On some mathematical aspects of deterministic classical electrodynamics

B. W. Stuck

Bell Laboratories, Murray Hill, New Jersey 07974
(Received 14 November 1973)

A new commutative diagram summarizing some of the mathematical structure of deterministic classical electrodynamics is presented. The diagram clearly delineates the fundamentally different roles played by the space–time differentiable manifold (*vis à vis* exterior calculus) and by matter or vacuum (*vis à vis* the constitutive relations for dielectric permittivity and magnetic permeability) in electrodynamics. Two different elliptic operators, called here the Laplace–Beltrami and Laplace–Poisson operators, arise naturally from this formulation. Some properties of eigenfunctions of elliptic operators with compact support are briefly reviewed with regard to potential application in numerical analysis of practical problems in electrodynamics. The action of the so-called inhomogeneous Lorentz group on electrodynamical functions is described. Several scalar inner products which remain invariant under the action of this group are seen to arise naturally from the mathematical structure discussed here. By using some of these invariant quantities, a new variational approach to deterministic classical electrodynamics is then developed. First, a new Lagrangian function is presented and used to derive the Euler–Lagrange equations for electrodynamics. Second, a series of new Hamiltonian functions are presented and used to derive the Hamiltonian equations for electrodynamics. All results are illustrated by a detailed examination of the electrodynamical structure of a model for an inhomogeneous nonisotropic medium.

I. INTRODUCTION

Although an extensive literature exists on mathematical aspects of deterministic classical electrodynamics, there is apparently no clear rigorous exposition on the relationship of exterior calculus and differential forms to Maxwell's equations. This is somewhat surprising, because exterior calculus would hopefully clarify some of the mathematical structure underlying electrodynamics, while offering a different (more formal but less physical) view of the structure than that of conventional or classical vector calculus. Such work hopefully would continue the interaction of mathematics and physics, which has been so fruitful in the past. Finally, the tremendous technological importance of electrodynamics lends added interest to such work.

This paper attempts to fill some of this gap in the literature. The scope is limited to a particular class of models for an inhomogeneous nonisotropic medium. Within this framework, a number of novel and well-known results are obtained more easily and naturally than by methods based on conventional vector calculus.

The goal here is to unify and simplify certain mathematical aspects of electromagnetism. An example of a successful attempt along similar lines can be found in modern communication and control theory, which have been greatly unified through the concept of state and state variable techniques. It is hoped the approach discussed here will find application in other branches of physics, just as state variables have found wide application (e.g., in electrical network theory and in control system theory). This hope must be tempered by the following observation: Many practical problems can be adequately modeled by a set of first order ordinary differential equations, where the state space is a finite-dimensional vector space. The analogous state space for a distributed parameter system, such as is discussed here, is a finite-dimensional differentiable manifold and the vector fields associated with the manifold. The dimensionality of the state space for lumped parameter systems can be anything in practice; the underlying manifold for distributed parameter systems might be only four-dimensional, three spatial and one temporal, in practice. This suggests that future work should be directed toward a better understanding of the peculiarities of the four-dimensional case, as well as toward generalizations in higher dimensions.

Although interesting in their own right, the results presented here are interesting from a purely pedagogical point of view as well. One need only known the operations or rules of exterior algebra, as well as how to compute the total differential of a function; then the calculation of gradient, curl, and divergence become routine formal manipulations, but unfortunately often devoid of physical insight into the nature of the calculation. The conventional or classical vector calculus approach, with its line integrals and pillboxes, compliments this method by offering great physical insight into the nature of the calculation, but often at the expense of algebraic complexity in computing the correct answer. Both approaches have their merits and disadvantages, offering different views on the same situation.

The initial motivation for this work is found in Flanders.[1] While it was felt his approach was basically sound, it seemed sketchy at points and could be considerably more detailed. Another impetus is found in Dyson,[2] who has observed that the foundations of exterior calculus were laid by Grassmann in the mid-nineteenth century, but the tools he developed were discarded when the mathematical structure of electrodynamics was considered, in favor of tools developed to describe the structure of Lie groups and Lie algebras.

An appendix is included sketching and illustrating the basic concepts of multilinear algebra, differentiable manifolds, and exterior calculus. The reader familiar with these topics can proceed directly to the main body of the paper; otherwise, this detour is advised.

The second section presents a commutative diagram which summarizes the electrodynamical equations for an inhomogeneous nonisotropic medium; this is analogous to a block diagram or signal flow graph in control and communication theory.

The third section discusses two different differential operators which arise from this formulation of electrodynamics; previous work has tended to ignore or obscure this point. Various properties of the spectrum and eigenfunctions of these operators are briefly reviewed, for the case where the operators are elliptic and compactly supported.

The fourth section dwells on a group of coordinate transformations which preserve the structure of the equations of electrodynamics. Several well-known and new scalar inner products are seen to arise naturally from the approach discussed here.

The fifth section develops an alternate calculus-of-variations approach to the mathematical structure of electrodynamics. A new Lagrangian function is discussed, and all of the electrodynamical equations are derived from it. A new series of Hamiltonian functions are derived from Legendre transformations on the Lagrangian, and all of the electrodynamical equations are rederived.

All these results are illustrated by examining again and again a model for an inhomogeneous nonisotropic medium.

II. A COMMUTATIVE DIAGRAM

Throughout this section, X is an oriented Riemannian differentiable manifold called space—time.[3] T^*_X denotes the cotangent bundle associated with X, and $\Lambda(T^*_X) = \sum_{k=0}^{4} \Lambda^k(T^*_X)$ the associated exterior algebra. This section is broken into two parts: first, a diagram is presented which summarizes Maxwell's equations for a particular class of inhomogeneous nonisotropic media (in effect, the equations can be read off with the aid of this diagram); second, an example is presented to illustrate more clearly this result.

Theorem (Classical electrodynamics—Maxwell's equations): If $M:\Lambda^2(T^*_X) \to \Lambda^2(T^*_X)$ is a smooth function of $\Lambda^2(T^*_X)$, and is invertible at every point of the manifold X, then the diagram shown below commutes

$$\begin{array}{ccccccccc}
\Lambda^0 & \xrightarrow{d} & \Lambda^1 & \xrightarrow{d} & \Lambda^2 & \xrightarrow{d} & \Lambda^3 & \xrightarrow{d} & \Lambda^4 \\
{\scriptstyle d^2\circ M\circ d^2}\downarrow & & \downarrow{\scriptstyle d\circ M\circ d} & & {\scriptstyle M}\downarrow\ \uparrow{\scriptstyle M^{-1}} & & \uparrow{\scriptstyle d\circ M^{-1}\circ d} & & \uparrow{\scriptstyle d^2\circ M^{-1}\circ d^2} \\
\Lambda^4 & \xleftarrow{d} & \Lambda^3 & \xleftarrow{d} & \Lambda^2 & \xleftarrow{d} & \Lambda^1 & \xleftarrow{d} & \Lambda^0
\end{array}$$

where d is the exterior derivative.

Proof: The proof proceeds in three steps: (i) all operations shown above must be well defined on all charts of the manifold, i.e., locally; (ii) all operations must be capable of being pieced together smoothly on overlapping charts; (iii) the diagram must commute, i.e., be independent of path. Since the exterior derivative and the linear transformation M are well defined, all operations shown in the diagram are valid on each chart of the manifold. On overlapping coordinate charts, the transition functions associated with these charts can be used to smoothly piece together the operators M, M^{-1}, $d\circ M\circ d$, $d\circ M^{-1}\circ d$, $d^2\circ M\circ d^2$, $d^2\circ M^{-1}\circ d^2$. Finally, the diagram commutes, because of the two preceding steps.

Remark: Since $d^2=0$, the maps from $\Lambda^0 \to \Lambda^4$ are trivial.

Example: Choose rectangular orthonormal basis vectors $\{dx, dy, dz, icdt\}$ for X, and orientation $+dx\wedge dy \wedge dz \wedge icdt$. ($c$ is the speed of light.) The physical nature of each differential form is well known:

(A) Λ^0—g_e, g_m—electric, magnetic gauge.

(B) Λ^1—(A_{ex}, A_{ey}, A_{ez}), (A_{mx}, A_{my}, A_{mz})—electric, magnetic vector potential; φ_e, φ_m—electrical, magnetic scalar potential.

(C) Λ^2—(D_x, D_y, D_z), (B_x, B_y, B_z)—electric displacement, magnetic flux; (E_x, E_y, E_z), (H_x, H_y, H_z)—electric, magnetic fields.

(D) Λ^3—(J_{mx}, J_{my}, J_{mz}), (J_{ex}, J_{ey}, J_{ez})—magnetic, electric current densities; ρ_m, ρ_e—magnetic, electric charge densities.

(E) Λ^4—s_m, s_e—magnetic, electric source.

The question arises of how to associate which differential form with which electromagnetic function. The choice adopted here offers a certain amount of physical appeal, and is self consistent and complete with respect to exterior calculus.

Since X is a four-dimensional differentiable manifold, the differential forms may be interpreted intuitively as follows:

(1) Λ^0—gauge—scalar functions of space—time.

(2) Λ^1—potentials—directed line elements or 1-volumes in space—time.

