is a one-form in space and  $\phi$  is a scalar function, equation (92) becomes

$$\begin{cases} B = dA \\ E = -d\phi - \dot{A}. \end{cases} \tag{94}$$

[Note that the minus sign in (93) and in some of the other equations considered, such as  $\Psi = D - H dt$ , is not particularly significant. These signs have been introduced to simplify the correspondence with the usual vectorial expressions.] A gauge transformation

$$\alpha \rightarrow \alpha' = \alpha + df \tag{95}$$

where  $f$  is a scalar function, gives another solution  $\Phi = d\alpha'$ , since  $d \cdot d = 0$ .

To show the importance of the topology on the global existence of a potential, we consider two examples: 1) the magnetic field of a steady linear electric current (a well-known problem) and 2) the field  $B$  (two-form) due to a hypothetical magnetic charge (Dirac [11], [12]).

In the first example, a steady unit current, along the  $z$ -axis,  $Oz$ , is represented by the two-form

$$J(x, y, z) = \delta(x, y) 1(z) dx dy \tag{96}$$

where  $1(z)$  is the function equal to 1 for all values of  $z$ . The magnetic field satisfies

$$dH = J \tag{97}$$

thus,  $dH = 0$  in the domain  $D$ , complement of the  $z$ -axis. This domain is not simply connected, hence, we cannot assert the existence of a function  $f$  (zero-form) such that

$$df = J \tag{98}$$

in  $D$ . However, if we remove from  $R^3$  the half-plane  $P_1: y = 0, x < 0$ , the resulting domain  $D_1 = R^3 \setminus P_1$  is simply connected and there exists a function  $f_1$  such that

$$df_1 = H \quad \text{in } D_1. \tag{99}$$

In fact,  $H$  can be obtained by a classical argument: integrating  $dH = J$  over a disk  $S$  of radius  $r$  and axis  $Oz$ , bounded by the circle  $\Gamma = \partial S$ :

$$J|S = dH|S = H|\partial S = H|\Gamma \quad (100)$$

and then invoking the rotational symmetry,

$$H = \frac{d\phi_1}{2\pi} \quad (101)$$

where  $\phi_1$  is the azimuthal angle  $\phi$  about  $Oz$  restricted to

$$-\pi < \phi \leq \pi. \quad (102)$$

Equation (101) corresponds to the well-known  $\vec{H} = \hat{\phi}_1/2\pi r$ . (Note that  $\phi$  without such a restriction is *not* a function over  $D$  since it is multivalued. Therefore, strictly speaking,  $d\phi$  is not an exact one-form over  $D$ .) A scalar potential for  $H$  in  $D$  is

$$f_1 = \frac{\phi_1}{2\pi}. \quad (103)$$

It is discontinuous across the cut represented by the half-plane  $P_1$ . If we had chosen the half-plane  $P_2: y = 0, x > 0$ , as a cut, we could have used the potential

$$f_2 = \frac{\phi_2}{2\pi} \quad (104)$$

where  $\phi_2$  is the azimuthal angle restricted to the range

$$0 \leq \phi_2 < 2\pi. \quad (105)$$

It is discontinuous across the cut represented by the half-plane  $P_2$ . In any given region of overlap the two potentials  $f_1$  and  $f_2$  must differ only by a constant since  $H = df_1 = df_2$ . Thus, for  $y > 0$ ,  $f_1 = f_2$ , while for  $y < 0$ ,  $f_2 = f_1 + 1$ . The constants are different in the two regions.

This simple example shows how to handle cases where, because of the topology of the domain  $D$ , a global potential does not exist. A set of potentials can be constructed, each valid over a "simple" region. In the overlap of the two regions, the two potentials, in order to give the same field, are related by a gauge transformation (in  $y < 0$ , it is simply  $f_1 \rightarrow f_2 = f_1 + 1$ ). When more than two regions overlap, the gauge transformations between each pair must satisfy obvious compatibility conditions (form a group). These considerations are relevant to the theory of gauge fields [13].

Using the two potentials  $f_1$  and  $f_2$ , one can evaluate the circulation  $H|\Gamma$  over a cycle  $\Gamma$  that surrounds the origin  $O$ . We can decompose it into  $\Gamma_1 \cup \Gamma_2$  (see Fig. 4) where  $\Gamma_1 \subset D_1$  and  $\Gamma_2 \subset D_2$  and the endpoints  $a$  and  $b$  have been chosen such that  $a$  is in the region  $y < 0$  and  $b$  is in the region  $y \geq 0$ . We have  $\partial\Gamma_1 = b - a$  and  $\partial\Gamma_2 = a - b$ . Applying Stokes' theorem to each path,

$$H|\Gamma_1 = f_1|(b - a) = f_1(b) - f_1(a) \quad (106)$$

and

$$H|\Gamma_2 = f_2|(a - b) = f_2(a) - f_2(b). \quad (107)$$

Thus

$$H|\Gamma = H|(\Gamma_1 + \Gamma_2) = (f_2 - f_1)(a) - (f_2 - f_1)(b) = 1. \quad (108)$$

The second example is provided by Dirac's analysis of the properties of a magnetic monopole, assuming that one exists

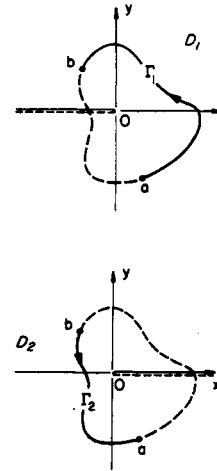


Fig. 4. Scalar potentials for the magnetic field of a linear electric current.

[11], [12]. Such a monopole of strength  $\mu$  would be represented by a three-form

$$\sigma = \mu \delta(x, y, z) dx dy dz. \quad (109)$$

The two-form  $B$  satisfying  $dB = \sigma$  can be obtained directly in the same manner as the  $D$  field of an electric charge. It is given by

$$B = \mu \frac{*dr}{4\pi r^2} = \frac{\mu}{4\pi} \sin \theta d\theta d\phi. \quad (110)$$

An interpretation of  $B$  is that for any surface  $S$ , the integral  $B|S$  is  $\mu$  times the solid angle under which  $S$  is seen from the origin (in units of  $4\pi$ ).

In order to describe the motion of an electron in that field, according to quantum mechanics, one seeks a wavefunction  $\psi$  that satisfies Schrödinger's equation, an equation that depends on a vector potential  $\vec{A}$  (or the corresponding one-form  $A$ ). Since  $dB$  is not zero over the entire space, but only on the domain  $D$ , complement of the origin, such a one-form does not exist globally. This is because  $D$  does not have the appropriate topology: a surface surrounding the origin can not be shrunk continuously to a point without getting out of  $D$ . Following the process used in the preceding example, we can consider the domains  $D_1$ , complement of the negative  $z$ -axis, and  $D_2$ , complement of the positive  $z$ -axis. (These removed half-axes have been called strings [12].) In these two domains potentials can be found such that  $dA_i = B$  over  $D_i$ ;  $i = 1, 2$ . It is easily verified that

$$A_1 = \frac{\mu}{4\pi} (1 - \cos \theta) d\phi \quad (111)$$

$$A_2 = -\frac{\mu}{4\pi} (1 + \cos \theta) d\phi \quad (112)$$

satisfy these conditions. One can also verify that  $A_1$  is regular for  $\theta = 0$  and singular for  $\theta = \pi$ . The converse is true for  $A_2$ . In the intersection  $D_c = D_1 \cap D_2$ , complement of the  $z$ -axis, the two potentials  $A_1$  and  $A_2$  are related by a gauge transformation

$$A_2 - A_1 = \frac{\mu}{2\pi} d\phi. \quad (113)$$

A potential  $A_c = -(\mu/4\pi) \cos \theta \, d\phi$  is also valid in that domain.

It can be shown (see standard texts in quantum mechanics) that a change in phase of the wave function combined with a gauge transformation of the potential one-form preserves the form of Schrödinger's equation. More precisely, if the potential is expressed in units of  $g = \hbar c/e$ , the transformation

$$A \rightarrow A' = A + \omega \tag{114}$$

where  $\omega$  is a closed one-form, corresponds to a change of the wave function

$$\psi \rightarrow \psi' = \psi e^{i\omega|\Gamma} \tag{115}$$

where  $\Gamma$  is a path ending at the observation point. When the path  $\Gamma$  is closed (initial and observation points coincide), the phase factor  $e^{i\omega|\Gamma}$  must be one in order to make the wave function single-valued; i.e.,  $\omega|\Gamma$  must be an integer multiple of  $2\pi$ .

