SPACETIME CALCULUS for GRAVITATION THEORY

David Hestenes
Arizona State University

Abstract

A new gauge theory of gravitation on flat spacetime has recently been developed by Lasenby, Doran, and Gull in the language of Geometric Calculus. This paper provides a systematic account of the mathematical formalism to facilitate applications and extensions of the theory. It includes formulations of differential geometry, Lie derivatives and integrability theorems which are coordinate-free and gauge-covariant. Emphasis is on use of the language to express physical and geometrical concepts.

CONTENTS

1996 PACS: 04.20.Cv, 11.15.-q, 02.40.-k, 11.30.Cp.

Introduction

Lasenby, Doran and Gull have recently created a powerful coordinate-free reformulation, refinement, and extension of general relativity [1,2]. It is a gauge theory on flat spacetime, but it retains the attractive geometric structure of Einstein's theory. The mathematical formalism which they employ comes mostly from [3], with additional pieces from [4-8] and elsewhere. It is not easy, however, to assimilate and adapt the formalism from these references. Although much of that has already been done in [1], further discussion and analysis will help make the whole approach more accessible. Also, the emphasis here is more strongly focused on notions of differential geometry in formulating the theory. Indeed, the method amounts to a new approach to differential geometry which could fairly be called *gauge geometry*.

The general mathematical formalism is called *geometric calculus* or, when it refers specifically to spacetime, *spacetime calculus*. This paper presents a systematic account of spacetime calculus for the purposes of gravitation theory. It is divided into three parts.

Part I reviews the fundamentals of spacetime calculus with emphasis on those relevant to gravitation theory. Many of the proofs are omitted, since they can easily be supplied or found in the references. Otherwise, the treatment is self-contained.

Part II develops gauge covariant Riemannian geometry on flat spacetime. The main objective is to clarify the fundamental ideas and provide a systematic account of the definitions, theorems, proofs, and computational techniques needed to apply the spacetime calculus efficiently to any physical problem. Specific physical applications are not addressed here; excellent examples, which amply demonstrate the computational power of the calculus, have been worked out in [1,2] and [9–13].

Part III extends the spacetime calculus to deal systematically with congruences of curves and associated concepts, such as Lie derivatives, Killing vectors and the theorem of Frobenius. The gauge covariant formulation of these concepts is a new idea with broad implications for mathematics as well as physics. The treatment here is primarily a translation of well known concepts and results into the new formalism; proofs are consequently sketchy or omitted. (For comparison, a standard "modern" approach to the same topics can be found in [14]). The translation is nontrivial and we concentrate here on consolidating the main ideas and results to facilitate applications. The main applications are likely to involve a "relativistically gauge covariant" continuum mechanics, including gravitational wave propagation. As a first step, a general definition and equation of motion for the deformation tensor is presented.

In a lengthy and profound analysis of the relation between physics and geometry, composed more than a decade before the advent of Einstein's General Theory of Relativity (GR), Henri Poincaré [15] concluded that: "One geometry cannot be more true than another; it can only be more convenient. Now, Euclidean geometry is and will remain, the most convenient." After more analysis he added: "What we call a straight line in astronomy is simply the path of a ray of light. If, therefore, we were to discover negative parallaxes, or to prove that all parallaxes are higher than a certain limit, we should have a choice between two conclusions: we could give up Euclidean geometry, or modify the laws of optics, and suppose that light is not rigorously propagated in a straight line. It is needless to add that every one would look upon this solution as the more advantageous." Applied to GR, this amounts to the claim that any curved-space formulation of the theory can be replaced by an equivalent and simpler flat-space formulation. Ironically, the curved-space formulation has been preferred by nearly everyone since the inception of GR, and competing flat-space formulations certainly have not exhibited any of the simplicity anticipated by Poincaré. One wonders if the trend might have been different if Poincaré were still alive to promote his view when GR made its spectacular appearance on the scene.

The situation has been changed dramatically by the new gauge formulation of GR [1,2]. In retrospect, the popular "vierbein" approach to GR can be seen as a step toward a flat-space formulation, but it did not exhibit clear theoretical or practical simplifications until it was expressed in the language of geometric calculus (GC) in [5]. The superiority of GC in practical calculations involving the curvature tensor was also demonstrated there. By reinterpreting the GC formulation the new gauge theory has greatly clarified GR and produced numerou examples of computational simplifications. All this amounts to compelling evidence that Poincaré was right in the first place.

Part I. MATHEMATICAL FUNDAMENTALS

1. Spacetime Algebra

We represent Minkowski *spacetime* as a real 4-dimensional vector space $\mathcal{M}^4$. The two properties of scalar multiplication and vector addition in $\mathcal{M}^4$ provide only a partial specification of spacetime geometry. To complete the specification we introduce an associative *geometric product* among vectors $a, b, c, \ldots$ with the property that the square of any vector is a (real) scalar. Thus for any vector a we can write

$$a^2 = aa = \epsilon |\, a\,|^2 \,, \tag{1.1}$$

where ϵ is the *signature* of a and $|\, a\,|$ is a (real) positive scalar. As usual, we say that a is *timelike*, *lightlike* or *spacelike* if its signature is positive ($\epsilon = 1$), null ($\epsilon = 0$), or negative ($\epsilon = -1$). To specify the signature of $\mathcal{M}^4$ as a whole, we adopt the axioms: (a) $\mathcal{M}^4$ contains at least one timelike vector; and (b) Every 2-plane in $\mathcal{M}^4$ contains at least one spacelike vector.

The vector space $\mathcal{M}^4$ is not closed under the geometric product just defined. Rather, by multiplication and (distributive) addition it generates a real, associative (but noncommutative), geometric algebra of dimension $2^4 = 16$, called the *spacetime algebra* (STA). The name is appropriate because all the elements and operations of the algebra have a geometric interpretation, and it suffices for the representation of any geometric structure on spacetime.

From the *geometric product* ab of two vectors it is convenient to define two other products. The *inner product* $a \cdot b$ is defined by

$$a \cdot b = \tfrac{1}{2}(ab + ba) = b \cdot a \,, \tag{1.2}$$

and the *outer product* $a \wedge b$ is defined by

$$a \wedge b = \tfrac{1}{2}(ab - ba) = -b \wedge a \,. \tag{1.3}$$

The three products are therefore related by the important identity

$$ab = a \cdot b + a \wedge b \,, \tag{1.4}$$

which can be seen as a decomposition of the product ab into symmetric and antisymmetric parts.

From (1.1) it follows that the inner product $a \cdot b$ is scalar-valued. The outer product $a \wedge b$ is neither scalar nor vector but a new entity called a *bivector*, which can be interpreted geometrically as an oriented plane segment, just as a vector can be interpreted as an oriented line segment.

Inner and outer products can be generalized. The outer product and the notion of *k-vector* can be defined iteratively as follows: Scalars are regarded as 0-vectors, vectors as 1-vectors and bivectors as 2-vectors. For a given k-vector K the positive integer k is called the *grade* (or *step*) of K. The outer product of a vector a with K is a $(k+1)$-vector defined in terms of the geometric product by

$$a \wedge K = \tfrac{1}{2}(aK + (-1)^k Ka) = (-1)^k K \wedge a \,, \tag{1.5}$$

The corresponding inner product is defined by

$$a \cdot K = \tfrac{1}{2}(aK + (-1)^{k+1} Ka) = (-1)^{k+1} K \cdot a, \tag{1.6}$$

and it can be proved that the result is a $(k-1)$-vector. Adding (1.5) and (1.6) we obtain

$$aK = a \cdot K + a \wedge K \,, \tag{1.7}$$

which obviously generalizes (1.4). The important thing about (1.7) is that it decomposes aK into $(k-1)$-vector and $(k+1)$-vector parts.

Manipulations and inferences involving inner and outer products are facilitated by a host of theorems and identities given in [3] of which the most important are recorded here. The outer product is associative, and

$$a_1 \wedge a_2 \wedge \ldots \wedge a_k = 0 \tag{1.8}$$

if and only if the vectors $a_1, a_2, \ldots, a_k$ are linearly dependent. Since $\mathcal{M}^4$ has dimension 4, (1.8) is an identity in STA for $k > 4$, so the generation of new elements by multiplication with vectors terminates at $k = 4$. A nonvanishing k-vector can be interpreted as a directed volume element for $\mathcal{M}^k$ spanned by the vectors $a_1, a_2, \ldots, a_k$. In STA 4-vectors are called *pseudoscalars*, and for any four such vectors we can write

$$a_1 \wedge a_2 \wedge a_3 \wedge a_4 = \lambda i \,, \tag{1.9}$$

where i is the unit pseudoscalar and λ is a scalar which vanishes if the vectors are linearly dependent.