(3) Λ^2—fields—directed areas or 2-volumes in space—time, in part directed along purely spatial directions ($dy\wedge dz$, $dz\wedge dx$, $dx\wedge dy$) and in part directed along a mixture of space—time directions ($dx\wedge icdt$, $dy\wedge icdt$, $dz\wedge icdt$).

(4) Λ^3—current densities—directed shells or 3-volumes in space—time in part directed along a purely spatial direction ($dx\wedge dy\wedge dz$) and in part directed along a mixture of space—time directions ($dy\wedge dz\wedge icdt$, $dz\wedge dx\wedge icdt$, $dx\wedge dy\wedge icdt$).

(5) Λ^4—sources—directed volumes or 4-volumes in space—time.

It is interesting to give a physical interpretation to the commutative diagram, much as in control and communication theory problems one gives a physical interpretation to a block diagram. Suppose, for example, a 1-form or potential is known at every point in X. The exterior derivative of this potential specifies a 2-form or field at every point; applying the constitutive relations plus the exterior derivative to the field specifies a 3-form or current density, which in turn feeds back to modify the potentials, and so on.

In this choice of coordinates, Maxwell's equations can be written using exterior calculus as:

(1) $d(g_e + ig_m) = \frac{\partial}{\partial x}(g_e + ig_m)\,dx \quad \frac{\partial}{\partial y}(g_e + ig_m)\,dy$

$$+\frac{\partial}{\partial z}(g_e+ig_m)\,dz+\frac{\partial}{ic\,\partial t}(g_e+ig_m)\,icdt$$
$$=c(A_{ex}+iA_{mx})\,dx+c(A_{ey}+iA_{my})\,dy$$
$$+c(A_{ez}+iA_{mz})\,dz+(\varphi_e+i\varphi_m)\,icdt,$$

(2) $$d[c(A_{ex}+iA_{mx})\,dx+c(A_{ey}+iA_{my})\,dy+c(A_{ez}+iA_{mz})\,dz+(\varphi_e+i\varphi_m)\,icdt]$$
$$=\left(\frac{\partial}{\partial y}c(A_{ez}+iA_{mz})-\frac{\partial}{\partial z}c(A_{ey}+iA_{my})\right)dy\wedge dz$$
$$+\left(\frac{\partial}{\partial z}c(A_{ex}+iA_{mx})-\frac{\partial}{\partial x}c(A_{ez}+iA_{mz})\right)dz\wedge dx$$
$$+\left(\frac{\partial}{\partial x}c(A_{ey}+iA_{my})-\frac{\partial}{\partial y}c(A_{ex}+iA_{mx})\right)dx\wedge dy$$
$$+\left(\frac{\partial}{\partial x}(\varphi_e+i\varphi_m)-\frac{\partial}{ic\,\partial t}c(A_{ex}+iA_{mx})\right)dx\wedge icdt$$
$$+\left(\frac{\partial}{\partial y}(\varphi_e+i\varphi_m)-\frac{\partial}{ic\,\partial t}c(A_{ey}+iA_{my})\right)dy\wedge icdt$$
$$+\left(\frac{\partial}{\partial z}(\varphi_e+i\varphi_m)-\frac{\partial}{ic\,\partial t}c(A_{ez}+iA_{mz})\right)dz\wedge icdt$$
$$=c(D_x+iB_x)\,dy\wedge dz+c(D_y+iB_y)\,dz\wedge dx$$
$$+c(D_z+iB_z)\,dx\wedge dy+(E_x+iH_x)\,dx\wedge icdt$$
$$+(E_y+iH_y)\,dy\wedge icdt+(E_z+iH_z)\,dz\wedge icdt,$$

(3) $$d[c(D_x+iB_x)\,dy\wedge dz+c(D_y+iB_y)\,dz\wedge dx+c(D_z+iB_z)\,dx\wedge dy+(E_x+iH_x)\,dx\wedge icdt+(E_y+iH_y)\,dy\wedge icdt+(E_z+iH_z)\,dz\wedge icdt]$$
$$=\left(\frac{\partial}{\partial y}(E_z+iH_z)-\frac{\partial}{\partial z}(E_y+iH_y)+\frac{\partial}{ic\,\partial t}c(D_x+iB_x)\right)$$
$$\times dy\wedge dz\wedge icdt+\left(\frac{\partial}{\partial z}(E_x+iH_x)-\frac{\partial}{\partial x}(E_z+iH_z)\right.$$
$$\left.+\frac{\partial}{ic\,\partial t}c(D_y+iB_y)\right)dz\wedge dx\wedge icdt+\left(\frac{\partial}{\partial x}(E_y+iH_y)\right.$$
$$\left.-\frac{\partial}{\partial y}(E_x+iH_x)+\frac{\partial}{ic\,\partial t}c(D_z+iB_z)\right)dx\wedge dy\wedge icdt$$
$$+\left(\frac{\partial}{\partial x}c(D_x+iB_x)+\frac{\partial}{\partial y}c(D_y+iB_y)+\frac{\partial}{\partial z}c(D_z+iB_z)\right)$$
$$\times dx\wedge dy\wedge dz$$
$$=(J_{mx}+iJ_{ex})\,dy\wedge dz\wedge icdt+(J_{my}+iJ_{ey})\,dz\wedge dx\wedge icdt$$
$$+(J_{mz}+iJ_{ez})\,dx\wedge dy\wedge icdt+c(\rho_e+i\rho_m)\,dx\wedge dy\wedge dz,$$

(4) $$d[(J_{mx}+iJ_{ex})\,dy\wedge dz\wedge icdt+(J_{my}+iJ_{ey})\,dz\wedge dx\wedge icdt+(J_{mz}+iJ_{ez})\,dx\wedge dy\wedge icdt+c(\rho_e+i\rho_m)\,dx\wedge dy\wedge dz]$$
$$=\left(\frac{\partial}{\partial x}(J_{mx}+iJ_{ex})+\frac{\partial}{\partial y}(J_{my}+iJ_{ey})+\frac{\partial}{\partial z}(J_{mz}+iJ_{ez})\right.$$
$$\left.-\frac{\partial}{ic\,\partial t}c(\rho_e+i\rho_m)\right)dx\wedge dy\wedge dz\wedge icdt$$
$$=(s_m+is_e)\,dx\wedge dy\wedge dz\wedge icdt.$$

The classical electrodynamics equations are found by equating real and imaginary parts of (1)–(4). The sign on φ_e and ρ_m must be reversed to conform to that standard in physics.[4] It is assumed here the transformation M can be written in matrix form for an inhomogeneous nonisotropic medium as

$$\begin{bmatrix} cD_x & dy\wedge dz\\ cD_y & dz\wedge dx\\ cD_z & dx\wedge dy\\ E_x & dx\wedge icdt\\ E_y & dy\wedge icdt\\ E_z & dz\wedge icdt\end{bmatrix}=\{p\circ\tilde M_E\circ*\circ+(1-p)\circ*\circ\tilde M_E\circ\}\begin{bmatrix} cD_x & dy\wedge dz\\ cD_y & dz\wedge dx\\ cD_z & dx\wedge dy\\ E_x & dx\wedge icdt\\ E_y & dy\wedge icdt\\ E_z & dz\wedge icdt\end{bmatrix}$$

$$\begin{bmatrix} cB_x & dy\wedge dz\\ cB_y & dz\wedge dx\\ cB_z & dx\wedge dy\\ H_x & dx\wedge icdt\\ H_y & dy\wedge icdt\\ H_z & dy\wedge icdt\end{bmatrix}=\{p\circ\tilde M_H\circ*\circ+(1-p)\circ*\circ\tilde M_H\circ\}\begin{bmatrix} cB_x & dy\wedge dz\\ cB_y & dz\wedge dx\\ cB_z & dx\wedge dy\\ H_x & dx\wedge icdt\\ H_y & dy\wedge icdt\\ H_z & dz\wedge icdt\end{bmatrix},$$

where $0\le p\le 1$. In this example, M is assumed to be a convex combination of the star operator, $\tilde M_E$ and $\tilde M_H$, where

$$\tilde M_E=\begin{bmatrix}0 & c\epsilon\\ c^{-1}\epsilon^{-1} & 0\end{bmatrix},\quad \tilde M_H=\begin{bmatrix}0 & c\mu\\ c^{-1}\mu^{-1} & 0\end{bmatrix},$$

where μ, ϵ are 3×3 matrices, 0 is the all zero 3×3 matrix. μ is called magnetic permeability, while ϵ is dielectric permittivity; the units are meter·kilogram·second.

In other treatments[5,6] a different set of units are often used: In these units the dielectric permittivity ϵ and magnetic permeability μ are rescaled, and it is frequently states that (in these units) $E_x=D_x$, $H_x=B_x$, and so forth. Strictly speaking, these equalities are quite ill-defined because the electric field (E_x,E_y,E_z) and electric displacement (D_x,D_y,D_z) lie in orthogonal subspaces of Λ^2, as does the magnetic field (H_x,H_y,H_z) and magnetic flux (B_x,B_y,B_z). To emphasize this often ignored fact, meter·kilogram·second units have been adopted.