In the present case, from (106)

$$\omega = \frac{\mu}{2\pi} \, d\phi. \tag{116}$$

Taking for  $\Gamma$  a circle with axis  $Oz$ ,

$$\omega|\Gamma = \mu. \tag{117}$$

Hence,  $\mu$  must be an integral multiple of  $2\pi$  or in terms of the units chosen, a multiple of  $g/2$ . Thus a unit magnetic monopole has charge  $g/2$ ; it is 57.5 times stronger than a unit electric charge  $e$ .

### CONCLUSION

The calculus of exterior differential forms offers an attractive alternative to the conventional vector calculus for the formulation and handling of the equations of electromagnetics. It plays an important role as well for a number of topics in physics: classical mechanics, geometrical optics, quantum mechanics, and more recently, gauge field theories.

This article was meant to introduce this subject for the particular application to electromagnetics. Exterior differential forms are particularly relevant because they represent electromagnetic quantities quite naturally. The relations between these quantities are expressed by means of the differential operator "d," which takes the place of the curl, grad, and div operators and which obeys simple rules. The equations have been displayed in flow diagrams. Their invariance under changes of coordinates in three or four dimensions simplifies the use of curvilinear coordinates. Physical dimensions of all forms related by d are the same; hence, two units (electric and magnetic charge) suffice for all the forms involved.

Only a few applications could be sketched in this paper, and they have been deliberately chosen among the simplest ones. They should give the reader at least a flavor of the exterior calculus, demonstrate its appropriateness, the automatic nature of the computations involved, the ease in changing variables, and the conciseness of the expressions obtained. Many interesting aspects had to be left out such as the consideration of weak forms (deRham's currents)—these are almost essential to electromagnetics when wires and surfaces are considered—the consideration of flows in optics, symplectic geometry and its applications to reciprocity (topics for future articles!).

In spite of its shortcomings, it is hoped that this article may help realize this prediction of H. Flanders about exterior calculus: "Physicists are beginning to realize its usefulness. Perhaps it will soon make its way into engineering."

### APPENDICES

#### A BRIEF REVIEW OF DIFFERENTIAL GEOMETRY

##### A. Vector Space

A vector space  $E$ , over the set of real numbers  $R$  (scalars) is a set of elements, called vectors, and two operations: addition, which assigns to a pair of vectors  $(x, y)$  a vector  $x + y$ , and multiplication by a scalar  $a \in R$ , which transforms vector  $x$  into vector  $ax$ . The set  $R$  may be replaced by any "field" of numbers, for instance the set  $C$  of complex numbers. The above operations satisfy well-known properties that will not be repeated here.

Vectors  $e_1, e_2, \dots, e_n$  are said to be linearly independent if  $\sum a_i e_i = 0$  implies that all  $a_i = 0$ . These vectors form a *basis* if any vector  $x$  is expressible as

$$x = x^i e_i. \tag{A.1}$$

(The Einstein summation convention with respect to the repeated index  $i$  is used throughout.)

The  $x^i$  coordinates of  $x$  are uniquely defined by  $x$ . The space  $E$  is thus isomorphic to the direct product  $R^n$  of  $n$  replicas of  $R$ . The number  $n$  is the dimension of  $E$ . A linear transformation  $\mu$  from space  $E$  of dimension  $n$  to space  $F$  of dimension  $m$  is defined if its action on all basis vectors  $e_i$  is known:

$$\mu e_i = \mu^j_i f_j \tag{A.2}$$

where  $\{f_j\}$  is a basis of  $F$ .

The numerical coefficients  $\mu^j_i$  form an  $m \times n$  matrix. When  $m = n$  and  $\det \mu^j_i \neq 0$ ,  $\mu^{-1}$  is defined, and  $\mu$  is an isomorphism of  $E$  and  $F$ .

##### B. Duality—Covectors

A linear function  $\xi$  over the space  $E$  with values in  $R$  is a function

$$\xi: E \rightarrow R: x \mapsto \xi(x) \tag{B.1}$$

which satisfies

$$\xi(ax + by) = a\xi(x) + b\xi(y) \tag{B.2}$$

for any scalars  $(a, b)$  and any vectors  $(x, y)$ .

The set of linear functions over  $E$  form a space  $E^*$  called the dual of  $E$ . This space takes the structure of a vector space by the natural definitions

$$(\xi + \eta)(x) = \xi(x) + \eta(x) \tag{B.3}$$

$$(a\xi)(x) = a\xi(x), \quad x \in E. \tag{B.4}$$

Its elements are called *covectors* and will be denoted here by greek letters with the exception of the electromagnetic quantities which in the main text are designated by the conventional notations:  $E, B, D, H$ , etc.

A basis of  $E^*$  is a set of linearly independent covectors  $\epsilon^1, \epsilon^2, \dots, \epsilon^n$  such that any  $\xi \in E^*$  can be expressed by

$$\xi = \xi_i \epsilon^i. \tag{B.5}$$

It can be shown that the dimension of  $E^*$  is  $n$  and that the dual of  $E^*$  may be taken as  $E$ .

To emphasize the symmetry between  $E$  and  $E^*$ , we shall denote  $\xi(x)$  by  $\xi|x$ , or simply by  $\xi x$  when no confusion is to be feared; that is, when it is clear that the two factors are, respectively, a covector and a vector. We shall call  $\xi|x$  the *duality product* of  $\xi$  and  $x$ . It is indeed a "product," i.e., a linear function of each of the two factors.

If we let  $\epsilon^i$  be the linear function that takes the value 1 for  $e_i$  and the value zero for  $e_j$  ( $j \neq i$ ), the set  $\{\epsilon^i\}$  is called the *dual basis* of  $\{e_i\}$ . If  $\xi$  and  $x$  are expressed in dual bases, their duality product

$$\xi|x = (\xi_i \epsilon^i)(x^j e_j) = \xi_i x^i. \quad (\text{B.6})$$

It is similar in form to a scalar product  $x \cdot y = x^i y^i$ . The difference is that for the scalar product the two factors belong to the same vector space while for the duality product they belong to two distinct dual spaces [observe the up and down positions of the indices].

### C. Exterior Algebra—Multivectors

In the  $n$ -dimensional vector space  $E$ , a vector  $x$  can be represented as a sum

$$x = x^i e_i \quad (\text{C.1})$$

where  $(e_1 \cdots e_n)$  are  $n$  vectors that form a *basis* of  $E$  and the  $x^i$  are real numbers.

An *algebra* is constructed on  $E$  by defining products of the basis vectors  $e_i$ , and by applying the familiar rules of associativity and distributivity. For example, the complex algebra  $C$  is a two-dimensional vector space over the field of real numbers with basis  $(e_0, e_1)$  and multiplication rules  $e_0^2 = e_0$ ;  $e_0 e_1 = e_1 e_0 = e_1$ ;  $e_1^2 = -e_0$ . The usual notation  $e_0 = 1$ ,  $e_1 = \sqrt{-1} = i$ , makes  $R$  a subspace of  $C$ . Another example is the quaternion algebra, a four-dimensional vector space over  $R$  based on  $(e_0, e_1, e_2, e_3)$  with the following multiplication rules:  $e_0^2 = e_0$ ;  $e_i^2 = -e_0$  and  $e_i e_0 = e_0 e_i = e_i$  (for  $i = 1, 2$  or  $3$ ); and  $e_i e_j = e_k$  (where  $(i, j, k)$  form an even permutation of  $(1, 2, 3)$ ).

In these two examples, the products formed remain within the vector space  $E$ . In contrast, the *exterior algebra* produces new elements that lie outside of  $E$ . The product of two basis vectors  $e_i$  and  $e_j$ , denoted by  $e_{ij}$ , satisfy the rule

$$e_j e_i = -e_i e_j$$

or

$$e_{ij} = -e_{ji}. \quad (\text{C.2})$$

Hence, in particular,  $e_i e_i = 0$ .

Linear combinations of these  $e_{ij}$  generate a vector space of dimension  $n(n-1)/2$ , denoted by  $\Lambda^2 E$  or  $E^2$ . Its elements are called *two-vectors*.

For instance, the product of vectors  $(x, y)$  in  $R^3$  is

$$xy = (x^2 y^3 - x^3 y^2) e_{23} + (x^3 y^1 - x^1 y^3) e_{31} + (x^1 y^2 - x^2 y^1) e_{12}. \quad (\text{C.3})$$

The coefficients have the same expressions as those of a vector product. The product  $xy$  would actually be the vector product if the basis  $e_i$  was orthonormal and if  $(e_{23}, e_{31}, e_{12})$  were replaced by  $(e_1, e_2, e_3)$ .

This gives a hint for an interpretation of a two-vector. It represents the parallelogram defined by  $x$  and  $y$  through its area, the direction of its plane, and a sign which indicates the order of the factors  $(x, y)$ .