The *unit pseudoscalar* is so important that the special symbol i is reserved for it. It has the properties

$$i^2 = -1 \,, \tag{1.10}$$

and for every vector a in $\mathcal{M}^4$

$$ia = -ai\,. \tag{1.11}$$

Of course, i can be interpreted geometrically as *the* (unique) unit oriented volume element for spacetime. A convention for specifying its orientation is given below. Multiplicative properties of the unit pseudoscalar characterize the geometric concept of *duality*. The *dual* of a k-vector K in STA is the $(4-k)$-vector defined (up to a sign) by iK or Ki. Trivially, every pseudoscalar is the dual of a scalar. Every 3-vector is the dual of a vector; for this reason 3-vectors are often called *pseudovectors*. The inner and outer products are dual to one another. This is easily proved by using (1.7) to expand the associative identity $(aK)i = a(Ki)$ in two ways:

$$(a\cdot K + a\wedge K)i = a\cdot(Ki) + a\wedge(Ki)\,.$$

Each side of this identity has parts of grade $(4-k\pm1)$ and which can be separately equated, because they are linearly independent. Thus, one obtains the *duality identities*

$$(a\cdot K)i = a\wedge(Ki)\,, \tag{1.12a}$$

$$(a\wedge K)i = a\cdot(Ki)\,, \tag{1.12b}$$

Note that (1.12a) can be solved for

$$a\cdot K = [a\wedge(Ki)]i^{-1}\,, \tag{1.13}$$

which could be used to define the inner product from the outer product and duality.

Unlike the outer product, the inner product is not associative. Instead, it obeys various identities, of which the following involving vectors, k-vector K and s-vector B are most important:

$$(b\wedge a)\cdot K = b\cdot(a\cdot K) = (K\cdot b)\cdot a = K\cdot(b\wedge a) \qquad \text{for grade } k\geq 2\,, \tag{1.14}$$

$$a\cdot(K\wedge B) = (a\cdot K)\wedge B + (-1)^k K\wedge(a\cdot B)\,. \tag{1.15}$$

The latter implies the following identity involving vectors alone:

$$a\cdot(a_1\wedge a_2\wedge\ldots\wedge a_k) = (a\cdot a_1)a_2\wedge\ldots\wedge a_k - (a\cdot a_2)a_1\wedge a_3\ldots\wedge a_k + \cdots + (-1)^{k-1}(a\cdot a_k)a_1\wedge a_2\ldots\wedge a_{k-1}\,. \tag{1.16}$$

This is called the *Laplace expansion*, because it generalizes and implies the familiar expansion for determinants. The simplest case is used so often that it is worth writing down:

$$a\cdot(b\wedge c) = (a\cdot b)c - (a\cdot c)b = a\cdot b\,c - a\cdot c\,b\,. \tag{1.17}$$

Parentheses have been eliminated here by invoking a *precedence convention*, that in ambiguous algebraic expressions inner products are to be formed before outer products, and both of those before geometric products. This convention allows us to drop parentheses on the right side of (1.16) but not on the left.

The entire spacetime algebra is generated by taking linear combinations of k-vectors obtained by outer multiplication of vectors in $\mathcal{M}^4$. A generic element of the STA is called a *multivector*. Any multivector M can be written in the *expanded form*

$$M = \alpha + a + F + bi + \beta i = \sum_{k=0}^{4} \langle M \rangle_k \,, \tag{1.18}$$

where α and β are scalars, a and b are vectors and F is a bivector. This is a decomposition of M into its "k-vector parts" $\langle M \rangle_k$, where

$$\langle M \rangle_0 = \langle M \rangle = \alpha \tag{1.19}$$

is the scalar part, $\langle M \rangle_1 = a$ is the vector part, $\langle M \rangle_2 = F$ is the bivector part, $\langle M \rangle_3 = bi$ is the pseudovector part and $\langle M \rangle_4 = \beta i$ is the pseudoscalar part. Duality has been used to simplify the form of the trivector part in (1.18) by expressing it as the dual of a vector. Like the decomposition of a complex number into real and imaginary parts, the decomposition (1.18) is significant because the parts with different grade are linearly independent of one another and have distinct geometric interpretations. On the other hand, multivectors of mixed grade often have geometric significance that transcends that of their graded parts.

Any multivector M can be uniquely decomposed into parts of even and odd grade as

$$M_{\pm} = \tfrac{1}{2}(M \mp iMi) \,. \tag{1.20}$$

In terms of the expanded form (1.18), the *even part* can be written

$$M_+ = \alpha + F + \beta i \,, \tag{1.21}$$

while the *odd part* becomes

$$M_- = a + bi \,. \tag{1.22}$$

The set $\{M_+\}$ of all even multivectors forms an important subalgebra of STA called the *even subalgebra*. Spinors can be represented as elements of the even subalgebra.

Computations are facilitated by the operation of reversion, defined for arbitrary multivectors M and N by

$$(MN)^{\sim} = \widetilde{N}\widetilde{M} \,, \tag{1.23a}$$

with

$$\widetilde{a} = a \tag{1.23b}$$

for any vector a. For M in the expanded form (1.18), the reverse $\widetilde{M}$ is given by

$$\widetilde{M} = \alpha + a - F - bi + \beta i \,. \tag{1.24}$$

Note that bivectors and trivectors change sign under reversion while scalars and pseudoscalars do not.

This mathematical apparatus makes it possible to formulate and solve fundamental equations *without using coordinates*. Nevertheless, STA facilitates the manipulation of coordinates, as shown below and in later sections of this document. Let $\{\gamma_\mu; \mu = 0, 1, 2, 3\}$ be a *right-handed orthonormal frame* of vectors in $\mathcal{M}^4$ with γ_0 in the *forward light cone*. In accordance with (1.2), we can write

$$\eta_{\mu\nu} \equiv \gamma_\mu \cdot \gamma_\nu = \tfrac{1}{2}(\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu)\,. \tag{1.25}$$

This may be recognized as the defining condition for the "Dirac algebra," which is a matrix representation of STA over a complex number field instead of the reals. Although the present interpretation of the $\{\gamma_\mu\}$ as an orthonormal frame of vectors is quite different from their usual interpretation as matrix componentsof a single vector, it can be shown that these alternatives are in fact compatible.

The *orientation* of the *unit pseudoscalar* i relative to the frame $\{\gamma_\mu\}$ is set by the equation

$$i = \gamma_0\gamma_1\gamma_2\gamma_3 = \gamma_0 \wedge \gamma_1 \wedge \gamma_2 \wedge \gamma_3\,. \tag{1.26}$$

We shall refer to $\{\gamma_\mu\}$ as a *standard frame*, without necessarily associating it with the reference frame of a physical observer. For manipulating coordinates it is convenient to introduce the *reciprocal frame* $\{\gamma^\mu\}$ defined by the equations

$$\gamma_\mu = \eta_{\mu\nu}\gamma^\nu \qquad \text{or} \qquad \gamma_\mu \cdot \gamma^\nu = \delta_\mu^\nu\,. \tag{1.27}$$

(summation convention assumed). "Rectangular coordinates" $\{x^\nu\}$ of a spacetime point x are then given by

$$x^\nu = \gamma^\nu \cdot x \qquad \text{and} \qquad x = x^\nu\gamma_\nu\,. \tag{1.28}$$

The γ_μ generate by multiplication a complete basis for STA, consisting of the $2^4 = 16$ linearly independent elements

$$1, \quad \gamma_\mu, \quad \gamma_\mu \wedge \gamma_\nu, \quad \gamma_\mu i, \quad i\,. \tag{1.29}$$

Any multivector can be expressed as a linear combination of these elements. For example, a bivector F has the expansion

$$F = \tfrac{1}{2}F^{\mu\nu}\gamma_\mu \wedge \gamma_\nu\,, \tag{1.30}$$

with its "scalar components" given by

$$F^{\mu\nu} = \gamma^\mu \cdot F \cdot \gamma^\nu = \gamma^\nu \cdot (\gamma^\mu \cdot F) = (\gamma^\nu \wedge \gamma^\mu) \cdot F\,. \tag{1.31}$$

However, such an expansion is seldom needed.

Besides the inner and outer products defined above, many other products can be defined in terms of the geometric product. The *commutator product* $A \times B$ is defined for any A and B by

$$A \times B \equiv \tfrac{1}{2}(AB - BA) = -B \times A\,. \tag{1.32}$$

Mathematicians classify this product as a "derivation" with respect to the geometric product, because it has the "distributive property"

$$A \times (BC) = (A \times B)C + B(A \times C)\,. \tag{1.33}$$

This implies the *Jacobi Identity*

$$A \times (B \times C) = (A \times B) \times C + B \times (A \times C)\,, \tag{1.34}$$

which is a derivation on the commutator product. The relation of the commutator product to the inner and outer products is grade dependent; thus, for a vector a,

$$a \times \langle\, M \,\rangle_k = a \wedge \langle\, M \,\rangle_k \qquad \text{if } k \text{ is odd}\,, \tag{1.35a}$$

$$a \times \langle\, M \,\rangle_k = a \cdot \langle\, M \,\rangle_k \qquad \text{if } k \text{ is even}\,. \tag{1.35b}$$

The commutator product is especially useful in computations with bivectors. With any bivector A this product is grade preserving:

$$A \times \langle\, M \,\rangle_k = \langle\, A \times M \,\rangle_k\,. \tag{1.36}$$

In particular, this implies that the space of bivectors is closed under the commutator product. It therefore forms a Lie algebra — which is, in fact, the Lie algebra of the Lorentz group. The geometric product of bivector A with M has the expanded form

$$AM = A \cdot M + A \times M + A \wedge M \qquad \text{for grade } M \geq 2\,. \tag{1.37}$$

This should be compared with the corresponding expansion (1.4) for the product of vectors.
For any bivector F, (1.37) assures us that we can write

$$F^2 = F \cdot F + F \wedge F = |\, f \,|^2 e^{i2\varphi}\,.$$

When $F^2 \neq 0$, this can be solved for the "invariant angle φ," given by

$$i \tan 2\varphi = \frac{F \wedge F}{F \cdot F}\,. \tag{1.38}$$

Then F can be written in the *canonical form*

$$F = f e^{i\varphi}\,. \tag{1.39}$$

When $F^2 = 0$, F can still be written in the form (1.39) with

$$f = k \wedge e = ke\,, \tag{1.40}$$

where k is a null vector and e is spacelike. In this case, the decomposition is not unique, and the exponential factor can always be absorbed in the definition of f. The canonical bivector decomposition (1.39) is especially useful in treating the electromagnetic field.