In order to model the inhomogeneity of the medium, matrix elements in ϵ and μ are smooth functions of x,y,z and ict. To account for the anisotropy of the medium, ϵ and μ are assumed not to be similar to scalar multiples of the identity matrix.

Clearly, this choice of assumed constitutive relations for $\tilde M$ is not the only one that can model an inhomogeneous nonisotropic medium: The only essential assumption is that $\tilde M$ must be invertible on its support, X. The example here was chosen as illustrative of linear constitutive relationships; it can be generalized in any number of ways. For example, the next section shows $\Lambda^2(X)$ can be considered as a Hilbert space, the space of all functions in $L^2(X)$; $\tilde M$ may now be defined as an invertible operator defined on Hilbert space. Other generalizations are possible.

III. ELLIPTIC OPERATORS

Two differential operators are seen to arise naturally from this formulation of electrodynamics, the Laplace–Beltrami operator and the Laplace–Poisson operator.

The Laplace–Beltrami operator $\Delta=d\delta+\delta d$, where

$\delta = {*}\circ d\circ{*}$, is elliptic (Warner, Ref. 7, pp. 250–251), and

$$\Delta: \Lambda^K \to \Lambda^K, \quad K=0,1,2,3,4.$$

Since Δ depends only on the underlying manifold X, a picturesque description of Δ is that it is totally geometric or topological in nature. If X is a compact manifold, then the Hodge decomposition theorem shows that any differential form $u_p \in \Lambda^p$ ($p=0,1,2,3,4$) can be written as the sum of an exact form, a coexact form, and a harmonic form which lies in the finite-dimensional kernel of Δ,

$$u_p = u_p{}^H \oplus du_{p-1}{}^E \oplus \delta u_{p+1}{}^C, \quad p=0,1,2,3,4,$$

where the superscripts H, E, C denote harmonic, exact, and coexact, respectively (Ref. 7, p. 223).

The Laplace–Poisson operator $d\circ M\circ d = d\circ M^{-1}\circ d$ depends partly on the underlying manifold X (via the exterior derivative d) and partly on the physics (embodied in $\tilde{M}$); this operator may be considered as partly geometric or topological and partly physical. In the special case which is of great practical interest where the Laplace–Poisson operator can be shown to be elliptic (e.g., constant permittivity ϵ and permeability μ, a homogeneous nonisotropic media) a great deal more can be ascertained. If X is compact, then any p-form may be written as the sum of a p-form lying in the finite-dimensional kernel of the operator, plus a term in the orthogonal complement of this vector space.

Since the Laplace–Beltrami operator is always elliptic, while the Laplace–Poisson operator is often elliptic, a brief review of some of the properties of the eigenfunctions and eigenvalues of elliptic operators is included. Let E be an elliptic operator whose support is on a compact manifold X; then it is well known that (Ref. 7, pp. 254–256)

(1) nontrivial eigenvalues and eigenfunctions of E exist,

(2) there are an infinite number of eigenfunctions,

(3) all eigenvalues are nonpositive,

(4) the eigenfunctions are complete in $L^2(X)$,

(5) any function in $L^2(X)$ can be uniformly approximated by a sequence of these eigenfunctions, on X,

(6) the eigenvalues have no finite accumulation point,

(7) the eigenspaces associated with each eigenvalue are finite-dimensional.

Example: From the Hodge decomposition theorem,

(A) $g_e + ig_m = f_0^H \oplus \delta f_1^C$,

(B) $c(A_{ex}+iA_{mx})\,dx + c(A_{ey}+iA_{my})\,dy + c(A_{ez}+iA_{mz})\,dz + (\varphi_e + i\varphi_m)\,icdt = f_1^H \oplus \delta f_0^E \oplus \delta f_2^C$,

(C) $c(D_x+iB_x)\,dy\wedge dz + c(D_y+iB_y)\,dz\wedge dx + c(D_z+iB_z)\,dx\wedge dy + (E_x+iH_x)\,dx\wedge icdt + (E_y+iH_y)\,dy\wedge icdt + (E_z+iH_z)\,dz\wedge icdt = f_2^H \oplus df_1^E \oplus \delta f_3^C$,

(D) $(J_{mx}+iJ_{ex})\,dy\wedge dz\wedge icdt + (J_{my}+iJ_{ey})\,dy\wedge dx\wedge icdt + (J_{mz}+iJ_{ez})\,dx\wedge dy\wedge icdt + c(\rho_e+i\rho_m)\,dx\wedge dy\wedge dz = f_3^H \oplus df_2^E \oplus \delta f_4^C$,

(E) $(s_m+is_e)\,dx\wedge dy\wedge dz\wedge icdt = f_4^H \oplus df_3^E$,

where $f_K^H, f_K^E, f_K^C \in \Lambda^K$ ($K=0,1,2,3,4$), and the superscripts H, E, C denote harmonic, exact, and coexact, respectively. If X is compact and simply connected, it can be shown that (Ref. 7, p. 158 and pp. 226–229)

$$f_0^H = \text{const},$$

$$f_1^H = 0, \quad f_2^H = 0, \quad f_3^H = 0,$$

$$f_4^H = (\text{const})\,dx\wedge dy\wedge dz\wedge icdt,$$

corresponding physically to a source-free region of space–time. The terms δf_K^C ($K=1,2,3,4$) and df_i^E ($i=0,1,2,3$) can be expressed as linear combinations of eigenfunctions of the Laplace–Beltrami operator. The exact and coexact forms are also called *Hertz vectors*.[4]

The Laplace–Poisson operator, since it is a different operator from the Laplace–Beltrami operator, will in general have different eigenfunctions. Note that any 3-form can be expressed as an infinite linear combination of these eigenfunctions denoted $\{\tilde{u}_k^3\}$, $k=1,2,\cdots$. Using the exterior derivative d, its adjoint δ, plus the Hodge star operator $*$, the following statements hold (recall the underlying manifold is four-dimensional, so $d\tilde{u}_k^3=0$) on a compact manifold:

(i) Any 0-form may be written as an infinite linear combination of

$\{{*}\circ d\circ\tilde{u}_k^3\}$, $k=1,2,\cdots$,

(ii) Any 1-form may be written as an infinite linear combination of

$\{{*}\circ\tilde{u}_k^3\}$, $k=1,2,\cdots$,

(iii) Any 2-form may be written as an infinite linear combination of

$\{\delta\tilde{u}_k^3\}$, $k=1,2,\cdots$,

(iv) Any 4-form may be written as an infinite linear combination of

$\{d\tilde{u}_k^3\}$, $k=1,2,\cdots$.

This finding may have practical application. In semiconductor device work, or in transmission of electromagnetic energy, Maxwell's equations plus real boundary conditions are often analytically intractable, and a numerical approximation to the true solution must often be used. One type of numerical approximation is to expand all functions as a sum of a finite number of orthonormal functions, and to truncate the sum when an error criteria is sufficiently small. The approach presented here makes it possible to choose from two different sets of orthonormal functions; under some circumstances, one set may be preferable to the other.

IV. SOME GROUP THEORETIC ASPECTS

In certain situations, a great deal of insight is gained by a change of coordinates. This section is concerned with a class of coordinate transformations which form a group, and quantities which remain invariant under this class of transformations.

Consider the semidirect product of the Lie group $SO(4)$ with an affine group $T, G = SO(4) \tilde{\times} T$ ($\tilde{\times}$ denotes semidirect product); G is called the inhomogeneous Lorentz group. One parameter subgroups of T correspond physically to translations of the origin of the space–time coordinate frame. One parameter subgroups of $SO(4)$ correspond physically to rotation about an axis or motion along an axis. It is straightforward to show G acts transitively on X: given $x \in X$, $g \in G$, then $gx \in X$. Since $G: G \times X \to X$, G is well defined on scalar functions $f \in \Lambda^0$, and this action is denoted $L_0, L_0: L_0 \times \Lambda^0 \to \Lambda^0$.

Since T_X and T_X^*, the tangent and cotangent bundles of X, are isomorphic to the direct product of X with itself, G has a well-defined transitive action on T_X and T_X^*. Since T_X can be identified with $\Lambda^1(T_X)$, while T_X^* can be identified with $\Lambda^1(T_X^*)$, G acts in a well-defined manner on $\Lambda^1(T_X^*)$, denoted L_1, $L_1: L_1 \times \Lambda^1 \to \Lambda^1$.

It is now necessary to extend the action of G to Λ^2, Λ^3, and Λ^4. To illustrate how this is accomplished, consider an orthonormal set of basis vectors $\{e_1, e_2, e_3, e_4\}$ for Λ^1 (the extension to a general basis is straightforward). ge_k is the action of g on e_k ($k = 1, 2, 3, 4$) for some $g \in G$; $\{e_k' = ge_k\}$ is a set of orthonormal basis vectors for $\Lambda^{1\prime}$. Since $\{e_i \wedge e_j \mid i = 1, 2, 3, j = 2, 3, 4\}$ is a basis for Λ^2, $\{ge_i \wedge ge_j \mid i = 1, 2, 3; j = 2, 3, 4\}$ is a basis for $\Lambda^{2\prime}$, and $L_2: L_2 \times \Lambda^2 \to \Lambda^2$ is the well-defined action of G on $\Lambda^2(T_X^*)$. Similarly, $\{ge_i \wedge ge_j \wedge ge_k,\ i = 1, 2,\ j = 2, 3,\ k = 3, 4\}$ is a basis for $\Lambda^{3\prime}$, and $\{ge_1 \wedge ge_2 \wedge ge_3 \wedge ge_4\}$ is a basis for $\Lambda^{4\prime}$, which lead to well-defined actions of G, $L_3: L_3 \times \Lambda^3 \to \Lambda_3$ and $L_4: L_4 \times \Lambda^4 \to \Lambda^4$. This can be summarized as follows.