The expression  $xy$  is more general than the vector product in two respects: 1) It applies to bases that are not orthonormal (the space  $E$  may not be endowed with a "metric" which is necessary to give a meaning to orthogonality and normality (see Appendix E). The elements  $e_{ij}$  that form a basis of  $\Lambda^2 E$  represent the "extent" of the oriented parallelograms defined by pairs  $(e_i, e_j)$ . 2) The exterior product is defined for any dimension of the space  $E$ . When  $n \neq 3$ , it is not possible to associate a vector in  $E$  to a two-vector. For instance, in  $R^4$  a two-vector has six components. Note that a general two-vector  $z = \sum z^{ij} e_{ij} \in \Lambda^2 E$  is not always representable as a product of two-vectors, but by a sum of such products.

Continuing the construction of products, we introduce a basis for three-vectors by defining  $e_{ijk} = e_i e_j e_k$ , a product that obeys associativity, but is anticommutative. There are  $C_n^3 = n(n-1)(n-2)/3!$  such independent products that span the vector space  $\Lambda^3 E = E^3$  of *three-vectors*. More generally,  $p$ -vectors form a space  $\Lambda^p E$  of dimension

$$C_n^p = n!/p!(n-p)!. \quad (\text{C.4})$$

In particular,  $\Lambda^n E = E^n$  has dimension one. Its elements are scalar multiples of  $e_1 e_2 \cdots e_n = e_N$ , where  $N = 1, 2, \dots, n$ . For  $p > n$  or  $p < 0$ ,  $\Lambda^p E$  is empty. For  $p = 0$ , we let  $\Lambda^0 E = R$  (or whatever "field" of scalars has been used to define  $E$ ).

A  $p$ -vector  $\alpha$  can be expressed as a sum

$$\alpha = \sum a^J e_J \quad (\text{C.5})$$

where  $J$  is a  $p$ -index; i.e., a sequence

$$J = j_1 j_2 \cdots j_p \quad (\text{C.6})$$

of indices  $j_k \in [1, 2, \dots, n]$ , and  $e_J$  stands for  $e_{j_1} e_{j_2} \cdots e_{j_p}$ .

Because of the skew symmetry property

$$e_i e_j = -e_j e_i \quad (\text{C.7})$$

the sum may be reduced to one where no two  $J$ 's are composed of the same elements in different order. One can achieve this by using only indices  $J$  such that  $j_1 < j_2 < \cdots < j_p$ . These will form a set  $\mathcal{J}_p$  and in (C.5) we may restrict the sum to run over that set. We may write

$$\alpha = \sum^0 a^J e_J \quad (\text{C.8})$$

or simply  $\alpha = a^J e_J$ , generalizing Einstein's convention to ordered multi-indices. Note that other conventions may be used to define the set  $\mathcal{J}_p$  that also lead to a reduced expression. In three dimensions, for example, it is convenient to choose for  $\mathcal{J}_2$  the set  $(23, 31, 12)$  rather than  $(23, 13, 12)$  as this leads to more harmonious expressions.

In the Minkowski space  $R^{3+1}$ , we can take for  $\mathcal{J}_3$  the set  $(234, 314, 125, 123)$  where the fourth coordinate (time) plays a special role.

The direct sum of all the  $\Lambda^p E$  forms the exterior algebra  $\Lambda E$ . The rules for adding and multiplying any two elements are those of ordinary algebra, except that one must pay attention to the order of the factors.

A scalar (or zero-vector)  $a$  commutes with any  $p$ -vector, and therefore can be moved in the sequence of factors:

$$a(xy) = (ax)y = x(ay). \quad (\text{C.9})$$

(The  $p$ -vectors may be considered as contravariant skew-symmetric tensors, but there is little to be gained from this

point of view as exterior calculus can be carried out without reference to it.)

An alternative expression for a  $p$ -vector results from allowing  $J$  to take all possible  $p!$  values obtained by permutation, instead of only those in  $\mathcal{J}_p$ , and from requiring that the  $a^j$  be skew symmetric with respect to the  $j$ 's. Then,

$$x = \frac{1}{p!} \sum a^J e_J. \quad (C.10)$$

To form the product of a  $p$ -vector,  $x = a^J e_J$ ,  $J \in \mathcal{J}_p$  and a  $q$ -vector,  $y = b^K e_K$ ,  $K \in \mathcal{J}_q$ , one uses distributivity and the property

$$e_J e_K = e_{JK} \quad (C.11)$$

$JK$  being the  $(p+q)$ -index obtained by juxtaposition of  $J$  and  $K$ . The resulting expression is then reduced.

It is easy to verify that

$$yx = (-)^{pq} xy. \quad (C.12)$$

The exterior product of  $xy$  is often denoted by  $x \wedge y$  and called a "wedge product." As long as only exterior products are considered (which is the case for most of this article), the sign can be omitted. This is a current practice when dealing with differentials, and we shall apply it to all multivectors and multivectors considered. Other types of products will be distinguished by a sign, for instance, a dot for the inner product (see Appendix E), a vertical bar for the duality product (see Appendices B and D).

An exterior algebra can also be constructed on the dual space  $E^*$ . It will be denoted by  $\Lambda(E^*)$  and is the direct sum of the spaces of  $p$ -covectors  $\Lambda^p(E^*) = E_p$ .

#### D. Extensions of Duality

The duality product defined at first for a pair of elements,  $\xi$  in  $E^*$  and  $x$  in  $E$

$$E^* \times E \rightarrow R: (\xi, x) \rightarrow \xi|x \quad (D.1)$$

is extended to the product of a  $p$ -covector by a  $p$ -vector

$$E_p \times E^p \rightarrow R: (\xi, x) \rightarrow \xi|x \quad (D.2)$$

by noting that  $E_p$  is the dual of  $E^p$ , dual bases for these spaces being  $\{e^J\}$  and  $\{e_J\}$  where  $J \in \mathcal{J}_p$ .

The duality product can be further extended to a bilinear operation

$$E_p \times E^q \rightarrow E_{p-q}, \quad q < p \quad (D.3)$$

defined by the condition that

$$(\xi, x) \rightarrow \xi|x \quad (D.4)$$

iff for every  $y \in E^{p-q}$

$$(\xi|x)|y = \xi|(xy). \quad (D.5)$$

(The vertical bars can be omitted if one takes into account the nature of the various factors.) The duality product is also called inner product, or *contraction*.

An important special case obtains when  $x$  is a vector ( $q = 1$ ). The notation  $i_x \xi$  is then often used instead of  $\xi|x$ . The operator  $i_x$  transforms  $p$ -forms into  $(p-1)$ -forms. Applied to an

exterior product of covectors  $\xi\eta$  it acts as a derivative, obeying the modified Leibnitz rule:

$$i_x(\xi\eta) = (i_x \xi)\eta + (-)^p \xi(i_x \eta) \quad (D.6)$$

where  $\xi \in E_p$ .

Note that if  $x = e_1$  and  $\xi = \epsilon^1 \alpha$ , where  $\alpha \in E_{p-1}$  does not contain any  $\epsilon^1$ , we have

$$(\epsilon^1 \alpha) e_1 = \alpha. \quad (D.7)$$

Thus  $i_{e_1}$  acts as a division by  $\epsilon^1$  from the left.

#### E. Metric Vector Space

The vector space  $E = R^n$  is endowed with a metric (more precisely, a Euclidean metric) if a scalar-valued symmetric bilinear function  $g(x, y)$  is defined for every pair of vectors  $(x, y)$  in  $E$ . This function is called the *scalar* or *dot* product of  $x$  and  $y$ , and is often denoted by  $x \cdot y$ . It is assumed to be a nondegenerate; i.e.,  $g(x, y) = 0$  for every  $y$  implies  $x = 0$ .

When the quadratic form  $g(x, x)$  is positive definite, i.e.,

$$g(x, x) > 0, \quad \text{for } x \neq 0 \quad (E.1)$$

one defines the length of  $x$  as the positive square root of  $g(x, x)$ . When the quadratic form is not positive definite, one can define the "extent" of  $x$  as the square root of  $|g(x, x)|$  and add some qualification that indicates the sign of  $g(x, x)$ . For example, in the Minkowski space, where for a vector  $v$  with coordinates  $(x, y, z, \tau = ct)$

$$g(v, v) = x^2 + y^2 + z^2 - \tau^2 \quad (E.2)$$

one calls the vector  $v$  space-like when  $g > 0$  and time-like when  $g < 0$ .