2. Vector Derivatives and Differentials

To extend spacetime algebra into a complete *spacetime calculus*, suitable definitions for derivatives and integrals are required. In terms of the coordinates (1.28), an operator $\nabla \equiv \partial_x$ interpreted as the *derivative* with respect to a spacetime point x can be defined by

$$\nabla = \gamma^\mu \partial_\mu \tag{2.1}$$

where

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \gamma_\mu \cdot \nabla \,. \tag{2.2}$$

The square of ∇ is the d'Alembertian

$$\nabla^2 = \eta^{\mu\nu}\partial_\mu\partial_\nu \,. \tag{2.3}$$

The matrix representation of (2.1) can be recognized as the "Dirac operator," originally discovered by Dirac by seeking a "square root" of the d'Alembertian (2.3) in order to find a first order "relativistically invariant" wave equation for the electron. In STA, however, where the γ^μ are vectors rather than matrices, it is clear that ∇ is a vector operator; indeed, it provides an appropriate definition for the derivative with respect to any spacetime vector variable.

Contrary to the impression given by conventional accounts of relativistic quantum theory, the operator ∇ is not specially adapted to spin-$\frac{1}{2}$ wave equations. It is equally apt for electromagnetic field equations: in STA an electromagnetic field is represented by a bivector-valued function $F = F(x)$ on spacetime. The field produced by a source with spacetime current density $J = J(x)$ is determined by *Maxwell's Equation*

$$\nabla F = J \,. \tag{2.4}$$

Since ∇ is a vector operator the identity (1.7) applies, so that we can write

$$\nabla F = \nabla \cdot F + \nabla \wedge F \,, \tag{2.5}$$

where $\nabla \cdot F$ is the *divergence* of F and $\nabla \wedge F$ is the *curl*. We can accordingly separate (2.4) into vector and trivector parts:

$$\nabla \cdot F = J \,, \tag{2.6a}$$

$$\nabla \wedge F = 0 \,. \tag{2.6b}$$

If F is derived from a vector potential $A = A(x)$ satisfying the "Lorentz condition" $\nabla \cdot A = 0$, then $F = \nabla \wedge A = \nabla A$ and (1.35) becomes the familiar "wave equation"

$$\nabla^2 A = J \,. \tag{2.7}$$

This demonstration illustrates the simplicity that the *vector derivative* ∇ brings to the formulation of basic equations in physics. However, to make it an efficient computational tool, its properties must be developed more fully.

As in conventional scalar calculus, the derivatives of elementary functions are often needed for computations. The most useful derivatives are listed here:

Table of vector derivatives: (2.8)

$$\nabla(x \cdot a) = a \cdot \nabla x = a \,, \qquad \nabla(x \cdot A) = A \cdot \nabla x = rA$$

$$\nabla|\,x\,|^2 = \nabla x^2 = 2x \,, \qquad \nabla(x \wedge A) = A \wedge \nabla x = (4-r)A$$

$$\nabla \wedge x = 0 \,, \qquad \nabla(Ax) = \gamma^\mu A \gamma_\mu = (-1)^r(4-2r)A \,,$$

$$\nabla x = \nabla \cdot x = 4$$

$$\nabla |\, x\,|^k = k|\, x\,|^{k-2}x\,, \qquad \nabla\left(\frac{x}{|\, x\,|^k}\right) = \frac{4-k}{|\, x\,|^k}\,.$$

In the table, $\nabla = \partial_x$, and obvious singularities at $x = 0$ are excluded; a is a "*free vector*" variable (i.e. independent of x), while A is a free r-vector.

The *directional derivative* $a \cdot \nabla$, that is, the "derivative in the direction of vector a" can be obtained from ∇ by applying the inner product. Alternatively, the directional derivative can be defined directly by

$$a \cdot \nabla F = a \cdot \partial_x F(x) = \frac{d}{d\tau}\, F(x + a\tau)\Big|_{\tau=0} = \lim_{\tau \to 0} \frac{F(x + a\tau) - F(x)}{\tau}\,, \tag{2.9}$$

where $F = F(x)$ is any multivector valued function. Then the general vector derivative can be obtained from the directional derivative by using (2.8): thus,

$$\nabla F = \partial_x F(x) = \partial_a a \cdot \partial_x F(x)\,. \tag{2.10}$$

This relation can serve as an alternative to the partial derivative in (2.1) for defining the vector derivative. Of course, the directional derivative has the same limit properties as the partial derivative, including the rules for differentiating sums and products, but the explicit display of the vector variable is advantageous in concept and calculation.

Equation (2.10) and the preceding equations above define the derivative ∂_a with respect to any spacetime vector a. As already indicated, we reserve the symbol "x" for a vector representing a position in spacetime, and we use the special symbol $\nabla \equiv \partial_x$ for the derivative with respect to this "spacetime point." When differentiating with respect to any other vector variable a, we indicate the variable explicitly by writing ∂_a. Mathematically, there is no difference between ∇ and ∂_a. However, there is often an important difference in physical or geometrical interpretation associated with these operators.

The directional derivative (2.9) produces from F a tensor field called the *differential* of F, denoted variously by

$$\underline{F}(a) = F_a \equiv a \cdot \nabla F\,. \tag{2.11}$$

As explained in the next section, the underbar notation serves to emphasize that $\underline{F}(a)$ is a linear function of a, though it is not a linear transformation unless it is vector valued. The argument a may be a free variable or a vector field $a = a(x)$.

The *second differential* $\underline{F}_b(a) = F_{ab}$ is defined by

$$\underline{F}_b(a) \equiv b \cdot \nabla \underline{F}(a) - \underline{F}(b \cdot \nabla a) = b \cdot \grave{\nabla} \grave{\underline{F}}(a)\,, \tag{2.12}$$

where the accents over $\grave{\nabla}$ and $\grave{F}$ serve to indicate that only F is differentiated and a is not. Of course, the accents can be dropped if a is a free variable. The second differential has the symmetry property

$$\underline{F}_b(a) = \underline{F}_a(b)\,. \tag{2.13}$$

Using (1.14) and (1.17), this can be expressed as an operator identity:

$$(a \wedge b) \cdot (\nabla \wedge \nabla) = [a \cdot \nabla, b \cdot \nabla] = 0\,, \tag{2.14}$$

where the bracket denotes the commutator. Differentation by ∂_a and ∂_b puts this identity in the form

$$\nabla \wedge \nabla = 0\,. \tag{2.15}$$

These last three equations are equivalent formulations of the *integrability condition* for vector derivatives.

Since the derivative ∇ has the algebraic properties of a vector, a large assortment of "differential identities" can be derived by replacing some vector by ∇ in any algebraic identity. The only caveat is to take proper account of the product rule for differentiation. For example, the product rule gives

$$\nabla \cdot (a \wedge b) = \grave{\nabla} \cdot (\grave{a} \wedge b) + \grave{\nabla} \cdot (a \wedge \grave{b})\,,$$

whence the algebraic identity (1.17) yields

$$a \cdot \nabla b - b \cdot \nabla a = \nabla \cdot (a \wedge b) + a \nabla \cdot b - b \nabla \cdot a\,, \tag{2.16}$$

The left side of this identity may be identified as a Lie bracket; a more general concept of the Lie bracket is introduced later on. Other identities will be derived as needed.

3. Linear Algebra

The computational and representational power of linear algebra is greatly enhanced by integrating it with geometric algebra. In fact, geometric calculus makes it possible to perform coordinate-free computations in linear algebra without resorting to matrices. Integration of the two algebras can be achieved with the few basic concepts, notations and theorems reviewed below. However, linear algebra is a large subject, so we restrict our attention to the essentials needed for gravitation theory.

Though the approach works for vector spaces of any dimension, we will be concerned only with linear transformations of Minkowski space, which map spacetime vectors into vectors. We need a notation which clearly distinguishes linear operators and their products from vectors and their products. To this end, we distinguish symbols representing a linear transformation (or operator) by affixing them with an underbar (or overbar). Thus, for a linear operator $\underline{f}$ acting on a vector a, we write

$$\underline{f}a = \underline{f}(a)\,. \tag{3.1}$$

As usual in linear algebra, the parenthesis around the argument of $\underline{f}$ can be included or omitted, either for emphasis or to remove ambiguity.