Proposition: The diagram shown below commutes

$$\begin{array}{ccccccccc}
\Lambda^0 & \xrightarrow{d} & \Lambda^1 & \xrightarrow{d} & \Lambda^2 & \xrightarrow{d} & \Lambda^3 & \xrightarrow{d} & \Lambda^4 \\
L_0 \downarrow & & L_1 \downarrow & & L_2 \downarrow & & L_3 \downarrow & & L_4 \downarrow \\
\Lambda^{0\prime} & \xrightarrow{d'} & \Lambda^{1\prime} & \xrightarrow{d'} & \Lambda^{2\prime} & \xrightarrow{d'} & \Lambda^{3\prime} & \xrightarrow{d'} & \Lambda^{4\prime}
\end{array}$$

Proof: Again, the proof has three parts. First, observe that d and d' (the exterior derivative in the new coordinates), as well as L_k ($k = 0, 1, 2, 3, 4$) are well defined on each chart of X. Second, note that d, d', and L_k ($k = 0, 1, 2, 3, 4$) are well defined globally, using the transition functions to smoothly piece together the operators on overlapping coordinate charts. Third, the verification the diagram commutes is straightforward, because of the two preceding steps.

Since X has a well-defined inner product $\langle a, b\rangle$ is well defined, where either $a \in \Lambda^K$, $b \in \Lambda^K$, or $a \in \Lambda^K$, $*b \in \Lambda^{4-K}$. Both the real and imaginary parts of all these inner products remain invariant under the action of G. The Hamiltonian and Lagrangian functions result from forming linear combinations of these inner products.[2,4,5]

Example: In rectangular coordinates, the inner products invariant under the action of G are

(i) $\langle (g_e + ig_m), *[(s_m + is_e)\, dx \wedge dy \wedge dz \wedge icdt]\rangle$
$= (g_e + ig_m)(s_m + is_e)$

(ii) $\langle c(A_{ex} + iA_{mx})\, dx + c(A_{ey} + iA_{my})\, dy + c(A_{ez} + iA_{mz})\, dz + (\varphi_e + i\varphi_m) \wedge icdt,$
$*[(J_{mx} + iJ_{ex})\, dy \wedge dz \wedge icdt + (J_{my} + iJ_{ey})\, dz \wedge dx \wedge icdt + (J_{mz} + iJ_{ez})\, dx \wedge dy \wedge icdt + c(\rho_e + i\rho_m)\, dx \wedge dy \wedge dz]\rangle$
$= c(A_{ex} + iA_{mx})(J_{mx} + iJ_{ex}) + c(A_{ey} + iA_{my})(J_{my} + iJ_{ey}) + c(A_{ez} + iA_{mz})(J_{mz} + iJ_{ez}) - (\varphi_e + i\varphi_m)c(\rho_e + i\rho_m),$

(iii) $\langle c(D_x + iB_x)\, dy \wedge dz + c(D_y + iB_y)\, dz \wedge dx + c(D_z + iB_z)\, dx \wedge dy + (E_x + iH_x)\, dx \wedge icdt + (E_y + iH_y)\, dy \wedge icdt + (E_z + iH_z)\, dz \wedge icdt,$
$*[c(D_x + iB_x)\, dy \wedge dz + (D_y + iB_y)\, dz \wedge dx + c(D_z + iB_z)\, dx \wedge dy + (E_x + iH_x)\, dx \wedge icdt + (E_y + iH_y)\, dy \wedge icdt + (E_z + iH_z)\, dz \wedge icdt]\rangle$
$= 2\{c(D_x + iB_x)(E_x + iH_x) + c(D_y + iB_y)(E_y + iH_y) + c(D_z + iB_z)(E_z + iH_z)\},$

(iv) $\|(g_e + ig_m)\|^2 = (g_e + ig_m)(g_e + ig_m),$

(v) $\|c(A_{ex} + A_{mx})\, dx + c(A_{ey} + iA_{my})\, dy + c(A_{ez} + iA_{mz})\, dz + (\varphi_e + i\varphi_m) icdt\|^2$
$= c^2[(A_{ex} + iA_{mx})^2 + (A_{ey} + iA_{my})^2 + (A_{ez} + iA_{mz})^2] + (\varphi_e + i\varphi_m)^2,$

(vi) $\|c(D_x + iB_x)\, dy \wedge dz + c(D_y + iB_y)\, dz \wedge dx + c(D_z + iB_z)\, dx \wedge dy + (E_x + iH_x)\, dx \wedge icdt + (E_y + iH_y)\, dy \wedge icdt + (E_z + iH_z)\, dz \wedge icdt\|^2$
$= c^2[(D_x + iB_x)^2 + (D_y + iB_y)^2 + (D_z + iB_z)^2] + (E_x + iH_x)^2 + (E_y + iH_y)^2 + (E_z + iH_z)^2,$

(vii) $\|(J_{mx} + iJ_{ex})\, dy \wedge dz \wedge icdt + (J_{my} + iJ_{ey})\, dz \wedge dx \wedge icdt + (J_{mz} + iJ_{ez})\, dx \wedge dy \wedge icdt + c(\rho_e + i\rho_m)\, dx \wedge dy \wedge dz\|^2$
$= (J_{mx} + iJ_{ex})^2 + (J_{my} + iJ_{ey})^2 + (J_{mz} + iJ_{ez})^2 + c^2(\rho_e + i\rho_m)^2,$

(viii) $\|(s_m + is_e)\, dx \wedge dy \wedge dz \wedge icdt\|^2 = (s_m + is_e)^2.$

Remark: (i), (iv), and (viii) are often overlooked invariants, (cf. Refs. 5,6).

V. VARIATIONAL PRINCIPLES

For the sake of completeness, as well as to have an alternate interesting way in which to view the mathematical structure in electrodynamics, a Lagrangian and Hamiltonian formulation will now be discussed. The results presented here are more complete than any other of which the author is aware,[5,6] and illustrate a new relationship between dynamics based on exterior calculus and dynamics based on a calculus of variations approach. Since many excellent treatments[8,9] can be found in the literature on Lagrangian and Hamiltonian dynamics, but few good examples can be found on how to apply this knowledge, the general discussion is cursory, while the example is dwelt on at length.

The *Lagrangian function* L is defined as

$$L:\ \Lambda(T^*(X)) \times \Lambda(T^*(X)) \to \mathrm{R},$$

$$L = \tfrac{1}{2} RE(+\langle u_0, *\circ u_4\rangle + \langle u_1, *\circ u_3\rangle - \langle u_2, *\circ u_2\rangle - \langle u_3, *\circ u_1\rangle$$

TABLE I

	Components of associated generalized momentum			
	x	y	z	ict
g_e	J_{mx}	J_{my}	J_{mz}	
ig_m	iJ_{ex}	iJ_{ey}	iJ_{ez}	

$+(M_q, \mathrm{rot} M_0))$,

where $RE(a+ib)=a$, $a, b \in \mathbb{R}$.

In order to specify L on a chart of a manifold, one must give the local coordinates of the chart, all elements in $\Lambda(T^*(X))$ and all partial derivatives of elements in $\Lambda(T^*(X))$ with respect to local coordinates. The elements in $\Lambda(T^*(X))$ are called *generalized coordinates* while the partial derivatives are called *generalized velocities*.

The *generalized momenta* are defined as the partials of L with respect to the generalized velocities. The *Hamiltonian function* H is derived from L by computing the inner product of the generalized velocities with their respective generalized momenta and then subtracting the Lagrangian L; this transformation is called a *Legendre transformation*. The Hamiltonian function H is specified on a chart of a manifold by specifying coordinates on the chart, the generalized coordinates and the generalized momenta.

Solutions to Maxwell's equations are extremals to the *action integral*

$$\int_1^2 L\, de_1 \wedge de_2 \wedge de_3 \wedge de_4,$$

where the integral is evaluated along a space–time trajectory beginning at point 1 and ending at point 2, and $de_1 \wedge de_2 \wedge de_3 \wedge de_4$ is a unit basis vector for Λ^4. For a more complete and precise discussion of how to evaluate this integral, the reader is referred to the bibliography (Spivak,[10] Loomis-Sternberg,[11] Warner[7]).

Given a Lagrangian function, a well-defined recipe due to Euler and Lagrange exists for finding the associated equations of motions whose solutions are extremals to the action integral. Given a Hamiltonian function, a well-defined formula due to Hamilton exists for finding the associated equations of motion. Since both these approaches are independent of the constitutive relations, but depend only on the underlying differentiable manifold and its associated vector fields, the resulting equations of motion are said, picturesquely, to be totally geometric or topological in nature, independent of matter or vacuum.