Two vectors  $(x, y)$  are said to be orthogonal if

$$g(x, y) = 0. \quad (E.3)$$

Besides lengths (or extents) of vectors, one can also define the angle between two vectors. When the metric is not positive definite, some of these angles are imaginary and correspond to real hyperbolic angles.

From an arbitrary basis for the space  $E$ , one can deduce special bases  $\{e_i\}$ ,  $i \in (1, 2, \dots, n)$  that are *semiorthonormal*. This means that any pair  $e_i, e_j$  is orthogonal:

$$e_i \cdot e_j = 0 \quad (E.4)$$

and that any  $e_i$  has a scalar square equal to  $\pm 1$ .

We shall express this by

$$e_i \cdot e_i = (-)^{s(i)} \quad (E.5)$$

introducing a function  $s$  on the set  $\{1, 2, \dots, n\}$  that takes values 0 or 1. For a  $p$ -index  $J = j_1 j_2 \dots j_p$ , we define

$$s(J) = s(j_1) + \dots + s(j_p). \quad (E.6)$$

The value  $s(J)$  is the number of elements  $e_j$  with  $j \in \{j_1, \dots, j_p\}$ , that have a negative square.

When the metric is *positive definite* ( $g(x, x) > 0$  for all  $x$  except 0), there exists *orthonormal bases*, i.e., bases such that  $s(i) = 0$ ; hence, all  $e_i$  are of unit length. When the metric is not positive definite, some of the  $e_i$  have a negative scalar square. The number of these particular units, which is  $s(N)$

for  $N = (12 \cdots n)$  is the same for all semiorthonormal bases corresponding to a given metric (Sylvester's law of inertia).

An alternative description of the metric is based on the observation that for a fixed vector  $x$ , the function

$$g(x, \cdot) : y \rightarrow g(x, y) \tag{E.7}$$

is linear over the space  $E$ . Hence, it may be considered as an element of the dual  $E^*$ . This element depends linearly on  $x$ . We shall denote it by  $G(x)$  and call it the *mate* of  $x$ . Thus

$$G : E \rightarrow E^* : x \rightarrow G(x) \tag{E.8}$$

and

$$g(x, y) = G(x)|y. \tag{E.9}$$

Thus the scalar product that combines two vectors  $x$  and  $y$  is replaced by a duality product that combines a vector  $G(x)$ , the mate of  $x$ , with the vector  $y$ . We shall designate the mate by an overbar:  $G(x) = \bar{x}$  and (E.9) becomes

$$x \cdot y = \bar{x}|y. \tag{E.10}$$

With respect to an arbitrary basis  $\{e_i\}$  the scalar product is defined if it is known for every pair  $(e_i, e_j)$ .

Letting

$$e_i \cdot e_j = g_{ij} \tag{E.11}$$

the scalar product of  $x = x^i e_i$  and  $y = y^j e_j$  is

$$g(x, y) = g_{ij} x^i y^j. \tag{E.12}$$

It is symmetric:  $g_{ij} = g_{ji}$ ; and nondegenerate:  $\det g_{ij} \neq 0$ .

If  $\{e^j\}$  is the dual basis of  $\{e_i\}$ , the linear mapping  $G$  is represented by the matrix  $g_{ij}$  as follows. If  $G(x) = x^i G(e_i)$  is expressed by  $\xi_j e^j$ , equation (E.12) implies

$$\xi_j = g_{jh} x^h \tag{E.13}$$

and

$$\bar{e}_i = G(e_i) = g_{ij} e^j. \tag{E.14}$$

Since  $g_{ij}$  is nondegenerate, the transformation (E.13) can be inverted. This defines a map

$$G^{-1} : E^* \rightarrow E : \xi \rightarrow G^{-1}(\xi) \tag{E.15}$$

which associates to every covector  $\xi$  a mate  $G^{-1}(\xi)$  that will also be designated by an overbar  $\bar{\xi}$ . Both operations  $G$  and  $G^{-1}$  (sometimes called  $g$  flat and  $g$  sharp) can be represented by a single involutive overbar operator<sup>2</sup>  $\bar{O}$ . No confusion results since  $E$  and  $E^*$  are disjoint. (The situation is similar to the use of a transpose operator to relate row vectors and column vectors.)

A natural choice for a metric in  $E^*$ , related to the one in  $E$ , is defined by the scalar product of two covectors,  $\xi$  and  $\eta$ :

$$\xi \cdot \eta = \xi|\bar{\eta}. \tag{E.16}$$

The metric can also be extended to multivectors and multicovectors in a natural manner. It is sufficient to define the overbar of the generators  $e_J$  (or  $e^J$  for covectors). If the  $\{e_j\}$  form a semiorthonormal basis and  $\{e^j\}$  is a dual basis,

the mate of  $e_j$  is  $(-)^{s(j)} e^j = \bar{e}_j$  and that of a product of several  $e_j$  is the product of the corresponding  $\bar{e}_j$ . Thus, a  $p$ -vector element  $e_J$  has for its mate  $(-)^{s(J)} e^J$ . By linearity, this defines the metric over  $\Lambda^p E$  and in a similar manner over  $\Lambda^p E^*$ .

### F. Star Operator

We have noted that the space  $E^p$  of  $p$ -vectors and the space  $E^{n-p}$  of  $(n-p)$ -vectors have the same dimension  $C_n^p$ . Consequently, there exist one-to-one linear transformations of the algebra  $\Lambda E$  onto itself that map each  $E^p$  onto  $E^{n-p}$ . Among these transformations a particular one, denoted by  $*$  and called the (Hodge) star operator, bears a close relationship to a given metric of  $E$ , extended to  $\Lambda E$  (see Appendix E). Thus among other applications, the star operator can replace the scalar product or the overbar operator as a characteristic of the given metric.

For a pair of multivectors  $(x, y)$ , the relation to the scalar product is

$$*(x \cdot y) = x * y \tag{F.1}$$

where  $x * y$ , to be read from right to left, is short for  $x(*y)$ . Since  $*$  is a linear transformation and (F.1) is linear in  $x$  and  $y$ , it is sufficient to define the effect of  $*$ , and to verify (F.1), for the elements of a basis. We shall assume this basis (semi) orthonormal and take  $x = e_I, y = e_J$ . The analysis of (F.1) then shows that both sides are zero unless  $I = J$ , and that a solution for  $*e_J$  is

$$*e_J = (-)^{s(K)} (-)^\sigma e_K \tag{F.2}$$

where  $K$  is a complement of  $J$ , i.e., a sequence such that  $JK$  is a permutation of  $N = (1, 2, \dots, n)$ ,  $(-)^s = +1$  for even  $\sigma$ ,  $-1$  for odd  $\sigma$ .

Special cases of (F.2) are

$$*1 = (-)^{s(N)} e_N \tag{F.3}$$

and

$$*e_N = 1. \tag{F.4}$$

Another solution would have been the negative of (F.2), which obviously also satisfies (F.1). It would have resulted from replacing  $e_N$  by  $e_{N'}$  where  $N'$  is an odd permutation of  $N$ . The particular ordering of indices in  $N$  defines the *orientation* of space. It must be considered as a part of the definition of  $*$ . For a positive definite metric  $s(N) = 0$  and the formulas simplify accordingly.

In its applications to electromagnetics, the star operator that is of most use is the one defined over the exterior algebra generated by differential forms; hence, instead of using  $\{e^j\}$  to represent the orthonormal basis of the algebra, we shall use for a basis the differentials  $(dx, dy, dz)$  of some orthonormal coordinates for space and  $(dx, dy, dz, d\tau)$ , with  $\tau = ct$ , for the basis in space-time. The bases (for vectors) dual of these are  $(\partial_x, \partial_y, \partial_z)$  and  $(\partial_x, \partial_y, \partial_x, \partial_\tau)$ , respectively.

For the three-dimensional space, the positive definite metric would be defined for vector

$$V = X\partial_x + Y\partial_y + Z\partial_z \tag{F.5}$$

by

$$\bar{O}V = \bar{V} = X dx + Y dy + Z dz \tag{F.6}$$

$$\bar{V}|V = X^2 + Y^2 + Z^2. \tag{F.7}$$

<sup>2</sup>Any operator that is represented by adding a subscript, a superscript, an overbar, a star, or any recognized "ornament" may be represented by the letter  $\bar{O}$  embellished by the same ornament. Thus

$$\bar{O} : x \rightarrow \bar{x}, O_\tau : f \rightarrow f_\tau, O^* : f \rightarrow f^*, \text{ etc.}$$

The star operator  $*$  computed from (F.2) is defined by

$$*: 1 \rightarrow dx \, dy \, dz \rightarrow 1, \quad dx \rightarrow dy \, dz \rightarrow dx \quad (F.8)$$

and relations obtained by circular permutations of the variables  $(x, y, z)$ . We note that in  $R^3$  or any space of odd dimension

$$*^2 = id, \quad *^{-1} = *. \quad (F.9)$$

For a two-dimensional space, which is useful in computations in waveguides, the star operator, designated by  $\perp$  (a two-branch star) is

$$\perp: 1 \rightarrow dx \, dy \rightarrow 1, \quad dx \rightarrow dy \rightarrow -dx \quad (F.10)$$

for orthonormal coordinates  $x$  and  $y$ , and the orientation defined by the order  $dx \, dy$ .