Every linear transformation $\underline{f}$ on Minkowski space has a unique extension to a linear function on the whole STA, called the *outermorphism* of $\underline{f}$ because it preserves outer products. It is convenient to use the same notation $\underline{f}$ for the outermorphism and the operator that "induces" it, distinguishing them when necessary by their arguments. The outermorphism is defined by the property

$$\underline{f}(A \wedge B) = (\underline{f}A) \wedge (\underline{f}B) \tag{3.2}$$

for arbitrary multivectors A, B, and

$$\underline{f}\alpha = \alpha \tag{3.3}$$

for any scalar α. It follows that, for any factoring $A = a_1 \wedge a_2 \wedge \ldots \wedge a_r$ of an r-vector A into vectors,

$$\underline{f}(A) = (\underline{f}a_1) \wedge (\underline{f}a_2) \wedge \ldots \wedge (\underline{f}a_r)\,. \tag{3.4}$$

This relation can be used to compute the outermorphism directly from the inducing linear operator.

Since the outermorphism preserves the outer product, it is grade preserving, that is

$$\underline{f}\langle M \rangle_k = \langle \underline{f}M \rangle_k \tag{3.5}$$

for any multivector M. This implies that $\underline{f}$ alters the pseudoscalar i only by a scalar multiple. Indeed

$$\underline{f}(i) = (\det \underline{f})i \qquad \text{or} \qquad \det \underline{f} = -i\,\underline{f}(i)\,, \tag{3.6}$$

which defines the *determinant* of $\underline{f}$. Note that the outermorphism makes it possible to define (and evaluate) the determinant without introducing a basis or matrices.

The "product" of two linear transformations, expressed by

$$\underline{h} = \underline{g}\,\underline{f}\,, \tag{3.7}$$

applies also to their outermorphisms. In other words, the outermorphism of a product equals the product of outermorphisms. It follows immediately from (3.6) that

$$\det(\underline{g}\,\underline{f}) = (\det \underline{g})(\det \underline{f})\,, \tag{3.8}$$

from which many other properties of the determinant follow, such as

$$\det(\underline{f}^{-1}) = (\det \underline{f})^{-1} \tag{3.9}$$

whenever $\underline{f}^{-1}$ exists.

Every linear transformation $\underline{f}$ has an *adjoint* transformation $\bar{f}$ which can be extended to an outermorphism denoted by the same symbol. The adjoint outermorphism can be defined in terms of $\underline{f}$ by

$$\langle M\bar{f}N \rangle = \langle N\underline{f}M \rangle\,, \tag{3.10}$$

where M and N are arbitrary multivectors and the bracket as usual indicates "scalar part." For vectors a, b this can be written

$$b \cdot \bar{f}(a) = a \cdot \underline{f}(b)\,. \tag{3.11}$$

Differentiating with respect to b we obtain, with the help of (2.8),

$$\bar{f}(a) = \partial_b\, a \cdot \underline{f}(b)\,. \tag{3.12}$$

This is the most useful formula for obtaining $\bar{f}$ from $\underline{f}$. Indeed, it might well be taken as the preferred definition of $\bar{f}$.

Unlike the outer product, the inner product is not generally preserved by outermorphisms. However, it obeys the fundamental transformation law

$$\bar{f}[\underline{f}(A) \cdot B] = A \cdot \bar{f}(B) \tag{3.13}$$

for (grade A) $\leq$ (grade B). Of course, this law also holds with an interchange of overbar and underbar. If $\underline{f}$ is invertible, it can be written in the form

$$\bar{f}[A\cdot B] = \underline{f}^{-1}(A)\cdot\bar{f}(B)\,. \tag{3.14}$$

For $B = i$, since $A\cdot i = Ai$, this immediately gives the general formula for the inverse outermorphism:

$$\underline{f}^{-1}A = [\bar{f}(Ai)](\bar{f}\,i)^{-1} = (\det\underline{f})^{-1}\bar{f}(Ai)i^{-1}\,. \tag{3.15}$$

This relation shows explicitly the double duality involved in computing the inverse.

Although all linear transformations preserve the outer product (by definition of the outermorphism (3.2)), only a restricted class preserves the inner product. This is called the *Lorentz group*, and its members are called *Lorentz transformations*. The defining property for a Lorentz transformation $\underline{L}$ is

$$(\underline{L}a)\cdot(\underline{L}b) = a\cdot(\bar{L}\underline{L}b) = a\cdot b\,. \tag{3.16}$$

This is equivalent to the operator condition

$$\bar{L} = \underline{L}^{-1} \tag{3.17}$$

STA makes it possible to express any $\underline{L}$ in the simple *canonical form*

$$\underline{L}(a) = \epsilon L a L^{-1}\,, \tag{3.18}$$

where the multivector L is either *even* with $\epsilon = 1$ or *odd* with $\epsilon = -1$. This defines a double-valued homomorphism between Lorentz transformations $\{\underline{L}\}$ and multivectors $\{\pm L\}$, where the composition of two Lorentz transformations $\underline{L}_1\underline{L}_2$ corresponds to the geometric product $\pm L_1L_2$. Thus, the Lorentz group has a double-valued representation as a multiplicative group of multivectors. This *multivector representation* of the Lorentz group greatly facilitates the analysis and application of Lorentz transformations in STA.

From (3.18) it follows immediately that, for arbitrary multivectors A and B,

$$\underline{L}(AB) = (\underline{L}A)(\underline{L}B)\,. \tag{3.19}$$

Lorentz transformations therefore preserve the geometric product. This implies that (3.16) generalizes to

$$\underline{L}(A\cdot B) = (\underline{L}A)\cdot(\underline{L}B)\,. \tag{3.20}$$

in agreement with (3.14) when (3.17) is satisfied.

From (3.18) it follows easily that

$$\underline{L}(i) = \epsilon i\,, \qquad \text{where} \qquad \epsilon = \det\underline{L} = \pm 1\,. \tag{3.21}$$

A Lorentz transformation $\underline{L}$ is said to be *proper* if $\epsilon = 1$, and *improper* if $\epsilon = -1$. It is said to be *orthochronous* if, for any timelike vector v,

$$v\cdot\underline{L}(v) > 0\,. \tag{3.22}$$

A *proper, orthochronous* Lorentz transformation is called a *Lorentz rotation* (or a *restricted* Lorentz transformation). For a Lorentz rotation $\underline{R}$ the canonical form can be written

$$\underline{R}(a) = Ra\widetilde{R}\,, \tag{3.23}$$

where the even multivector R is called a *rotor* and is normalized by the condition

$$R\widetilde{R} = 1\,. \tag{3.24}$$

The rotors form a multiplicative group called the *Rotor group*, which is a double-valued representation of the Lorentz rotation group (also called the restricted Lorentz group).

The most elementary kind of Lorentz transformation is a *reflection* $\underline{n}$ by a (non-null) vector n, according to

$$\underline{n}\,(a) = -nan^{-1}\,. \tag{3.25}$$

This is a reflection with respect to a hyperplane with normal n. A reflection

$$\underline{v}(a) = -vav \tag{3.26}$$

with respect to a timelike vector $v = v^{-1}$ is called a *time reflection*. Let n_1, n_2, n_3 be spacelike vectors which compose the trivector

$$n_3 n_2 n_1 = iv\,. \tag{3.27}$$

A *space inversion* $\underline{v}_s$ can then be defined as the composite of reflections with respect to these three vectors, so it can be written

$$\underline{v}_s(a) = n_3 n_2 n_1 a n_1 n_2 n_3 = ivavi = vav\,. \tag{3.28}$$

Note the difference in sign between the right sides of (3.26) and (3.28). Although $\underline{v}_s$ is determined by v alone on the right side of (3.28), the multivector representation of $\underline{v}_s$ must be the trivector iv in order to agree with (3.18). The composite of the time reflection (3.26) with the space inversion (3.28) is the *spacetime inversion*

$$\underline{v}_{st}(a) = \underline{v}_s v(a) = -iai^{-1} = -a\,, \tag{3.29}$$

which is represented by the pseudoscalar i. Note that spacetime inversion is proper but not orthochronous.

Two basic types of Lorentz rotation can be obtained from the product of two reflections, namely *timelike rotations* (or *boosts*) and *spacelike rotations*. For a *boost*

$$\underline{V}\,(a) = Va\widetilde{V}\,, \tag{3.30}$$

the rotor V can be factored into a product

$$V = v_2 v_1 \tag{3.31}$$

of two unit timelike vectors v_1 and v_2. The boost is a rotation in the timelike plane containing v_1 and v_2. The factorization (3.31) is not unique. Indeed, for a given V any timelike vector in the plane can be chosen as v_1, and v_2 then computed from (3.31). Similarly, for a *spacelike rotation*

$$\underset{\sim}{Q}(a) = Qa\widetilde{Q}\,, \tag{3.32}$$

the rotor Q can be factored into a product

$$Q = n_2 n_1 \tag{3.33}$$

of two unit spacelike vectors in the *spacelike plane* of the rotation.