TABLE II

	Components of associated generalized momenta			
	x	y	z	ict
cA_{ex}	0	E_z	$-E_y$	
cA_{ey}	$-E_z$	0	E_x	
cA_{ez}	E_y	$-E_x$	0	
$i\varphi_m$	$-icB_x$	$-icB_y$	$-icB_z$	
icA_{mx}	0	iH_z	$-iH_y$	
icA_{my}	$-iH_z$	0	iH_y	
icA_{mz}	iH_y	$-iH_z$	0	
φ_e	$-cD_x$	$-cD_y$	$-cD_z$	

TABLE III

	Components of associated generalized momentum			
	x	y	z	ict
E_x	0	$-cA_{ez}$	cA_{ey}	
E_y	cA_{ez}	0	$-cA_{ex}$	
E_z	$-cA_{ey}$	cA_{ex}	0	
icB_x	$-i\varphi_m$	0	0	
icB_y	0	$-i\varphi_m$	0	
icB_z	0	0	$-i\varphi_m$	
iH_x	0	$-icA_{mz}$	icA_{my}	
iH_y	icA_{mz}	0	$-icA_{mx}$	
iH_z	$-icA_{my}$	icA_{mx}	0	
cD_x	$-\varphi_e$	0	0	
cD_y	0	$-\varphi_e$	0	
cD_z	0	0	$-\varphi_e$	

Example: The Lagrangian function L is

$$L = -(cD_x \cdot E_x + cD_y \cdot E_y + cD_z \cdot E_z) + (cB_x \cdot H_x + cB_y \cdot H_y + cB_z \cdot H_z) - (g_m s_e - g_e s_m) - (cA_{ex} J_{ex} + cA_{ey} J_{ey} + cA_{ez} J_{ez} + \varphi_e \cdot c\rho_e) + (cA_{mx} J_{mx} + cA_{my} J_{my} + cA_{mz} J_{mz} + \varphi_m \cdot c\rho_m).$$

(A) The generalized coordinates are g_e and ig_m. The generalized velocities are all partials of g_e and ig_m with respect to x, y, z and ict. The x component of the generalized momentum associated with g_e is

$$x \text{ component} = \frac{\partial L}{\partial(\partial g_e/\partial x)} = \frac{\partial L}{\partial cA_{ex}} \cdot \frac{\partial cA_{ex}}{\partial(\partial g_e/\partial x)} = J_{mx}.$$

Note that to compute the generalized momentum it is necessary not only to compute $(\partial L/\partial cA_{ex})$ but also to know from Maxwell's equations that $\partial cA_{ex}/\partial(\partial g_e/\partial x) = +1$. In like manner it is straightforward to find all the generalized momenta, and the results are summarized in the Table I.

The Hamiltonian function H is

$$H_A = -(cB_x \cdot H_x + cB_y \cdot H_y + cB_z \cdot H_z) + (cD_x \cdot E_x + cD_y \cdot E_y + cD_z \cdot E_z) - (g_e s_m - g_m s_e)$$

and is independent of the generalized momentum. The Euler–Lagrange equations of motion are

$$g_e:\; s_m - \frac{\partial}{\partial x}(J_{mx}) - \frac{\partial}{\partial y}(J_{my}) - \frac{\partial}{\partial z}(J_{mz}) - \frac{\partial}{\partial ict} - ic\rho_m = 0,$$

TABLE IV

Generalized coordinate	components of associated generalized momentum			
		y	z	ict
J_{mx}	g_e	0	0	
J_{my}	0	g_e	0	
J_{mz}	0	0	g_e	0
$ic\rho_m$	0	0	0	$-g_e$
iJ_{ex}	ig_m	0	0	0
iJ_{ey}	0	ig_m	0	0
iJ_{ez}	0	0	ig_m	0
$c\rho_e$	0	0	0	$-ig_m$

TABLE V

Generalized Coordinate	Components of Associated Generalized momentum			
	x	y	z	ict
s_m	0	0	0	
is_e	0	0	0	

$$ig_m:\ is_e - \frac{\partial}{\partial x}(iJ_{ex}) - \frac{\partial}{\partial y}(iJ_{ey}) - \frac{\partial}{\partial z}(iJ_{ez}) - \frac{\partial}{\partial ict} - c\rho_e = 0$$

The Hamiltonian equations of motion are identical:

$$\frac{\partial H}{\partial g_e} = -s_m = -\left(\frac{\partial}{\partial x}J_{mx} + \frac{\partial}{\partial y}J_{my} + \frac{\partial}{\partial z}J_{mz} + \frac{\partial}{\partial ict} - ic\rho_m\right),$$

$$\frac{\partial H}{\partial ig_m} = -is_e = -\left(\frac{\partial}{\partial x}iJ_{ex} + \frac{\partial}{\partial y}iJ_{ey} + \frac{\partial}{\partial z}iJ_{ez} + \frac{\partial}{\partial ict} - c\rho_e\right).$$

Since the Hamiltonian is independent of the generalized momentum, the dual equations involved derivatives of H with respect to momenta are all zero. Note these equations are identical to those in Sec. 2, Eq. (4).

(B) The generalized coordinates and generalized momenta are tabulated in Table II.

The Hamiltonian function H is

$$H_B = -(cD_x\cdot E_x + cD_y\cdot E_y + cD_z\cdot E_z) + (H_x\cdot cB_x + H_y\cdot cB_y + H_z\cdot cB_z) - (g_e s_m - g_m s_e) - (cA_{ex}\cdot J_{mx} + cA_{ey}\cdot J_{my} + cA_{ez}\cdot J_{mz} + \varphi_m\cdot c\rho_m) + (cA_{mx}J_{ex} + cA_{my}J_{ey} + cA_{mz}J_{ez} + \varphi_e\cdot c\rho_e).$$

The Euler–Lagrange and Hamiltonian equations of motion are found in Sec. 2, Eq. (3).

(C) The generalized coordinates and generalized momenta are tabulated in Table III.

The Hamiltonian function H is

$$H_C = -(cB_x\cdot H_x + cB_y\cdot H_y + cB_z\cdot H_z) + (cD_x\cdot E_x + cD_y\cdot E_y + cD_z\cdot E_z) - (g_e s_m - g_m s_e)$$

The Euler–Lagrange and Hamiltonian equations of motion are found in Sec. 2, Eq. (2).

(D) The generalized coordinates and generalized momenta are tabulated in Table IV.

The Hamiltonian function H is

$$H_D = (cD_x\cdot E_x + cD_y\cdot E_y + cD_z\cdot E_z) - (cB_x\cdot H_x + cB_y\cdot H_y + cB_z\cdot H_z) + (cA_{mx}J_{ex} + cA_{my}J_{ey} + cA_{mz}J_{ez} + \varphi_e\cdot c\rho_e) - (cA_{ex}J_{mx} + cA_{ey}J_{my} + cA_{ez}J_{mz} + \varphi_m\cdot c\rho_m).$$

The Euler–Lagrange and Hamiltonian equations of motion are found in Sec. 2, Eq. (1).

(E) The generalized coordinates and generalized momenta are summarized in Table V.

The Hamiltonian function H is

$$H_E = -L = -(cB_x\cdot H_x + cB_y\cdot H_y + cB_z\cdot H_z) + (cD_x\cdot E_x + cD_y\cdot E_y + cD_z\cdot E_z) - (g_e s_m - g_m s_e) - (cA_{ex}J_{mx} + cA_{ey}J_{my} + cA_{ez}J_{mz} + \varphi_m c\rho_m) + (cA_{mx}J_{ex} + cA_{my}J_{ey} + cA_{mz}J_{ez} + \varphi_e c\rho_e).$$

The Euler–Lagrange and Hamiltonian equations of motion are

$$g_e = 0, \quad ig_m = 0.$$

ACKNOWLEDGMENTS

The author appreciates several conversations with Professor R.W. Brockett of Harvard University for stimulating this work. Professor R. Hermann of Rutgers University and Dr. R. Brammer of NASA Goddard Spaceflight Center offered several concrete suggestions for improving the exposition of the ideas presented here, as did Dr. R. Ludwig and Dr. I. Sandberg of Bell Laboratories.

APPENDIX: MATHEMATICAL PRELIMINARIES

This section is largely tutorial, sketching some of the fundamental concepts of multilinear algebra, differentiable manifolds, and exterior calculus. Suitable references can be found in the bibliography (Nelson,[12] Warner,[7] Spivak,[10] Loomis-Sternberg[11]).

A. Multilinear algebra

Let R denote the real numbers, and let V and W be finite-dimensional real linear vector spaces. V^* denotes the dual space of V, consisting of all real valued functions of V. The *direct product* of V with W is denoted $V\times W$ and consists of all linear combinations of pairs (v, w), with $v\in V$ and $w\in W$. The k-fold direct product of V with itself, denoted V^k, the l-fold direct product of V^* with itself, denoted V^{*l}, and the mixed direct product of $V^k\times V^{*l}$ are defined in an identical fashion.