We note that  $\perp$  is not reciprocal:

$$\perp^2 = (-)^p \quad \text{on } p\text{-forms.} \quad (F.11)$$

In Minkowski space-time  $R^{3+1}$ , a vector being

$$V = X\partial_x + Y\partial_y + Z\partial_z + T\partial_\tau \quad (F.12)$$

the metric (not positive) can be defined by

$$\bar{O}V = \bar{V} = X \, dx + Y \, dy + Z \, dz - T \, d\tau. \quad (F.13)$$

Thus

$$\bar{V}|V = X^2 + Y^2 + Z^2 - T^2. \quad (F.14)$$

The following table gives the star operator  $*$

$$\star \left\{ \begin{array}{l} 1 \rightarrow d\tau \, dx \, dy \, dz = -dx \, dy \, dz \, d\tau \rightarrow -1 \\ dx \rightarrow -dy \, dz \, d\tau \rightarrow dx \\ dy \, dz \rightarrow -dx \, d\tau \rightarrow -dy \, dz \\ dx \, dy \, dz \rightarrow -d\tau \rightarrow dx \, dy \, dz \end{array} \right. \quad (F.15)$$

and relations obtained by circular permutations of the variables  $(x, y, z)$ .

Note that the set of ordered subscripts is not based on the natural order (any more than it was in  $R^3$ ). Rather, we choose to make it invariant under circular permutation of the space variables, with a special role played by the time variable.

Differential forms are often decomposed into space and time components. They may be purely spatial, say  $\alpha$  that does not contain  $d\tau$ , or mixed, in which case  $d\tau$  can be factored, yielding an expression  $(d\tau) \alpha$  where  $\alpha$  is purely spatial. It is convenient when transforming equations from four to three dimensions (see (F.15)) to make use of the following properties:

$$\star \left\{ \begin{array}{l} \alpha \rightarrow -(*\alpha) \, d\tau \\ (d\tau)\alpha \rightarrow -*\alpha. \end{array} \right. \quad (F.16)$$

Given two forms  $\alpha$  and  $\beta$  of the same degree, on a compact manifold  $M$  of dimension  $n$ , the expression  $\alpha(*\beta) = \alpha * \beta$  is an  $n$ -form which can be integrated over  $M$ , producing an (integrated) scalar product  $(\alpha * \beta)|M$ , which we shall denote

$$\langle \alpha * \beta \rangle. \quad (F.17)$$

Since  $\beta * \alpha = \alpha * \beta$  for  $\alpha, \beta$  of the same degree, this scalar product is symmetric.

Let  $L$  be an operator of degree  $q$ , i.e., one that maps  $p$ -forms onto  $(p+q)$ -forms. If  $\alpha$  is a  $p$ -form and  $\beta$  a  $(p-q)$ -form, the adjoint  $L^*$  of  $L$  is defined by

$$\langle \alpha * (L\beta) \rangle = \langle (L^*\alpha) * \beta \rangle. \quad (F.18)$$

Naturally, the degree of  $L^*$  is  $-q$ .

Taking for  $L$  the differential operator  $d$  (of degree  $+1$ , to be defined in Appendix H), its adjoint  $d^*$  (of degree  $-1$ ) is called the *codifferential* operator. Its explicit expression is

$$d^* = *^{-1} \, d \, * (-)^p \quad (F.19)$$

when applied to  $p$ -forms. (This operator is often denoted by  $\delta$  in the literature.) An obvious property, shared with the exterior differential is

$$d^* \cdot d^* = 0 \quad (F.20)$$

The following combination

$$-\Delta = d \cdot d^* + d^* \cdot d \quad (F.21)$$

is an operator of degree zero which equals the negative of the Laplacian.

### G. Vector Fields and Differential Forms

Up to this point we have considered a single vector space  $E$  (or its dual  $E^*$ ) and the exterior algebras  $\Lambda E$  and  $\Lambda E^*$  constructed on these spaces. Now we shall deal with manifolds such as the surface  $S$  of a sphere in  $R^3$ . These are not vector spaces, but we can attach to every point  $r$ , of  $S$  for instance, the tangent space denoted by  $T_r S$  (tangent at  $r$ , to  $S$ ). This produces a collection of vector spaces which is called the *tangent bundle*  $TS$  to  $S$ . This concept is beginning to play an important role in theoretical physics. It extends immediately to a manifold  $M$  of any dimension, provided  $M$  is embedded in a linear vector space. When this is not the case, one must seek an intrinsic definition of vector on a manifold.

Roughly speaking, a manifold  $M$  is a topological space that locally "looks like" an Euclidean space. More precisely [1, pp. 49, ff.],  $M$  can be covered by a collection of open sets  $U_i$ , each of which can be mapped continuously by a one-to-one function  $\phi_i$  into an Euclidean space  $R^n$  ( $n$  is the dimension of  $M$ ,  $\phi_i^{-1}$  is a *parameterization* of  $U_i$ ). The collection of maps or *charts*  $(U_i, \phi_i)$  is said to constitute an *atlas*. A compatibility condition must hold on the overlap  $U_i \cap U_j$  of two sets: the function  $\phi_i \cdot \phi_j^{-1}$ , defined on  $\phi_j(U_i \cap U_j)$  must be  $C^\infty$ , i.e., have continuous derivatives of all orders. The manifold is called *differentiable* and is the only type considered here. The compatibility condition enables one to carry on globally the operations of differential geometry. The need for several maps to define a manifold is clear from the example of the sphere, which obviously cannot be mapped in its entirety on a plane by a one-to-one function.

A tangent vector  $V_r \in T_r M$ , when the manifold is embedded into a linear space, may be considered as the velocity vector, at time  $t=0$ , for a point describing a trajectory  $\gamma(t)$  passing through  $r$  at time, 0, ( $\gamma(0) = r$ ). To  $\gamma(t)$  is associated a *directional derivative* which indicates the rate of variation at time 0 of a function defined on  $M$  as the observation point moves along the trajectory

$$\partial_{t/0} f(\gamma(t)). \quad (G.1)$$



(The notation  $\partial_{t/0} h(t)$  is short for  $(dh/dt)_{t=0}$ .) It can be shown that this rate of change depends only on the vector  $V_r$ . Consequently, we shall denote this rate by  $V_r f$ . Furthermore, when the manifold is not embedded, we shall consider that the directional derivative itself defines a vector on the manifold.

The assignment of a vector  $V_r$  to every point  $r$  of  $M$  defines a *vector field* over  $M$ . Any vector field is identified with a linear differential operator that transforms  $\mathfrak{F}_0$  into  $\mathfrak{F}_0$ , where  $\mathfrak{F}_0$  is the set of functions over  $M$ . For example, on a manifold parametrized by  $(u, v, w)$  the operators  $\partial_u, \partial_v, \partial_w$  are three vector fields.

To apply the results from the previous section, one can replace  $(e_1, e_2, e_3)$  by  $(\partial_u, \partial_v, \partial_w)$ . ( $e_i$  by  $\partial_i$  on a manifold of dimension  $n$ .)

The dual  $T_r^*M$  of  $T_r M$  is composed of *tangent* covectors, also called *cotangent vectors*. The collection of  $T_r^*M$  for all  $r$  is the *cotangent bundle* on  $M$ :  $T^*M$ . The basis, dual of the  $\{e_i\} = \{\partial_i\}$ , is  $\{\epsilon^i\}$  and the  $\epsilon^i$  are precisely the differential of the variables  $u^i$ —they are denoted by  $du^i$ . This definition might have seemed strange to Leibnitz himself! The  $du^i$  are *not increments* of the variable  $u^i$ , as is often stated in classical texts. From the modern point of view, the differential of coordinate  $u$ ,  $du$ , is a linear operator which applied to a vector

$$V = \delta u \partial_u + \delta v \partial_v + \delta w \partial_w \quad (G.2)$$

produces the  $u$  component of the vector  $V$

$$du|V = \delta u. \quad (G.3)$$

(The  $\delta u$  may be considered as small (infinitesimal) or finite.) The one-form

$$df = f_u du + f_v dv + f_w dw \quad (G.4)$$

applied to  $V$  gives

$$df|V = f_u \delta u + f_v \delta v + f_w \delta w. \quad (G.5)$$

The similarity with (G.4) explains the confusion that often exists between increment and differential.