Note that the product, say $n_2 v_1$, of a spacelike vector with a timelike vector is not a rotor, because the corresponding Lorentz transformation is not orthochronous. Likewise, the pseudoscalar i is not a rotor, even though it can be expressed as the product of two pairs of vectors, for it does not satisfy the rotor condition (3.24).

Any Lorentz rotation $\underline{R}$ can be decomposed uniquely into the product

$$\underline{R} = \underline{V}\,\underset{\sim}{Q} \tag{3.34}$$

of a boost $\underline{V}$ and spacelike rotation $\underset{\sim}{Q}$ with respect to a given timelike vector $v_0 = v_0^{-1}$. To compute $\underline{V}$ and $\underset{\sim}{Q}$ from $\underline{R}$, first compute the vector

$$v = \underline{R}v_0 = Rv_0\widetilde{R}\,. \tag{3.35}$$

The timelike vectors v and v_0 determine the timelike plane for the boost, which can therefore be defined by

$$v = \underline{V}v_0 = Vv_0\widetilde{V} = V^2 v_0\,. \tag{3.36}$$

This can be solved for

$$V = \big(vv_0\big)^{\frac{1}{2}} = vw = wv_0\,, \tag{3.37a}$$

where the unit vector

$$w = \frac{v+v_0}{|\,v+v_0\,|} = \frac{v+v_0}{\Big[2(1+v\cdot v_0)\Big]^{\frac{1}{2}}} \tag{3.37b}$$

"bisects the angle" between v and v_0. The rotor Q can then be computed from

$$Q = \widetilde{V}R\,, \tag{3.38}$$

so that the spacelike rotation satisfies

$$\underset{\sim}{Q}v_0 = Qv_0\widetilde{Q} = v_0\,. \tag{3.39}$$

This makes (3.36) consistent with (3.35) by virtue of (3.34).

Equations (3.31) and (3.32) show how to parametrize boosts and spacelike rotations by vectors in the plane of rotation. More specifically, (3.37a,b) parametrizes a boost in terms of initial and final velocity vectors. This is especially useful, because the velocity vectors are often given, or are of direct relevance, in a physical problem. Another useful parametrization is in terms of angle (Appendix B of [4]). Any rotor R can be expressed in the exponential form

$$\pm R = e^{\frac{1}{2}F} = \sum_{k=0}^{\infty}\frac{1}{k!}\left(\tfrac{1}{2}F\right)^k, \tag{3.40}$$

where F is a bivector parametrizing the rotation. The positive sign can always be selected when $F^2 \neq 0$, and, according to (1.39), F can be written in the canonical form

$$F = (\alpha + i\beta)f \qquad \text{where} \qquad f^2 = 1, \tag{3.41}$$

α and β being scalar parameters. Since the timelike unit bivector f commutes with its dual if, which is a spacelike bivector, the rotor R can be decomposed into a product of commuting timelike and spacelike rotors. Thus

$$R = VQ = QV\,, \tag{3.42}$$

where

$$V = e^{\frac{1}{2}\alpha f} = \cosh \tfrac{1}{2}\alpha + f \sinh \tfrac{1}{2}\alpha\,, \tag{3.43}$$

and

$$Q = e^{\frac{1}{2}i\beta f} = \cos \tfrac{1}{2}\beta + if \sin \tfrac{1}{2}\beta\,. \tag{3.44}$$

The parameter α is commonly called the *rapidity* of the boost. The parameter β is the usual angular measure of a spatial rotation.

When $F^2 = 0$, equation (3.40) can be reduced to the form

$$R = e^{\frac{1}{2}\alpha f} = 1 + \tfrac{1}{2}\alpha f\,, \tag{3.45}$$

where f is a null bivector, and it can be expressed in the factored form (1.40). The two signs are inequivalent cases. There is no choice of null F which can eliminate the minus sign. The *lightlike rotor* in (3.45) represents a *lightlike Lorentz rotation*.

The spacelike rotations that preserve a timelike vector v_0 are commonly called *spatial rotations* without mentioning the proviso (3.38). The set of such rotations is the 3-parameter *spatial rotation group* (of v_0). More generally, for any given vector n, the subgroup of Lorentz transformations $\underline{N}$ satisfying

$$\underline{N}(n) = Nn\widetilde{N} = n \tag{3.46}$$

is called the *little group* of n. The little group of a lightlike vector can be parametrized as a lightlike rotor (3.40) composed with timelike and spacelike rotors.

The above classification of Lorentz transformations can be extended to more general linear transformations. For any linear transformation $\underline{f}$ the composite $\bar{f}\underline{f}$ is a symmetric transformation. If the latter has a well defined square root $\underline{S} = (\bar{f}\underline{f})^{\frac{1}{2}} = \overline{S}$, then $\underline{f}$ admits to the *polar decomposition*

$$\underline{f} = \underline{R}\,\underline{S} = \underline{S}'\underline{R}\,, \tag{3.47}$$

where $\underline{R}$ is a Lorentz rotation and $\underline{S}' = \underline{R}\,\underline{S}\,\underline{R}^{-1}$. Symmetric transformations are, of course, defined by the condition

$$\overline{S} = \underline{S}\,. \tag{3.48}$$

On Euclidean spaces every linear transformation has a polar decomposition, but on Minkowski space there are symmetric transformations (with null eigenvectors) which do not possess square roots and so do not admit a polar decomposition. A complete classification of symmetric transformations is given in the next section.

4. Tensors and their classification

Our development of the "tensor concept" is somewhat unconventional, to allow us to integrate it into the spacetime calculus and take full advantage of the algebra. We begin with a definition which is more general than the usual one.

A *tensor* T of *degree* k is a multilinear, multivector-valued function $T(a_1, a_2, \ldots, a_k)$ of k vectors $a_1, a_2, \ldots, a_k$. This means that it is a linear function of each vector variable separately. If it is skewsymmetric under interchange of any pair of variables, it can be regarded as a linear function L of a single k-vector variable; thus

$$L(a_1 \wedge a_2 \wedge \ldots \wedge a_k) = T(a_1, a_2, \ldots, a_k). \tag{4.1}$$

This kind of function is called a *k-form*. Ordinarily k-forms are scalar-valued. The term *multiform* can be used for the present generalization to multivector-valued functions. In general, then, a k-form

$$L(A) = L(\langle A \rangle_k) \tag{4.2}$$

is a *multiform of degree* k. This kind of function is fundamental to the theory of integration set out in Section 6.

In ordinary tensor algebra, *tensor addition* is defined only for tensors with the same rank and degree. A *tensor* T *of degree* k has a definite *rank* $r > k$ if it is $(r-k)$-vector valued, that is, if

$$T(a_1, a_2, \ldots, a_k) = \langle T(a_1, a_2, \ldots, a_k) \rangle_{r-k}. \tag{4.3}$$

We call this a tensor of type $r-k$. All the tensors ordinarily considered in physics are of this type. Among the tensors to be encountered in later sections, the energy-momentum tensor $T(a)$ and the gauge tensor $\underline{h}(a)$ are of type 2-1; the connection $\omega(a)$ is of type 3-1; and the curvature tensor $R(a \wedge b)$ is of type 4-2. Of course, every multivector is trivially a tensor of degree zero. Every scalar-valued k-form is equivalent to a k-vector, as can be proved using relations laid out below.

Conventional tensor algebra employs the concept of rank alone, while the property of degree is attributed to forms alone; moreover, tensors and forms are treated as completely separate entities. Geometric algebra brings them under one roof. Two tensors (of degree r and s, say) can be multiplied using any of the various geometric products to form a new tensor of degree $r+s$. The conventional *tensor product* of "ranked tensors" is a tensor with rank equal to the sum of the ranks of the multiplicands, without any particular assignment of degree. Since the tensor product can always be expressed in terms of the geometric product [3], it need not be introduced as a separate concept. As conceived here a tensor is a purely algebraic object, without any of the transformation rules attached to its arguments as in conventional tensor analysis. From the algebraic viewpoint, the linear transformations and outermorphisms of the preceding section are all tensors. Of course, the "product" of linear transformations by functional composition is not a tensor product, although construction of an outermorphism by outer products is. It should also be recognized that the terms "tensor" and "linear transformation" often indicate an important physical distinction when applied to objects of identical algebraic structure. As a rule, physicists apply the term *tensor* to quantities representing some property of a physical system, while the term *transformation* is applied to changes in properties or their representation. For the purely algebraic purposes of this section such a distinction is irrelevant. The coupling of tensors to transformations will be considered later.

Tensors inherit algebraic properties automatically from their definition in terms of STA. The only problem outstanding is to give suitable names to these properties to facilitate description. The inner product $a \cdot T(a_1, \ldots)$ with a vector a reduces the grade of T but preserves its rank. In fact, successive inner products with $s = r - k$ vectors produces a tensor of grade zero:

$$(a_1, \ldots, a_s) \cdot T(a_{s+1}, \ldots, a_r) . \tag{4.4}$$

We can regard (4.3) and (4.4) as different forms of one and the same tensor, for (4.3) can be recovered from (4.4) by successive protractions. The *protraction* of T (with respect to the first variable) is defined by

$$\partial_a \wedge T(a, \ldots) . \tag{4.5}$$

Unless the result vanishes, protraction raises the grade but preserves the rank. *Contraction*, defined by

$$\partial_a \cdot T(a, \ldots) , \tag{4.6}$$

lowers the rank by 2 and the grade by 1. Of course, this operation is not defined if the grade of T is zero.