Let $S(V^k)$ be the subspace of V^k generated by the set of all elements of the form

$$\left.\begin{matrix}(v_1, \ldots, v_i' + v_i'', \ldots, v_k) - (v_1, \ldots, v_i', \ldots, v_k) - (v_1, \ldots, v_i'', \ldots, v_k) \\ (v_1, \ldots, av_i, \ldots, v_k) - a(v_1, \ldots, v_i, \ldots, v_k)\end{matrix}\right\} \begin{matrix} v_i, v_i', v_i'' \in V, \quad i = 1\cdots k, \\ a\in \mathrm{R}.\end{matrix}$$

The quotient space $T^k = V^k/S(V^k)$, the set $\{x: (y-x) \in S(V^k)$, for all $y\in V^k\}$, is called the set of *kth order contravariant tensors*. In a similar manner, the quotient space $T^{*m} = V^{*m}/S(V^{*m})$ can be defined, and is called the set of mth *order covariant tensors*. Finally, the quotient space of mixed tensors, $T^{k,*m} = V^{k,*m}/S(V^{k,*m})$ can be defined in an analogous fashion. The terms contravariant and covariant will in general be dropped, being clear from context (cf. Spivak,[13] pp. 4–8 to 4–12). The direct sum $T(V)$, denoted by $\oplus$,

$$T(V) = T^{0,*0} \oplus T^{1,*0} \oplus T^{0,*1} \oplus \cdots = \bigoplus_{k=0}^{\infty} \bigoplus_{m=0}^{\infty} T^{k,*m}$$

where $T^{0,*0} = \mathrm{R}$, is called the *tensor algebra* of V. Consider an element in $T^k(V)$, denoted A; A is called *alternating* or *skew-symmetric* if

$$A(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -A(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k) \quad \forall v_1, \ldots, v_k \in V.$$

The set of all alternating kth order tensors is denoted $\Lambda^k(V)$, and is clearly a subspace of $T^k(V)$. If $L: V \to W$ is a linear transformation, then $L^*: T^k(W) \to T^k(V)$ is defined by $(L^* \circ T^k)(v_1, \ldots, v_k) = T^k[L(v_1), \ldots, L(v_k)]$. In particular, if $L: V \to W$, then

$$L^*(u \wedge v) = (L^*(u)) \wedge (L^*(v));$$

$\Lambda(V) = \Lambda^0(V) \oplus \cdots \oplus \Lambda^n(V)$, n equals dimension of V, is the *contravariant exterior algebra* of V, while $\Lambda(V^*) = \Lambda^0(V^*) \oplus \cdots \oplus \Lambda^n(V^*)$ is the *covariant exterior algebra* of V [which is defined in a manner entirely analogous to $\Lambda(V)$]. This work will concentrate entirely on exterior algebra. Multiplication in the exterior algebra $\Lambda(V)$ is denoted by "$\wedge$" the *exterior* or *wedge product*, a natural generalization of the three-dimensional cross product operation on two vectors. The exterior algebra is a graded algebra: if $u \in \Lambda^k(V)$, $v \in \Lambda^l(V)$, then $u \wedge v \in \Lambda^{k+l}(V)$. The exterior product obeys the following properties:

$$\left.\begin{aligned} u \wedge v &= (-1)^{kl} v \wedge u, \\ (u_1 + u_2) \wedge v &= (u_1 \wedge v) + (u_2 \wedge v), \\ u \wedge (v_1 + v_2) &= (u \wedge v_1) + (u \wedge v_2), \\ u \wedge (v \wedge w) &= (u \wedge v) \wedge w, \\ (au) \wedge v &= u \wedge (av) = a[u \wedge v] \end{aligned}\right\} \begin{aligned} &u, u_1, u_2 \in \Lambda^k(V), \\ &v, v_1, v_2 \in \Lambda^l(V), \\ &w \in \Lambda^m(V), \\ &a \in \mathrm{R}. \end{aligned}$$

If $\{e_1, \ldots, e_n\}$ is a basis for V, then $\{e_{i_1} \wedge \cdots \wedge e_{i_l} \mid i_1, \ldots, i_l = 1, \ldots, n\}$ is a basis for $\Lambda^l(V)$. In particular, note that $e_{i_1} \wedge \cdots \wedge e_{i_n}$ $(i_1, \ldots, i_n = 1, \ldots, n)$ is a basis vector for $\Lambda^n(V)$. Since $\Lambda^n(V)$ is one-dimensional, the sign on this basis vector can be either positive or negative, corresponding to a choice in *orientation* (cf. "right-handed" and "left-handed" coordinates in R^3).

Let $\langle\,,\,\rangle: V \times V \to \mathrm{R}$ be the standard sum-of-squares *inner product* on V, positive definite and symmetric in its arguments. Choose an orthonormal basis for V, $\{e_1, \ldots, e_n\}$. Let $a \in \Lambda^k(V)$, $b \in \Lambda^l(V)$,

$$a = \sum_{i_1 < \cdots < i_k} a(i_1, \ldots, i_k) e_{i_1} \wedge \cdots \wedge e_{i_k}, \quad a(i_1, \ldots, i_k) \in \mathrm{R}$$

$$b = \sum_{i_1 < \cdots < i_l} b(i_1, \ldots, i_l) e_{i_1} \wedge \cdots \wedge e_{i_l}, \quad b(i_1, \ldots, i_l) \in \mathrm{R},$$

the inner product of a and b, denoted $\langle a, b\rangle$ is defined by

$$\langle a, b\rangle = \begin{cases} \sum_{i_1 < \cdots < i_k} a(i_1, \ldots, i_k) b(i_1, \ldots, i_k) \langle e_{i_1}, e_{i_1}\rangle \cdots \langle e_{i_k}, e_{i_k}\rangle & k = l \\ 0, & k \neq l. \end{cases}$$

The *Hodge star operator*, denoted $*$, $*: \Lambda^k(V) \to \Lambda^{n-k}(V)$, is well defined by the requirement that for any orthonormal basis $e_1, \ldots, e_n$ of V,

$$*: (e_1 \wedge \cdots \wedge e_k) = \pm (e_{k+1} \wedge \cdots \wedge e_n),$$

where the plus sign is chosen if $+e_1 \wedge \cdots \wedge e_k \wedge e_{k+1} \wedge \cdots \wedge e_n$ is in the basis for $\Lambda^n(V)$, and the minus sign is otherwise chosen.

The requirement on the inner product and Hodge star operator that the basis be orthonormal can be relaxed, and the interested reader is referred to the bibliography (Warner,[7] Flanders,[1] Loomis-Sternberg,[11] Spivak[10]).

Example (R^3): Choose a rectangular set of orthonormal basis vectors $\{u_x, u_y, u_z\}$:

<u>$\Lambda(\mathrm{R}^3)$ Basis</u>

$\Lambda^0(\mathrm{R}^3)$ 1

$\Lambda^1(\mathrm{R}^3)$ u_x, u_y, u_z

$\Lambda^2(\mathrm{R}^3)$ $u_y \wedge u_z, u_z \wedge u_x, u_x \wedge u_y$

$\Lambda^3(\mathrm{R}^3)$ $u_x \wedge u_y \wedge u_z$

<u>Dual Forms</u>

$*1 = u_x \wedge u_y \wedge u_z$,	$*u_y \wedge u_z = u_x$,
$*u_x = u_y \wedge u_z$,	$*u_z \wedge u_x = u_y$,
$*u_y = u_z \wedge u_x$,	$*u_x \wedge u_y = u_z$,
$*u_z = u_x \wedge u_y$,	$*u_x \wedge u_y \wedge u_z = -1$

A zero form may be interpreted physically as a scalar, while a 1-form may be interpreted as a directed line segment, a 2-form as a directed area, and a 3-form as a directed volume.

Example (Space–Time): (For an extensive discussion of the mathematics underlying space–time, the reader is referred to Penrose.[3]) Choose a rectangular set of orthonormal basis vectors $\{dx, dy, dz, icdt\}$ where $i = \sqrt{-1}$ and c = speed of light, with orientation $+dx \wedge dy \wedge dz \wedge icdt$:

Λ (Space–Time)	Basis
Λ^0	1
Λ^1	$dx,\ dy,\ dz,\ icdt$
Λ^2	$dy\wedge dz,\ dz\wedge dx,\ dx\wedge dy,\ dx\wedge icdt,\ dy\wedge icdt,\ dz\wedge icdt$
Λ^3	$dy\wedge dz\wedge icdt,\ dz\wedge dx\wedge icdt,\ dx\wedge dy\wedge icdt,\ dx\wedge dy\wedge dz$
Λ^4	$dx\wedge dy\wedge dz\wedge icdt$

Dual Forms

$*1=dx\wedge dy\wedge dz\wedge icdt,$	$*dx\wedge icdt=dy\wedge dz,$
$*dx=dy\wedge dz\wedge icdt,$	$*dy\wedge icdt=dz\wedge dx,$
$*dy=dz\wedge dx\wedge icdt,$	$*dz\wedge icdt=dx\wedge dy,$
$*dz=dx\wedge dy\wedge icdt,$	$*dy\wedge dz\wedge icdt=-dx,$
$*icdt=-dx\wedge dy\wedge dz,$	$*dz\wedge dx\wedge icdt=-dy,$
$*dy\wedge dz=dx\wedge icdt,$	$*dx\wedge dy\wedge icdt=-dz$
$*dz\wedge dx=dy\wedge icdt,$	$*dx\wedge dy\wedge dz=icdt,$
$*dx\wedge dy=dz\wedge icdt,$	$*dx\wedge dy\wedge dz\wedge icdt=1.$