Note that this last equation (G.5) is also an expression of  $Vf$ . Once the correspondence between the notations for bases

$$e_i \rightarrow \partial_i \quad \epsilon^i \rightarrow du^i$$

have been understood, all the results of preceding sections can be translated from a single vector space to the tangent and cotangent spaces at one point, and to fields of multivectors or multicovectors. The latter are known as *differential forms*.

## H. Exterior Differential

The differential of a scalar-valued function  $f$

$$f: R^n \rightarrow R: x = (x^1, x^2, \dots, x^n) \rightarrow f(x) \quad (H.1)$$

is the one-form

$$df = f_i dx^i \quad (H.2)$$

where  $f_i$  is the partial derivative of  $f$  with respect to  $x^i$ .

The differential of a  $p$ -form  $\alpha$

$$\alpha = \sum^0 a_J(x) dx^J, \quad J \in \mathfrak{J}_p \quad (H.3)$$

is by definition the  $(p+1)$ -form

$$d\alpha = \sum^0 da_J(x) dx^J \quad (H.4)$$

where  $da_J(x)$  is computed as in (H.2) and the expression (H.4) is then simplified according to the rule of exterior algebra. The main justification for introducing this operation is its role in expressing the Stokes' theorem (see Appendix J). We shall consider its application to functions and differential forms in space  $R^3$ , using  $(x, y, z)$  as coordinates instead of  $(x^1, x^2, x^3)$ .

For the function  $f(x, y, z)$ , equation (H.2) becomes

$$df = f_x dx + f_y dy + f_z dz. \quad (H.5)$$

If  $(x, y, z)$  are Euclidean coordinates, i.e., if the metric is defined by  $x^2 + y^2 + z^2$ , then the vector associated to  $df$  is

$$\overline{0} df = \overline{df} = f_x \partial_x + f_y \partial_y + f_z \partial_z \quad (H.6)$$

which is recognized as the *gradient* of  $f$ .

For a one-form

$$\alpha = X dx + Y dy + Z dz \quad (H.7)$$

its differential is the two form

$$d\alpha = (Z_y - Y_z) dy dz + (X_z - Z_x) dz dx + (Y_x - X_y) dx dy. \quad (H.8)$$

The coefficients are those of the *curl* of the vector  $\overline{\alpha} = \overline{0}\alpha$  associated with  $\alpha$ . Using the star operator, we convert  $d\alpha$  into a one-form:

$$*d\alpha = (Z_y - Y_z) dx + (X_z - Z_x) dy + (Y_x - X_y) dz. \quad (H.9)$$

and then transform this one-form into a vector  $\overline{*d\alpha}$ , obtaining the formula

$$\text{curl } \overline{\alpha} = \overline{*d\alpha}. \quad (H.10)$$

For example, if  $E$  is the electric field one-form and  $\vec{E}$  the electric field vector,  $\text{curl } \vec{E}$  and  $*dE$  are mates, i.e., are related by  $\overline{0}$ .

For a two-form

$$\beta = U dy dz + V dz dx + W dx dy \quad (H.11)$$

Its differential, after simplification, is the three-form

$$d\beta = (U_x + V_y + W_z) dx dy dz \quad (H.12)$$

whose coefficient is the *divergence* of the vector  $\overline{*}\beta = U\partial_x + V\partial_y + W\partial_z$  associated with  $\beta$ :

$$*d\beta = \text{div } \overline{*}\beta. \quad (H.13)$$

For example, the current two-form  $J$  and the current density vector  $\vec{J}$  are such that

$$\text{div } \vec{J} = *dJ. \quad (H.14)$$

Thus, we recognize the kinship of the exterior differentials with the operators curl, grad, and div. Their connection is complicated by the intervention of metrical concepts, associated with the latter and that are often spurious to the problem considered. Thus formulas (H.5), (H.8), and (H.12), which are independent of a metric, have the same expression if  $(x, y, z)$  are arbitrary curvilinear coordinates. The formulas for curl, grad, and div on the other hand are more complicated, as they involve the metrical coefficients  $g_{ij}$ . For instance, even in orthogonal coordinates  $(u, v, w)$ , where only the three Lamé coefficients  $(h_1, h_2, h_3)$  are needed, the expression for the curl is

as follows

$$\begin{aligned} \text{curl}(U\partial_u + V\partial_v + W\partial_w) &= \frac{1}{h_2 h_3} (\partial_v(h_3 w) - \partial_w(h_2 v)) \partial_u \\ &+ \frac{1}{h_3 h_1} (\partial_w(h_1 u) - \partial_u(h_3 w)) \partial_v \\ &+ \frac{1}{h_1 h_2} (\partial_u(h_2 v) - \partial_v(h_1 u)) \partial_w \end{aligned} \tag{H.15}$$

as compared to an expression such as (H.8) valid when  $(x, y, z)$  are replaced by arbitrary coordinates  $(u, v, w)$  and  $(X, Y, Z)$  by functions  $(U, V, W)$  of these coordinates. Many other examples could be cited where the introduction of an irrelevant metric only complicates formulas. The expression (H.15) is even more complex for arbitrary coordinates, and is unnecessary in many cases where the metric does not matter.

Some useful properties of the exterior differential are  $d(d\alpha) = 0$ , discussed in Appendix K, and the modified Leibnitz rule

$d(\alpha\beta) = (d\alpha)\beta + (-)^p \alpha(d\beta)$

(H.16)

where  $\alpha$  is a  $p$ -form. This equation encompasses the equations of vector calculus for  $\text{grad}(fg)$ ,  $\text{curl}(f\vec{v})$ ,  $\text{div}(f\vec{v})$ , and  $\text{div}(\vec{u} \times \vec{v})$ .

**I. Change of Variables—Pullbacks, Integration**

Differential forms are particularly well-suited to perform changes of variables. This is essentially because their expressions include the differentials of the variables  $x = (x^1, \dots, x^n)$ . Let the variables  $x$  be known in terms of new variables  $u = (u^1, \dots, u^m)$  through a set of  $n$  functions  $x^i = f^i(u^1, \dots, u^m)$  which define the mapping

$$f: R^m \rightarrow R^n: u \mapsto x \tag{I.1}$$

and let

$$\alpha = \sum^0 a_J(x) dx^J \tag{I.2}$$

be a  $p$ -form ( $J \in \mathcal{J}_p, |J| = p$ ).

The change of variables from  $x$  to  $u$  in  $\alpha$  is done by substituting  $f^i(u)$  for  $x^i$  both in the coefficients  $a_J$  and in the differentials  $dx^J \equiv dx^{j_1} \dots dx^{j_p}$ . The differential  $df^i(u)$  is computed as in (H.2):

$$df^i(u) = \sum \frac{\partial f^i}{\partial u^j} du^j. \tag{I.3}$$

The products in (I.2) are then simplified according to the rules of exterior algebra. The result is a  $p$ -form  $\alpha'$  which is called the *pullback* of  $\alpha$ , this relation<sup>3</sup> being denoted by

$$\alpha' = f^* \alpha. \tag{I.4}$$

<sup>3</sup>The form  $\alpha$  with variables  $x$  corresponds to the form  $\alpha'$  with variables  $u$  through (I.4). It is sometimes the practice to use the same symbol  $\alpha$  for the two forms, writing  $\alpha(x)$  for the first,  $\alpha(u)$  for the second, and thus relying on the variable to indicate which of the two forms is meant. This is a common "abuse of language" which leads to writing  $f(u)$  for  $f(x(u))$ . This, of course, strictly is incorrect and somewhat dangerous when other variables, or numerical values, are substituted in  $f(\cdot)$ .

(The star here is not to be confused with the conjugate sign or the star operation described in Appendix F.)