It will be recognized that the *protraction* (4.5) is just a curl while the contraction (4.6) is just a divergence. This introduction of new names is justified, nevertheless, by their unique application to multilinear functions. In particular, note that these operations completely remove a tensor's functional dependence on one variable. Another type of *contraction*, which eliminates two explicit variables, is defined by

$$\partial_a \cdot \partial_b T(\ldots a, \ldots, b \ldots) = T(\ldots \partial_a, \ldots, a \ldots) = T(\ldots \gamma_\mu, \ldots, \gamma^\mu \ldots) . \tag{4.7}$$

This is equivalent to the standard definition of contraction in tensor analysis. Note that the contraction using the derivative is equivalent to the sum over a frame and its reciprocal, as defined by (1.27). In later sections this type of contraction will arise mainly from combinations of protractions and contractions of the "divergence type" (4.6), to which it is equivalent in some cases.

We now have enough mathematical apparatus for systematic classification of all tensors. However, we restrict our attention to those tensors of importance in gravitation theory — in particular, the energy-momentum tensor and the curvature tensor. Algebraic classifications of these tensors have been thoroughly worked out in the literature [16]. However, it will be seen that STA brings a new perspective to the subject. The algebraic classification of curvature tensors will be discussed in Section 11. The rest of this section is devoted to the classification of type 2-1 tensors; this includes all linear transformations as well as the energy-momentum tensor.

For a type 2–1 tensor $\underline{f}$, protraction produces an "invariant" bivector

$$F = \partial_a \wedge \underline{f}(a) , \tag{4.8}$$

while the contraction produces the trace

$$f_0 \equiv \mathrm{Tr}\, \underline{f} = \partial_a \cdot \underline{f}(a) . \tag{4.9}$$

The inner product of (4.8) with a vector recovers the skewsymmetric part of $\underline{f}$:

$$\underline{f}_-(a) = \tfrac{1}{2}\big[\,\underline{f}(a) - \bar{f}(a)\big] = \tfrac{1}{2}a\cdot F\,. \tag{4.10}$$

From the symmetric part

$$\underline{f}_+(a) = \tfrac{1}{2}\big[\,\underline{f}(a) + \bar{f}(a)\big]\,, \tag{4.11}$$

we can extract a traceless part

$$\underline{f}_\oplus = \underline{f}_+ - \frac{1}{n}\,f_0\underline{1}\,, \tag{4.12}$$

where $\underline{1}$ is the identity operator and $n = \mathrm{Tr}\,\underline{1}$. Putting all the parts together, we have the well known invariant decomposition

$$\underline{f} = \underline{f}_\oplus + \frac{1}{n}\,f_0\underline{1} + \underline{f}_-\,. \tag{4.13}$$

This kind of decomposition is especially useful when the bivector F in (4.8) and (4.10) has a direct physical interpretation. An important example, to be considered in Part III, is the deformation rate tensor, for which the protraction can be identified with vorticity.

As an alternative to the decomposition (4.13), the polar decomposition (3.47) is significant in other physical problems. In both cases a symmetric tensor appears — additively in one case, and multiplicatively in the other. A complete classification of symmetric tensors involves much more than simply separating off a traceless part as in (4.12). There are three major types, as is now described.

Type I

A symmetric tensor $\underline{S}$ is said to be of *type I* if it admits to the *"spectral decomposition"*

$$\underline{S} = \sum_k \lambda_k\,\underline{e}_k\,, \tag{4.14}$$

where the $\lambda_k = \langle\,\lambda_k\,\rangle_0$ are real scalars and the $\underline{e}_k$ are irreducible projection operators with the properties

Idempotence:	$\underline{e}_k^2 = \underline{e}_k,$	(4.15a)
Orthogonality:	$\underline{e}_j\,\underline{e}_k = 0 \qquad \text{for} \qquad j \neq k,$	(4.15b)
Completeness:	$\sum_k \underline{e}_k = \underline{1},$	(4.15c)
Eigenvectors:	$\underline{e}_j(e_k) = \delta_{jk}e_k.$	(4.15d)

Therefore $\underline{S}$ has a unique set of eigenvectors e_k satisfying

$$\underline{S}e_k = \lambda_k e_k\,. \tag{4.16}$$

The eigenvectors compose an orthogonal basis $\{e_0,\, e_1,\, e_2,\, e_3\}$ for Minkowski space.

Type I tensors $\underline{S}$ and $\underline{S}'$, related by a Lorentz rotation $\underline{R}$ according to

$$\underline{S}' = \underline{R}\,\underline{S}\,\underline{R}^{-1} = \sum_k \lambda_k \underline{e}'_k\,, \tag{4.17}$$

are equivalent in the sense that they have the same spectrum of eigenvalues $\{\lambda_k\}$. Subtypes are distinguished by special properties of the eigenvalues, such as positivity or degeneracy. This gives a complete classification of type I tensors.

Geometric algebra makes it possible to derive the projection operators from the eigenvectors; thus,

$$\underline{e}_k(a) = e_k^{-1} e_k \cdot a = e_k e_k^{-1} \cdot a = \tfrac{1}{2}(a + e_k a e_k^{-1})\,. \tag{4.18}$$

Also, outermorphisms facilitate the analysis and applications of symmetric transformations. For example, the bivectors

$$e_{ij} \equiv e_i \wedge e_j = e_i e_j \tag{4.19}$$

can be regarded as *eigenbivectors* of $\underline{S}$ satisfying (for $i \neq j$)

$$\underline{S}(e_{ij}) = \lambda_i \lambda_j e_{ij}\,. \tag{4.20}$$

Similarly, there are *eigentrivectors*

$$\underline{S}(e_{ijk}) = \lambda_i \lambda_j \lambda_k e_{ijk}\,. \tag{4.21}$$

Finally, for the pseudoscalar i we have

$$\underline{S}(i) = \lambda_0\lambda_1\lambda_2\lambda_3 i\,, \qquad \text{so} \qquad \det \underline{S} = \lambda_0\lambda_1\lambda_2\lambda_3\,. \tag{4.22}$$

For *degenerate eigenvalues* $\lambda_i = \lambda_j$, the spectral form (4.14) can be simplified by introducing the bivector projection

$$\underline{e}_{ij}(a) \equiv e_{ij}^{-1}(e_{ij} \cdot a) = \underline{e}_i(a) + \underline{e}_j(a), \tag{4.23}$$

so that two terms in the sum are reduced to one according to

$$\lambda_i \underline{e}_i + \lambda_j \underline{e}_j = \lambda_i \underline{e}_{ij}\,, \tag{4.24}$$

and the arbitrariness in choosing eigenvectors in the e_{ij}-plane is eliminated. Note the outermorphism properties

$$\underline{e}_k(a \wedge b) = 0\,, \tag{4.25}$$

$$\underline{e}_{ij}(a \wedge b) = e_{ij}{}^{-1} e_{ij} \cdot (a \wedge b)\,, \tag{4.26}$$

$$\underline{e}_{ij}(a \wedge b \wedge c) = 0\,. \tag{4.27}$$

For vector spaces with a Euclidean inner product, all symmetric tensors are of type I. However, for spacetime the existence of null vectors leads to other types.

Before continuing the classification, it is convenient to make a short excursion into the properties of null vectors. Two null vectors k, k^* are said to be a *conjugate pair* if

$$k^2 = k^{*2} = 0\,, \qquad k \cdot k^* = 1\,. \tag{4.28}$$

These vectors determine a unique timelike bivector

$$K \equiv k \wedge k^*\,, \tag{4.29}$$

with

$$K^2 = (k \cdot k^*)^2 = 1\,. \tag{4.30}$$

Also,

$$Kk = K \cdot k = k \qquad \text{and} \qquad Kk^* = -k^*\,. \tag{4.31}$$

Therefore k and k^* are eigenvectors of the tensor

$$\underline{K}_-(a) \equiv K \cdot a \tag{4.32}$$

with real eigenvalues ± 1, even though $\underline{K}_-$ is a skew symmetric tensor. Relations (4.28) and (4.29) can be combined into

$$kk^* = 1 + K\,, \tag{4.33}$$

which has the idempotence property

$$(kk^*)^2 = (1 + K)^2 = 2kk^*\,. \tag{4.34}$$

All the above relations between k and k^* are invariant under the *rescaling*

$$k \to \alpha k\,, \qquad k^* \to \frac{1}{\alpha}\,k^*\,, \tag{4.35}$$

by an arbitrary nonzero scalar α.

By analogy with the structure of the projection operators $\underline{e}_k$ in (4.18), we can construct from k and k^* the elementary tensors

$$\underline{k}(a) \equiv kk \cdot a = \tfrac{1}{2}kak \qquad \text{and} \qquad \underline{k}^*(a) \equiv k^*k^* \cdot a\,. \tag{4.36}$$

These tensors are not projections because, instead of being idempotent, they are *nilpotent*, satisfying

$$\underline{k}^2 = 0 = \underline{k}^{*2}\,. \tag{4.37}$$

However, they are also symmetric tensors and are obviously not of type I.