Note: $*\circ*\circ u_k=(-1)^k u_k,\ u_k\in\Lambda^k,\ k=0,1,2,3,4.$

B. Differentiable manifolds

Let X be a set, U an open subset of X, and m a map, $m:U\to V\subset R^n$ where m is bijective (one-one and onto). The pair (m, U) defines a *chart* on X; m specifies *local coordinates* on a subset of X. Consider two charts on X, (m_1, U_1) and (m_2, U_2); suppose $m_1m_2^{-1}, m_2m_1^{-1}: R^n\to R^n$ are C^k functions, i.e., differentiable k times but not $(k+1)$. $m_1m_2^{-1}$ and $m_2m_1^{-1}$ are called *transition functions*. A collection of charts on a set X is denoted A; A is called an *atlas* for X if the chart domains cover X, and the associated transition functions have open domains and are C^∞. A *complete atlas* is the union of all possible atlases for a set X. A *differentiable manifold* is a set X together with a complete atlas. Intuitively, a differentiable manifold is a union of nondisjoint sets, each of which is locally diffeomorphic to R^n, which is pieced together by the transition functions.

Let X and Y be differentiable manifolds. Choose any chart on X and Y with coordinate maps m_x and m_y, respectively. Then $f:X\to Y$ is defined by the composite map $m_y^{-1}\circ f\circ m_x$. Let p be a point in R^n, v a vector in R^n. To every function f defined in the neighborhood of p, associate a number called the *directional derivative* of f in the direction v at p, denoted $D_vf(p)$ and defined by

$$D_vf(p)=\frac{d}{dt}f(p+tv)\Big|_{t=0}$$

Consider now the manifold X; the *tangent vector* to X at p in the direction v is a map which associates with every C^∞ function f, defined on a neighborhood of p, a real number $D_vf(p)$ such that

(i) $f_1=f_2$ implies $D_vf_1(p)=D_vf_2(p)$,

(ii) $D_v(f+g)(p)=D_vf(p)+D_vg(p)$,

(iii) $D_v(f\cdot g)(p)=[D_vf(p)]g(p)+f(p)[D_vg(p)]$.

The *tangent space* of X at p is the set of all tangent vectors, for all $v\in R^N$. The tangent space of X at p can be shown to be a vector space, and thus has an associated dual vector space, called the *cotangent space* of X at p, the set of all linear functionals on the tangent space. The *tangent bundle* of a manifold X is the direct product of the set of all tangent spaces at all points $p\in X$; the *cotangent bundle* is the direct product of the set of all cotangent spaces at all points $p\in X$ with X. A *Riemannian differentiable manifold* is a differentiable manifold with a prescribed norm on the tangent bundle.

Example: Let X be a finite-dimensional vector space. Choose a basis for X, $\{e_1,\dots,e_n\}$, so $x\in X$ can be expressed as $x=x_1e_1+\cdots+x_ne_n$. Define the coordinate map $m(x_1e_1+\cdots+x_ne_n)=(x_1,\dots,x_n)$. An atlas for X is the set of coefficients, with respect to the basis $\{e_1\cdots e_n\}$, of all points $x\in X$. A second atlas for X is the set of coefficients, with respect to a different basis $\{e_1',\dots,e_n'\}$, of all points $x\in X$. The transition functions are given by a C^∞ linear transformation describing the change of basis. A complete atlas can be generated by considering all possible sets of basis vectors for X; thus, X is a differentiable manifold. The tangent space and cotangent space of X at a point p are clearly n-dimensional, so the tangent bundle and cotangent bundles are locally diffeomorphic to R^{2n}. Together with the standard Euclidean norm on the tangent bundle, X is a Riemannian differentiable manifold.

C. Exterior calculus

If $f:R^n\to R$ is a scalar differentiable function of n variables, then f is a *zero differential form*, or 0-form. The total differential of f, $df(x_1,\dots,x_n)=(\partial f/\partial x_1)dx_1+\cdots+(\partial f/\partial x_n)dx_n$ is called a *one differential form* or *1-form* if each component $\partial f/\partial x_k$, $k=1,\dots,n$ is differentiable. Note that f may be considered in Λ^0, while df is an element of Λ^1. The *exterior derivative* generalizes the concept of a total differential using exterior algebra:

Theorem[10]: Let $u\in\Lambda^k$. Then the exterior derivative of u is $du\in\Lambda^{k+1}$, and is defined by

$$du=\sum_{i_1<\cdots<i_k} du_{i_1,\dots,i_k}\ dx_{i_1}\wedge\cdots\wedge dx_{i_k},$$

where $du_{i_1,\dots,i_k}$ is the total differential of the i_j component of u, and the exterior derivative d obeys the following properties:

$$\left.\begin{array}{l}\text{(i) } d(u+v)=du+dv\\ \text{(ii) } d(u\wedge v)=du\wedge v+(-1)^k u\wedge dv\\ \text{(iii) } d(du)=0\to d\circ d\cong 0\end{array}\right\}\ u\in\Lambda^k,\ \ v\in\Lambda^l.$$

A differential form u is called *closed* if $du=0$, *exact* if $dv=u$; it can be shown every exact form is closed, but the converse is not true. The *adjoint* δ of the exterior derivative is defined such that

$$\langle du, v\rangle=\langle u, \delta v\rangle, \quad u\in\Lambda^k, \quad v\in\Lambda^{k+1}.$$

It can be shown $\delta=*d*$. A differential form is called *coclosed* if $\delta u=0$, *coexact* if $\delta v=u$. The *Laplace—Beltrami operator* is defined as $\Delta=d\delta+\delta d$, and is linear, $\Delta:\Lambda^k\rightarrow\Lambda^k$ $(k=0,\ldots,n)$. Elements in the kernel of Δ are called *harmonic*, and the set of all such k-forms is denoted $H^k=\{u:\Delta u=0, \ u\in\Lambda^k\}$. It can be shown the Laplace—Beltrami operator is *elliptic* (Warner[7], pp. 250—251).

A question of great practical interest is solving $\Delta u=v$, given v subject to suitable boundary conditions. For the special case where the underlying manifold X is compact, this question has been answered by

Theorem[7] (Hodge—DeRham—Kodaira): $\Delta u=v$ has a unique solution $u\in\Lambda^k$ iff $v\in\Lambda^k$ is orthogonal to H^k. Furthermore, Λ^k can be decomposed into a direct sum of three mutually orthogonal vector spaces,

$$\begin{aligned}\Lambda^k&=H^k\oplus\Delta(\Lambda^k)\\&=H^k\oplus(d\delta+\delta d)(\Lambda^k)\\&=H^k\oplus d(\Lambda^{k-1})\oplus\delta(\Lambda^{k+1})\end{aligned}$$

and H^k is finite-dimensional.

Example ($\mathbf{R}^3$): $\Lambda^0\xrightarrow{d}\Lambda^1\xrightarrow{d}\Lambda^2\xrightarrow{d}\Lambda^3$.

For simplicity choose a rectangular orthonormal basis $\{dx,dy,dz\}$ with orientation $+dx\wedge dy\wedge dz$. Then

$$f_0\in\Lambda^0, \quad df_0=\frac{\partial f_0}{\partial x}dx+\frac{\partial f_0}{\partial y}dy+\frac{\partial f_0}{\partial z}dz\in\Lambda^1;$$

$$f_1=f_{1x}dx+f_{1y}dy+f_{1z}dz\in\Lambda^1,$$

$$\begin{aligned}df_1&=\left(\frac{\partial f_{1x}}{\partial x}dx+\frac{\partial f_{1x}}{\partial y}dy+\frac{\partial f_{1x}}{\partial z}\right)\wedge dx\\&\quad+\left(\frac{\partial f_{1y}}{\partial x}dx+\frac{\partial f_{1y}}{\partial y}dy+\frac{\partial f_{1y}}{\partial z}\right)\wedge dy\\&\quad+\left(\frac{\partial f_{1z}}{\partial x}dx+\frac{\partial f_{1z}}{\partial y}dy+\frac{\partial f_{1z}}{\partial z}dz\right)\wedge dz\\&=\left(\frac{\partial f_{1z}}{\partial y}-\frac{\partial f_{1y}}{\partial z}\right)dy\wedge dz+\left(\frac{\partial f_{1x}}{\partial z}-\frac{\partial f_{1z}}{\partial x}\right)dz\wedge dx\\&\quad+\left(\frac{\partial f_{1y}}{\partial x}-\frac{\partial f_{1x}}{\partial y}\right)dx\wedge dy\in\Lambda^2;\end{aligned}$$