As an example, consider the electric field one-form in  $R^2$ :

$$E = X(x, y) dx + Y(x, y) dy \tag{I.5}$$

and "new" variables  $(\rho, \theta)$  related to  $(x, y)$  by

$$f: (\rho, \theta) \rightarrow (x, y) = (\rho \cos \theta, \rho \sin \theta). \tag{I.6}$$

Then

$$(dx, dy) = (\cos \theta d\rho - \rho \sin \theta d\theta, \sin \theta d\rho + \rho \cos \theta d\theta). \tag{I.7}$$

Substituting (I.7) in (I.5) and simplifying gives

$$f^*E = R(\rho, \theta) d\rho + \Theta(\rho, \theta) d\theta \tag{I.8}$$

with

$$R(\rho, \theta) = X'(\rho, \theta) \cos \theta + Y'(\rho, \theta) \sin \theta \tag{I.9}$$

$$\Theta(\rho, \theta) = -X'(\rho, \theta) \sin \theta + Y'(\rho, \theta) \cos \theta \tag{I.10}$$

$$X' = f^*X = X \cdot f \quad Y' = f^*Y = Y : f. \tag{I.11}$$

The name *pullback* given to this change of variable process is understood by the diagram representing the pullback of a function  $F$  (zero-form). The mapping  $f$  transforms a set  $U$  into a set  $V$ , and  $F$  maps  $V$  into  $R$ . The composition  $F \circ f$  maps  $U$  into  $R$  and is the pullback  $f^*F$ :

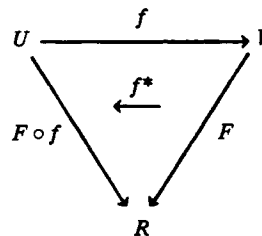


Fig. I.1. Pullback of a function  $F$ .

The following problem is left to the reader. Consider in vacuum the electromagnetic force field  $\Phi = E d\tau + B$  and the source field  $\Psi = D - H d\tau$ , where  $\tau = ct$ . Evaluate the pullbacks  $\Phi' = f^*\Phi$  and  $\Psi' = f^*\Psi$ , under the Lorentz transformation corresponding to a uniform motion with velocity  $c \cdot \text{th } \theta$  in the  $x$ -direction. This transformation is expressed by

$$f: (x', y', z', \tau') \rightarrow (x, y, z, \tau) \tag{I.12}$$

$$f: \begin{cases} x = \tau' \text{sh } \theta + x' \text{ch } \theta \\ y = y' \\ z = z' \\ \tau = \tau' \text{ch } \theta + x' \text{sh } \theta. \end{cases} \tag{I.13}$$

Verify that  $\Phi'$  and  $\Psi'$  are related by the same space-time star operator (defined in Appendix F) as  $\Phi$  and  $\Psi$ . In other terms, this operator is Lorentz invariant.

Even with little experience the reader will realize how easy and automatic these computations are. To quote Flanders on that subject [1, p. 25], "One can carry on fearlessly with the most obvious kind of calculations."

The following properties of the pullback can be verified:

$$f^*(\alpha + \beta) = f^*\alpha + f^*\beta \tag{I.14}$$

$$f^*(\alpha\beta) = (f^*\alpha)(f^*\beta) \tag{I.15}$$

$$f^*(d\alpha) = d(f^*\alpha) \tag{I.16}$$

$$(f \cdot g)^* = g^* \cdot f^* \tag{I.17}$$

The property (I.16) is particularly important, as it implies the invariance properties of equations expressed by means of exterior differentials.

An important application of pullbacks is the computation of integrals over a domain (or a chain) in an arbitrary manifold. If the domain  $D$  is parameterized by

$$\gamma: U \rightarrow D$$

where  $U$  is a Euclidean domain, the integral of the form  $\omega$  over  $D$  equals that of  $\gamma^*\omega$  over  $U$ .

Take, for instance, a curve in  $R^3$  defined by the parametrization  $\gamma$ , i.e., by a mapping from an interval  $I = [a, b]$  of  $R$  to the space  $R^3$ :

$$\gamma: I \rightarrow R^3: t \mapsto (x, y, z) = (f(t), g(t), h(t)). \tag{I.19}$$

The image  $\gamma(I) = \Gamma$  is oriented from  $\gamma(a)$  to  $\gamma(b)$ . Consider the problem of finding the work done by the electric field  $E = X dx + Y dy + Z dz$ , where  $X, Y, Z$  are functions of  $(x, y, z)$ , when a unit test charge is moved along  $\Gamma$ . The work  $W$  is the integral  $\int_{\Gamma} E$ , which we have denoted  $E|\Gamma$  to emphasize the linear dependence of the result on both factors  $E$  and  $\Gamma$ , and the fact that the integral is the limit of a Riemann sum of duality products  $E_i|\delta_i\gamma$  (see equation (4)).

This integral reduces to an integral over  $I$  by the pullback  $\gamma^*$ :

$$E|\Gamma = E|\gamma(I) = (\gamma^*E)|I. \tag{I.20}$$

The one-form  $\gamma^*E$  is a function of  $t$  and its integral is taken along an interval  $I$  of the real axis  $R$ .

Computations of multiple integrals can be handled similarly [1, p. 63].

*J. Stokes' Theorem*

Stokes' theorem is the single most important theorem of integral calculus. It does contain as special cases all known theorems that relate integrals over a domain and integrals over its boundary such as those due to Gauss, Poincaré, Ostrogradski, Green, and even Newton and Leibnitz! The statement of Stokes' theorem is particularly elegant in terms of differential forms: if  $\alpha$  is a  $p$ -form and  $D$  is a  $(p + 1)$ -domain of integration having  $\partial D$  as a boundary,

$$\boxed{\alpha|\partial D = d\alpha|D} \tag{J.1}$$

The definition of  $\partial D$  specifies the orientation of  $\partial D$  in relation to that of  $D$ . Both  $D$  and  $\partial D$  may be replaced by *chains* [1, p. 61].

In brief, a  $p$ -chain  $D$  on a manifold  $M$  is a linear combination  $\sum a_i \sigma_i$  of  $p$ -simplices. A  $p$ -simplex is a continuous map  $\sigma$  of an Euclidean  $p$ -simplex  $\Delta^p$  into  $M$ . A Euclidean  $p$ -simplex is defined by  $(p + 1)$  vertices  $[A_0 A_1 \cdots A_p]$  in  $R^p$  and consists of the centers of mass of those vertices with variable positive

weights that add up to 1. A simplex is oriented by selecting a particular order of the vertices. The boundary

$$\partial[A_0 A_1 \cdots A_p] = \sum (-)^i [A_0 \cdots \hat{A}_i \cdots A_p], \tag{I.2}$$

where  $\hat{A}_i$  means  $A_i$  is omitted, as extended by linearity and a continuous mapping onto the manifold. For precise definitions and discussion of these matters (see [1, pp. 61].)

*K. Poincaré Lemma*

The differential of a function  $f(x, y, z)$  is the one-form

$$df = f_x dx + f_y dy + f_z dz \tag{K.1}$$

where  $(f_x, f_y, f_z)$  are partial derivatives of  $f$  with respect to  $(x, y, z)$ . The exterior differential  $d(df)$  of this one-form is null because the mixed partial derivatives such as  $f_{xy}$  and  $f_{yx}$  are equal. This property is easily generalized to any  $p$ -form on a manifold of any dimension: the operation  $d$  applied twice in succession gives zero [1, pp. 2, 27].

$$d \circ d = 0 \tag{K.2}$$

(For two operators  $T$  and  $S$  the notation  $T \circ S$  represents their composition, i.e., the operation resulting from applying  $S$  first, then  $T$ .) Special cases of (K.2) in  $R^3$  correspond to the vector calculus properties  $\text{curl grad} = 0$  and  $\text{div curl} = 0$ .

The boundary operator  $\partial$ , applied to chains, enjoys a similar property:

$$\partial \circ \partial = 0. \tag{K.3}$$

This can be understood intuitively: the curve that bounds a surface has no end points, the surface that bounds a volume has no edge. The property is proven rigorously on a curved manifold after defining properly the boundary operator  $\partial$ , and the objects to which it applies, namely the *chains* which generalize integration domains. (See [1, pp. 61–63].)

By definition, a form  $d$  is *closed* if  $d\alpha = 0$ ; it is *exact* if there exists a form  $\beta$  such that  $d\beta = \alpha$ . Similarly, a chain  $C$  is a *cycle* if  $\partial C = 0$ , and it is a *boundary* if there exists a chain  $S$  such that  $\partial S = C$ . With these nomenclatures, properties (K.2) and (K.3) are expressed by

$$\begin{cases} \text{"Every exact form is closed."} & \text{(K.2')} \\ \text{"Every boundary chain is a cycle."} & \text{(K.3')} \end{cases}$$

In these two statements the conclusion holds over a domain  $D$  if the premise holds over that domain.

Converse properties of (K.2) and (K.3) are

$$\begin{cases} \text{"Every closed form is exact."} & \text{(K.4)} \\ \text{"Every cycle is a boundary."} & \text{(K.5)} \end{cases}$$

These are true over a specified domain  $D$  provided  $D$  has appropriate topological properties. For instance,  $D$  may be the whole space  $R^n$ , the inside of a sphere, or a star-like domain, i.e., one such that a segment joining a point  $O$  in  $D$  to any other point in  $D$  is entirely inside the domain  $D$ .