Two other elementary tensors can be formed from k and k^*, namely

$$\underline{k}_*(a) \equiv kk^* \cdot a \tag{4.38a}$$

and its adjoint

$$\bar{k}_*(a) \equiv k^*k \cdot a\,. \tag{4.38b}$$

These tensors are idempotent, but they are not symmetric. However, they do have the symmetric combination

$$\underline{K}(a) \equiv KK \cdot a = \underline{k}_*(a) + \bar{k}_*(a)\,, \tag{4.39}$$

which is just the projection onto the K-plane.

We are now ready to continue the classification of symmetric tensors. To simplify comparison with type I we identify the K-plane with the e_{30}-plane in the notation of (4.19).

Type II

A symmetric tensor of type II has the canonical form

$$\underline{S} = \underline{S}_0 + \lambda_1 \underline{e}_1 + \lambda_2 \underline{e}_2 \,, \tag{4.40}$$

where the $\underline{e}_k$ satisfy (4.15a–d) and

$$\underline{S}_0 \underline{e}_k = \underline{e}_k \underline{S}_0 = 0 \,. \tag{4.41}$$

It has an *irreducible* eigenbivector K satisfying

$$\underline{S}(K) = \lambda_0^2 K \,. \tag{4.42}$$

In other words, $\underline{S}$ leaves a timelike plane invariant. The eigenbivector K is irreducible in the sense that, unlike the eigenbivectors e_{ij} of a type I tensor, it cannot be factored into a product of two eigenvectors. The dual of K is proportional to e_{12}, so it is a reducible eigenbivector satisfying

$$\underline{S}(iK) = \lambda_1 \lambda_2 iK \,. \tag{4.43}$$

The type II tensors fall into three distinct subtypes, which can be characterized by specifying the operator $\underline{S}_0$ in terms of the three distinct elementary symmetric operators defined by (4.36) and (4.39).

Type II$_\pm$

$$\underline{S}_0 = \underline{K}[\lambda_0 \pm \underline{k}] = \lambda_0 \underline{K} \pm \underline{k} \,. \tag{4.44}$$

In this case there is exactly one null eigenvector when $\lambda_0 \neq 0$,

$$\underline{S}k = \underline{S}_0 k = \lambda_0 k \,. \tag{4.45}$$

If both k and k^* were eigenvectors we would have a type I tensor. The coefficient of the last term in (4.44) has been set to unit magnitude without loss of generality, because $\underline{k}$ can be rescaled using (4.35) without affecting $\underline{K}$ in the other term. Because rescaling cannot change the sign of the coefficient, there are two distinct cases II$_\pm$.

Type II$_0$

$$\underline{S}_0 = \lambda_0 \underline{K} + \beta(\underline{k}^* - \underline{k}) \,. \tag{4.46}$$

In this case, with β nonzero, there are no eigenvectors in the K-plane. Rescaling has been used to set the coefficients of $\underline{k}$ and $\underline{k}^*$ to the same magnitude, but they are necessarily opposite in sign. The characteristic equation for $\underline{S}_0$ has complex roots $z = \lambda_0 \pm i\beta$, though the unit imaginary here is devoid of geometric meaning and cannot be identified with the unit pseudoscalar. Instead this "imaginary" originates from the combination of nilpotent operators on the right side of (4.46).

Type III

In this case there is a single irreducible eigentrivector K_1, which can be written in the form

$$K_1 = k \wedge k^* \wedge e_1 = K e_1 = -i e_2 \,. \tag{4.47}$$

Defining

$$\underline{K}_1(a) \equiv K_1 K_1^{-1} \cdot a \,, \tag{4.48}$$

$$\underline{k}_1(a) \equiv k e_1^{-1} \cdot a \qquad \text{and} \qquad \bar{\underline{k}}_1(a) \equiv e^{-1} k \cdot a \,, \tag{4.49}$$

a type III tensor can be put in the canonical form

$$\underline{S} = \lambda_1 \underline{K}_1 + \underline{k}_1 + \bar{\underline{k}}_1 + \lambda_2 \underline{e}_2 \,. \tag{4.50}$$

Here the symmetric tensor $\underline{k}_1 + \bar{\underline{k}}_1$ has been scaled to unity without loss of generality. It is readily verified that

$$\underline{S}(K_1) = \lambda_1^3 K_1 \,, \tag{4.51}$$

and the 3-dimensional hyperplane "spanned" by K_1 "contains" exactly one eigenvector

$$\underline{S}(k) = \lambda_1 k \tag{4.52}$$

and one eigenbivector

$$\underline{S}(k \wedge e_1) = \lambda_1^2 k \wedge e_1 \,. \tag{4.53}$$

In addition, there is one other eigenvector

$$\underline{S}(e_2) = \lambda_2 e_2 \tag{4.54}$$

and one other eigenbivector

$$\underline{S}(k \wedge e_2) = \lambda_1 \lambda_2 k \wedge e_2 \,. \tag{4.55}$$

Equations (4.53) and (4.55) characterize two invariant null planes intersecting in the null line of k.

This completes the classification of the main types of symmetric tensors on Minkowski space. An important application is to the classification of energy-momentum tensors. For any nonspacelike vector u, a given energy-momentum tensor $T(u)$ is usually assumed (at least tacitly) to satisfy the conditions

$$u \cdot T(u) \geq 0 \,, \tag{4.56}$$

$$[\, T(u) \,]^2 \geq 0 \,. \tag{4.57}$$

The first condition ensures that the energy density is never negative. This is formally equivalent to the condition (3.22) for a Lorentz transformation to be orthochronous, though equality never occurs in that case. Accordingly, we refer to any tensor $T(a)$ as *orthochronous* if it satisfies (4.56). The conditions (4.56) eliminate classes II_-, II_0 and III. Moreover, for classes I and II_+ they require that

$$\lambda_0 \geq |\,\lambda_k\,| \tag{4.58}$$

for $k = 1, 2, 3$ in the first case, or $k = 1, 2$ in the second.

As an important example, consider the energy-momentum tensor for an arbitrary electromagnetic field F, which can be put in the form

$$T(a) = -\tfrac{1}{2}FaF = -\tfrac{1}{2}faf\,, \tag{4.59}$$

where the last equality arises on using the canonical form (1.39). There are two cases. When $f = e_3e_0$ is a timelike bivector, (4.49) has two pairs of doubly degenerate eigenvalues

$$\lambda_0 = \lambda_3 = \tfrac{1}{2}f^2 = -\lambda_1 = -\lambda_2\,, \tag{4.60}$$

so that it is a type I tensor. When f is a null bivector with the form (1.40), (4.59) can be written

$$T(a) = \tfrac{1}{2}ekake = ekk\cdot ae = kk\cdot a\,, \tag{4.61}$$

so that it is a type II_+ tensor of the most elementary kind (4.36).

5. Transformations on spacetime

This section describes the apparatus of geometric calculus for handling transformations of spacetime and the induced transformations of multivector fields on spacetime. We concentrate on the mappings of 4-dimensional regions, including the whole of spacetime, but the apparatus applies with only minor adjustments to the mapping of any submanifold in spacetime. Throughout, we assume whatever differentiability is required to perform indicated operations, so that we might as well assume that the transformations are diffeomorphisms and defer the analysis of discontinuities in derivatives. We therefore assume that all transformations are invertible unless otherwise indicated.

Let f be a diffeomorphism which transforms each point x in some region of spacetime into another point x', as expressed by

$$f:\ x \quad\to\quad x' = f(x)\,. \tag{5.1}$$

This induces a linear transformation of tangent vectors at x to tangent vectors at x', given by the differential

$$\underline{f}:\quad a \quad\to\quad a' = \underline{f}(a) = a\cdot\nabla f\,. \tag{5.2}$$

More explicitly, it determines the transformation of a vector field $a = a(x)$ into a vector field

$$a' = a'(x') \equiv \underline{f}[a(x);\, x] = \underline{f}[a(f^{-1}(x'));\, f^{-1}(x')]\,. \tag{5.3}$$

The outermorphism of $\underline{f}$ determines an induced transformation of specified multivector fields. In particular,

$$\underline{f}(i) = J_f i, \qquad \text{where} \qquad J_f = \det\underline{f} = -i\underline{f}i \tag{5.4}$$

is the *Jacobian* of f.

The transformation f also induces an *adjoint* transformation $\bar{f}$ which takes tangent vectors at x' back to tangent vectors at x, as defined by

$$\bar{f}: \quad b' \quad \to \quad b = \bar{f}(b') \equiv \grave{\nabla}\grave{f} \cdot b' = \partial_x f(x) \cdot b' \,. \tag{5.5}$$

More explicitly, for vector fields

$$\bar{f}: \quad b'(x') \quad \to \quad b(x) = \bar{f}[\,b'(x');\, x\,] = \bar{f}[\,b'(f(x));\, x\,] \,. \tag{5.6}$$

The differential is related to the adjoint by

$$b' \cdot \underline{f}(a) = a \cdot \bar{f}(b') \,. \tag{5.7}$$

According to (3.15), $\bar{f}$ determines the inverse transformation

$$\underline{f}^{-1}(a') = \bar{f}(a'i)(J_f i)^{-1} = a \,. \tag{5.8}$$

Also, however,

$$\underline{f}^{-1}(a') = a' \cdot \partial_{x'} f^{-1}(x') \,. \tag{5.9}$$

Thus, the inverse of the differential equals the differential of the inverse.