$$f_2=f_{2x}dy\wedge dz+f_{2y}dz\wedge dx+f_{2z}dx\wedge dy\in\Lambda^2,$$

$$\begin{aligned}df_2&=\left(\frac{\partial f_{2x}}{\partial x}dx+\frac{\partial f_{2x}}{\partial y}dy+\frac{\partial f_{2x}}{\partial z}dz\right)\wedge dy\wedge dz\\&\quad+\left(\frac{\partial f_{2y}}{\partial x}dx+\frac{\partial f_{2y}}{\partial y}dy+\frac{\partial f_{2y}}{\partial z}dz\right)\quad dz\quad dx\\&\quad+\left(\frac{\partial f_{2z}}{\partial x}dx+\frac{\partial f_{2z}}{\partial y}dy+\frac{\partial f_{2z}}{\partial z}dz\right)\wedge dx\wedge dy\\&=\left(\frac{\partial f_{2x}}{\partial x}+\frac{\partial f_{2y}}{\partial y}+\frac{\partial f_{2z}}{\partial z}\right)\wedge dx\wedge dy\wedge dz\in\Lambda^3;\end{aligned}$$

$$f_3\,dx\wedge dy\wedge dz\in\Lambda^3, \quad df_3=0.$$

Similarly, it can be shown that

$$\delta f_3=\frac{\partial f_3}{\partial x}dy\wedge dz+\frac{\partial f_3}{\partial y}dz\wedge dx+\frac{\partial f_3}{\partial z}dx\wedge dy\in\Lambda^2,$$

$$\delta f_2=\left(\frac{\partial f_{2z}}{\partial y}-\frac{\partial f_{2y}}{\partial z}\right)dx+\left(\frac{\partial f_{2x}}{\partial z}-\frac{\partial f_{2z}}{\partial x}\right)dy+\left(\frac{\partial f_{2y}}{\partial x}-\frac{\partial f_{2x}}{\partial y}\right)\in\Lambda^1,$$

$$\delta f_1=\frac{\partial f_{1x}}{\partial x}+\frac{\partial f_{1y}}{\partial y}+\frac{\partial f_{1z}}{\partial z}\in\Lambda^0,$$

$$\delta f_0=0.$$

Thus, the exterior derivative subsumes the operations of gradient, curl, and divergence. The Laplace—Beltrami operator $\Delta=d\delta+\delta d$ simplifies for this choice of basis. Define $D=-(\partial^2/\partial x^2+\partial^2/\partial y^2+\partial^2/\partial z^2)$, so

$$\Delta f_0=-\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)f_0=Df_0,$$

$$\Delta f_1=(Df_{1x})dx+(Df_{1y})dy+(Df_{1z})dz,$$

$$\Delta f_2=(Df_{2x})dy\wedge dy+(Df_{2y})dz\wedge dx+(Df_{2z})dx\wedge dy,$$

$$\Delta f_3=-\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)f_3\,dx\wedge dy\wedge dz=(Df_3)\,dx\wedge dy\wedge dz.$$

If the manifold is restricted to a compact subset of $\mathbf{R}^3$, then the Hodge—DeRham—Kodaira decomposition theorem shows that

$$\begin{aligned}f_0&=f_0^H\oplus\delta f_1^C,\\f_1&=f_1^H\oplus df_0^E\oplus\delta f_2^C,\\f_2&=f_2^H\oplus df_1^E\oplus\delta f_3^C,\\f_3&=f_3^H\oplus df_2^E,\end{aligned}$$

where the superscripts H, E, and C denote harmonic, exact, and coexact, respectively. To be more explicit,

$$\begin{aligned}f_0&=f_0^H\oplus\delta(f_{1x}^C dx+f_{1y}^C dy+f_{1z}^C dz)\\&=f_0^H\oplus\left[\frac{\partial}{\partial x}f_{1x}^C+\frac{\partial}{\partial y}f_{1y}^C+\frac{\partial}{\partial z}f_{1z}^C\right],\\f_1&=f_1^H\oplus df_0^E\oplus\delta[f_{2x}^C dy\wedge dz+f_{2y}^C dz\wedge dx+f_{2z}^C dx\wedge dy]\\&=f_1^H\oplus\left[\frac{\partial}{\partial x}f_0^E dx+\frac{\partial}{\partial y}f_0^E dy+\frac{\partial}{\partial z}f_0^E dz\right]\\&\quad\oplus\left[\left(\frac{\partial}{\partial y}f_{2z}^C-\frac{\partial}{\partial z}f_{2y}^C\right)dx+\left(\frac{\partial}{\partial z}f_{2x}^C-\frac{\partial}{\partial x}f_{2z}^C\right)dy\right.\\&\quad\left.+\left(\frac{\partial}{\partial x}f_{2y}^C-\frac{\partial}{\partial y}f_{2x}^C\right)dz\right],\\f_2&=f_2^H\oplus d[f_{1x}^E dx+f_{1y}^E dy+f_{1z}^E dz]\oplus\delta[f_3^C dx\wedge dy\wedge dz]\\&=f_2^H\oplus\left[\left(\frac{\partial}{\partial y}f_{1z}^E-\frac{\partial}{\partial z}f_{1y}^E\right)dy\wedge dz\right.\\&\quad\left.+\left(\frac{\partial}{\partial z}f_{1x}^E-\frac{\partial}{\partial x}f_{1z}^E\right)dz\wedge dx+\left(\frac{\partial}{\partial x}f_{1y}^E-\frac{\partial}{\partial y}f_{1x}^E\right)dx\wedge dy\right.\\&\quad\oplus\left[\frac{\partial}{\partial x}f_3^C dy\wedge dz+\frac{\partial}{\partial y}f_3^C dz\wedge dx+\frac{\partial}{\partial z}f_3^C dx\wedge dy\right],\\f_3&=f_3^H\oplus d[f_{2x}^E dy\wedge dz+f_{2y}^E dz\wedge dx+f_{2z}^E dx\wedge dy]\\&=f_3^H\oplus\left[\left(\frac{\partial}{\partial x}f_{2x}^E+\frac{\partial}{\partial y}f_{2y}^E+\frac{\partial}{\partial z}f_{2z}^E\right)dx\wedge dy\wedge dz\right].\end{aligned}$$

For the special case where the manifold is simply connected,

$$f_0^H=\text{const}, \quad f_1^H=0,$$

$$f_2^H=0,\quad f_3^H=(\text{const})\,dx\wedge dy\wedge dz.$$

In other words, any scalar 0-form can be written as the sum of a constant function plus the divergence of a function, any 1-form can be expressed as the curl of a vector valued function plus the gradient of a scalar function, any 2-form can be written as the curl of a vector valued function plus the gradient of a scalar function, and any 3-form can be written as the sum of a constant function plus the divergence of a vector valued function.

A second approach to this decomposition is to expand each f_k $(k=0,1,2,3)$ in eigenfunctions of the Laplace—Beltrami operator:

$$f_0=f_0^H+\sum_{i=1}^{\infty}\langle f_0,\delta u_i^C\rangle\delta u_i^C,$$

$$f_1=\sum_{i=1}^{\infty}\langle f_1,du_i\rangle du_i^{0E}+\sum_{j=1}^{\infty}\langle f_1,\delta u_j^{2C}\rangle\delta u_j^{2C},$$

$$f_2=\sum_{i=1}^{\infty}\langle f_2,du_i^{1E}\rangle du_i^{1E}+\sum_{j=1}^{\infty}\langle f_2,\delta u_j^{3C}\rangle\delta u_j^{3C},$$

$$f_3=f_3^H+\sum_{i=1}^{\infty}\langle f_3,du_i^{2E}\rangle du_i^{2E}.$$

The $\{u_j^i\}$, $j=1,2,\ldots,$ $i=0,1,2,3$ are eigenfunctions of the Laplace—Beltrami operator

$$\Delta:\Lambda^i\rightarrow\Lambda^i,\quad i=0,1,2,3\ ,$$

$$\Delta u_j^i=\lambda_j u_j^i.$$

Various properties of these igenfunctions are discussed in the text.

[1]H. Flanders, *Differential Forms* (Academic, New York, 1963), pp. 44—48.
[2]F.J. Dyson, Bull. Am. Math. Soc. 78, 635 (1972).
[3]R. Penrose, "Structure of Space—Time," in *Battelle Rencontres* (Benjamin, New York, 1968), pp. 121—235.
[4]A. Sommerfeld, *Electrodynamics* (Academic, New York, 1964), pp. 212—241.
[5]A.O. Barut, *Electrodynamics and Classical Theory of Fields and Particles* (MacMillan, New York, 1964), p. 121.
[6]C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1970, 1973), pp. 79—129.
[7]F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups* (Scott Foresman, Glenview, Ill., 1971).
[8]R. Abraham, *Foundations of Mechanics* (Benjamin, New York, 1967).
[9]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 1950), Chap. 11.
[10]M. Spivak, *Calculus on Manifolds* (Benjamin, New York, 1965).
[11]L. Loomis and S. Sternberg, *Advanced Calculus* (Addison-Wesley, Reading, Mass., 1968).
[12]E. Nelson, *Tensor Analysis* (Princeton, U.P., Princeton, N.J., 1967).
[13]M. Spivak, *Differential Geometry* (Publish or Perish, Boston, 1970), Vol. I.