Taking  $O$  as an origin, the transformation

$$h_t: r \rightarrow tr, \quad \text{for } t \in [0, 1] \tag{K.6}$$

maps  $D$  continuously into  $D$  such that  $h_1$  is the identity and  $h_0$  maps every point in  $D$  onto  $O$ . It defines a *homotopy* from the identity map  $h_1$  to the map  $h_0$ . (Other homotopies  $h_t: r \rightarrow h_t r$

sharing the above properties could be used to derive a solution of (K.4) and (K.5).)

The importance of this homotopy is that it makes it possible to construct a solution of (K.4) and (K.5). For (K.5), given a cycle  $C$ , the surface  $S$  of the cone with apex  $O$  and base  $C$  (generated by  $h_t(C)$  as  $t$  varies from 1 to 0), is such that  $\partial S = C$ .

For (K.4), one is given a  $p$ -form  $\alpha$ , closed over  $D$ . This form is exact if there exists a  $(p - 1)$ -form  $\beta$  such that  $d\beta = \alpha$ . The actual construction of  $\beta$  results essentially from integrating along segments from  $O$  to the point where  $\beta$  is evaluated. See [1, pp. 27-31] for such a construction. The property (K.4) is usually known as the Poincaré lemma (or theorem), although some authors (Flanders in particular) take (K.2) as the Poincaré lemma and call (K.4) its converse.

Note that the solution  $\beta$  is not unique. It is obvious that  $\beta$  can be augmented by any closed form or by  $d\gamma$ , for any form  $\gamma$ , without changing its differential. Conversely, two solutions of (K.5),  $\beta_1$  and  $\beta_2$ , are such that  $d(\beta_2 - \beta_1) = 0$ ; hence,  $\beta_2 - \beta_1 = d\gamma$  for some  $(p - 2)$ -form  $\gamma$ , provided the conditions under which the lemma applies are satisfied.

*L. Flow and Lie Derivatives*

The steady motion of a fluid in space  $R^3$  may be described by a vector-field  $V(r)$  which gives at every point  $r \in R^3$  the velocity of the fluid particle passing through that point. An integral curve of the vectorfield  $V$  is a curve

$$\gamma: R \rightarrow R^3: t \mapsto \gamma(t) \tag{L.1}$$

such that

$$\partial_t \gamma(t) = V(\gamma(t)). \tag{L.2}$$

The space over which the motion takes place may be a manifold  $M$ , its points representing the states of a system whose evolution is described by an integral curve, or trajectory in  $M$ .

If  $V(r)$  is smooth enough, it is possible to find for every point  $r$  an integral curve  $\gamma(t, r)$  such that

$$\gamma(\cdot, r): I \rightarrow R^3: t \mapsto \gamma(t, r): O \mapsto r \tag{L.3}$$

where  $I$  is an interval in  $R$ , that contains  $t = 0$ . (The interval  $I$  may depend on  $r$ .)

The mapping

$$\gamma(t, \cdot): R^3 \rightarrow R^3: r \mapsto \gamma(t, r) \tag{L.4}$$

defines the flow associated with the vectorfield  $V$ . It will be denoted by  $V^t$ .

Thus,  $V^t(r) = r_t$  is the position at time  $t$  of the fluid particle that was at point  $r$  at time 0. It is easy to see that

$$V^t \circ V^s = V^{t+s} \tag{L.5}$$

and

$$V^0 = id \tag{L.6}$$

which justifies the exponential notation and shows that the transformations  $V^t$  form a one-parameter Abelian group if  $V^t$  is defined for all real values of  $t$ . (The transformation  $V^t$  defined only for positive values of  $t$  forms a semigroup.) Given a flow  $V^t$  over a manifold  $M$  (for instance  $R^3$ ) and a scalar function

$$f: M \rightarrow R: r \mapsto f(r) \tag{L.7}$$

one may consider the function

$$(V^t)^* f: R \rightarrow R: t \mapsto f(V^t(r)) \tag{L.8}$$

which is the pullback (see Appendix I) of function  $f$ . It does represent the function  $f$  observed at a point  $r_t = V^t(r)$  carried by the flow. The derivative of this function at time 0 is called the Lie derivative  $L_V f$  of the function  $f$ . We have encountered its expression as the direction derivative  $Vf = df|V$  (see Appendix G, equation (G.5)). More generally, however, it is possible to define a derivative along the flow for other than scalar-valued functions: any type of tensor, covariant, contravariant, or mixed, can be transported by the flow, hence compared at two different points of a trajectory. The difference between its values at  $t$  and  $t = 0$ , divided by the time  $t$  has a limit as  $t \rightarrow 0$  called the Lie derivative of the tensor. (This derivative should not be confused with the covariant derivative that requires the concept of parallel displacement related to a Riemannian metric or to a connection.)

For covariant tensors and in particular for exterior differential forms, which are of primary interest for our purpose, the transport is performed by a pullback. At point  $r$  on the manifold the Lie derivative of the form  $\alpha$  is the differential form

$$(L_V \alpha)_r = \partial_{t/0} (V^t)^* \alpha(V^t r). \tag{L.9}$$

If  $\alpha$  is a  $p$ -form, we can integrate this equation over a  $p$ -domain  $D$ , carried along by the flow to

$$D_t = V^t D. \tag{L.10}$$

This gives

$$\partial_{t/0} (\alpha|D_t) = (L_V \alpha)|D \tag{L.11}$$

expressing the rate of change at time  $t = 0$  of the integral of  $\alpha$  over a moving domain  $D_t$ .

A useful formula for  $L_V \alpha$  is

$$L_V \alpha = d(\alpha|V) + (d\alpha)|V \tag{L.12}$$

or in operator form

$$L_V = d \circ i_V + i_V \circ d. \tag{L.13}$$

With the help of Stokes' theorem, equations (L.11) and (L.12) yield

$$\partial_{t/0} (\alpha|D_t) = ((d\alpha)|V)|D + (\alpha|V)|\partial D. \tag{L.14}$$

This formula expresses and generalizes formulas of vector calculus relative to integration over domains of dimensions 1, 2, and 3. For example, in dimension 2, (L.14) corresponds to the vector flux theorem of Helmholtz [7, pp. 285-287].

ACKNOWLEDGMENT

The author wishes to thank R. W. Ziolkowski for his painstaking review of the manuscript and for many excellent suggestions.

## REFERENCES

- [1] H. Flanders, *Differential Forms with Applications to the Physical Sciences*. New York: Academic Press, 1963.
- [2] V. I. Arnold, *Mathematical Methods of Classical Mechanics*. New York: Springer Graduate Texts in Math, 1978.
- [3] C. W. Misner and J. A. Wheeler, "Classical geometrodynamics," *Ann. Phys.*, vol. 2, pp. 525-603, Dec. 1957.
- [4] J. A. Wheeler, *Geometrodynamics*. New York: Academic Press, 1962.
- [5] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. San Francisco, CA: Freeman, 1970.
- [6] G. A. Deschamps *et al.*, *Mathematics Applied to Physics*, ch. III, "Exterior differential forms." New York: Springer-Verlag, 1970.
- [7] A. Sommerfeld, *Electrodynamics*. New York: Academic Press, 1952.
- [8] D. van Dantzig, "Electromagnetism, independent of metrical geometry," *Proc. Cambridge Phil. Soc.*, vol. 37, pp. 521-531, 1934.
- [9] G. deRham, *Variétés Différentiables*. Paris, France, Hermann, 1955.
- [10] Y. Aharonov and D. Bohm, "Significance of electromagnetic potential in the quantum theory," *Phys. Rev.*, vol. 115, pp. 485-491, 1959.
- [11] P. A. M. Dirac, "Quantised singularities in the electromagnetic field," *Proc., Roy. Soc.*, vol. A133, pp. 60-72, 1931.
- [12] —, "The theory of magnetic poles," *Phys. Rev.*, vol. 74, p. 817, 1949.
- [13] T. T. Wu and C. N. Yang, "Concept of nonintegrable phase factors and global formulation of gauge fields," *Phys. Rev. D*, vol. 12, pp. 3865-3857, Dec. 1975.
- [14] A. Rubinowicz, "The Miyamoto-Wolf diffraction wave," *Progr. Phys.*, vol. IV, E. Wolf, Ed. Amsterdam, The Netherlands: North Holland, 1965, pp. 201-240.

Two recently published books, that deal with Electromagnetics from the viewpoint of exterior (or Cartan's) calculus came to the attention of the author after this paper had been prepared:

- [15] K. Meetz and W. L. Engl, *Elektromagnetische Felder*. Heidelberg, Germany: Springer-Verlag, 1980.
- [16] W. Thirring, *Classical Field Theory*. New York: Springer-Verlag, 1979.