Since the adjoint maps "backward" instead of "forward," it is often convenient to deal with its inverse

$$\overline{f^{-1}}: \quad a(x) \quad \to \quad a'(x') = \overline{f^{-1}}\,[\,a(f^{-1}(x'))\,] \,. \tag{5.10}$$

This has the advantage of being directly comparable to $\underline{f}$. Note that it is not necessary to distinguish between $\overline{f^{-1}}$ and $\bar{f}^{-1}$.

Thus, we have two kinds of induced transformations for multivector fields: The first, by $\underline{f}$, is commonly said to be *contravariant*, while the second, by $\bar{f}$ or $\bar{f}^{-1}$, is said to be *covariant*. The first is said to "transform like a vector," while the second is said to "transform like a *covector*." The term "vector" is thus associated with the differential while "covector" is associated with the adjoint. This linking of the vector concept to a transformation law is axiomatic in ordinary tensor calculus. In geometric calculus, however, the two concepts are kept separate. The algebraic concept of vector is determined by the axioms of geometric algebra without reference to any coordinates or transformations. Association of a vector or a vector field with a particular transformation law is a separate issue.

The transformation of a multivector field can also be defined by the rule of *direct substitution:* A field $F = F(x)$ is transformed to

$$F'(x') \equiv F'(f(x)) = F(x) \,. \tag{5.11}$$

Thus, the values of the field are unchanged — although they are associated with different points by changing the functional form of the field. For the purposes of "gauge gravity theory" discussed in the next section, it is very important to note that the alternative definition $F'(x) \equiv F(x')$ is adopted in [1]. Each of these two alternatives has something to recommend it.

Directional derivatives of the two different functions in (5.11) are related by the *chain rule:*

$$a\cdot\nabla F = a\cdot\partial_x F'(f(x)) = (a\cdot\nabla_x f(x))\cdot\nabla_{x'}F'(x') = \underline{f}(a)\cdot\nabla' F' = a'\cdot\nabla' F' \,. \tag{5.12}$$

The chain rule is more simply expressed as an operator identity

$$a\cdot\nabla = a\cdot\bar{f}(\nabla') = \underline{f}(a)\cdot\nabla' = a'\cdot\nabla' \,. \tag{5.13}$$

Differentiation with respect to the vector a yields the general transformation law for the vector derivative:

$$\nabla = \bar{f}(\nabla') \qquad \text{or} \qquad \nabla' = \bar{f}^{\,-1}(\nabla)\,. \tag{5.14}$$

This is the most basic formulation of the chain rule, from which its many implications are most easily derived. All properties of induced transformations are essentially implications of this rule, including the transformation law for the differential, as (5.13) shows.

The rule for the induced transformation of the curl is derived by using the integrability condition (2.15) to prove that the adjoint function has vanishing curl; thus, for the adjoint of a vector field,

$$\grave{\nabla}\wedge\grave{\bar{f}}(a') = \nabla_b\wedge\bar{f}_b(a') = \nabla_b\wedge\nabla_c f_{cb}\cdot a' = \nabla\wedge\nabla f\cdot a' = 0\,. \tag{5.15}$$

The transformation rule for the curl of a vector field $a = \bar{f}(a')$ is therefore

$$\nabla\wedge a = \nabla\wedge\bar{f}(a') = \bar{f}(\nabla'\wedge a')\,. \tag{5.16}$$

To extend this to multivector fields, note that the differential of an outermorphism is not itself an outermorphism; rather it satisfies the "product rule"

$$\bar{f}_b(A'\wedge B') = \bar{f}_b(A')\wedge\bar{f}(B') + \bar{f}(A')\wedge\bar{f}_b(B')\,. \tag{5.17}$$

Therefore, it follows from (5.15) that the curl of the adjoint outermorphism vanishes, and (5.16) generalizes to

$$\nabla\wedge A = \bar{f}(\nabla'\wedge A') \qquad \text{or} \qquad \nabla'\wedge A' = \bar{f}^{\,-1}(\nabla\wedge A)\,, \tag{5.18}$$

where $A = \bar{f}(A')$. Thus, the outermorphism of the curl is the curl of an outermorphism.

The transformation rule for the divergence is more complex, but it can be derived from that of the curl by exploiting the duality of inner and outer products (1.12) and the transformation law (3.14) relating them. Thus,

$$\bar{f}(\nabla'\wedge(A'i)) = \bar{f}[(\nabla'\cdot A')i] = \bar{f}^{\,-1}(\nabla'\cdot A')\bar{f}(i)\,.$$

Then, using (5.18) and (5.4) we obtain

$$\nabla\wedge\bar{f}(A'i) = \nabla\wedge[\,\underline{f}^{-1}(A')\bar{f}(i)] = \nabla\cdot(J_f A)i\,.$$

For the divergence, therefore, we have the transformation rule

$$\nabla'\cdot A' = \nabla'\cdot\underline{f}(A) = J_f{}^{-1}\underline{f}[\nabla\cdot(J_f A)] = \underline{f}[\nabla\cdot A + (\nabla\ln J_f)\cdot A]\,, \tag{5.19}$$

where $A' = \underline{f}(A)$. This formula can be separated into two parts:

$$\dot{\nabla}' \cdot \dot{\underline{f}}(A) = \underline{f}[(\nabla \ln J_f) \cdot A] = (\nabla' \ln J_f) \cdot \underline{f}(A), \tag{5.20}$$

$$\dot{\nabla}' \cdot \underline{f}(\dot{A}) = \underline{f}(\nabla \cdot A). \tag{5.21}$$

The whole may be recovered from the parts by using the following generalization of (5.13) (which can also be derived from (3.14)):

$$\underline{f}(A) \cdot \nabla' = \underline{f}(A \cdot \nabla). \tag{5.22}$$

6. Directed Integrals and the Fundamental Theorem

In the theory of integration, geometric calculus absorbs, clarifies and generalizes the calculus of differential forms. Only the essentials are sketched here; details are given in [3], and [7] discusses the basic concepts at greater length with applications to physics.

The integrand of any integral over a k-dimensional manifold is a *differential k-form*

$$L = L(d^k x) = L[\, d^k x;\, x\,], \tag{6.1}$$

where $d^k x$ is a k-vector-valued measure on the manifold. If the surface is not null at x, we can write

$$d^k x = I_k \,|\, d^k x\,|, \tag{6.2}$$

where $I_k = I_k(x)$ is a unit k-vector field tangent to the manifold at x, and $|\, d^k x\,|$ is an ordinary scalar-valued measure. Thus, $d^k x$ describes the direction of the tangent space to the manifold at each point. For this reason it is called a *directed measure*. Since the integrals are defined from weighted sums, the integrand $L(d^k x)$ must be a linear function of its k-vector argument; accordingly it is a *k-form* as defined in Section 4. Of course, the values of L may vary with x, as indicated by the explicit x-dependence shown on the right side of (2.17).

The *exterior differential* of a k-form L is a $(k+1)$-form dL defined by

$$dL = \dot{L}[\,(d^{k+1}x) \cdot \dot{\nabla}\,] = L[\,(d^{k+1}x) \cdot \dot{\nabla};\, \dot{x}\,], \tag{6.3}$$

where the accent indicates that only the implicit dependence of L on x is differentiated. The exterior derivative of any "k-form" which is already the exterior derivative of another form necessarily vanishes, as is expressed by

$$d^2 L = 0. \tag{6.4}$$

This is an easy consequence of the integrability condition (2.15); thus,

$$d^2 L = d\dot{L}[(d^{k+1}x) \cdot \dot{\nabla}\,] = \dot{L}[(d^{k+1}x) \cdot (\dot{\nabla} \wedge \dot{\nabla})] = 0.$$

The *Fundamental Theorem of Integral Calculus* (also known as the "Boundary Theorem" or the "Generalized Stokes' Theorem") can now be written in the compact form

$$\int dL = \int dL\,(d^{k+1}x) = \oint L(d^k x) = \oint L\,. \tag{6.5}$$

This says that the integral of any k-form L over a *closed* k-dimensional manifold is equal to the integral of its exterior derivative over the enclosed $(k+1)$-dimensional manifold. It follows from (6.4) that this integral vanishes if $L = dN$ where N is a $(k-1)$-form.

To emphasize their dependence on a *directed measure*, the integrals in (6.5) may be called *directed integrals*. In conventional approaches to differential forms this dependence is disguised and all forms are scalar-valued. For that special case we can write

$$L = \langle\, A\, d^k x\,\rangle = (d^k x)\cdot A(x)\,, \tag{6.6a}$$

where $A = A(x)$ is a k-vector field. Then

$$dL = [\,(d^{k+1}x)\cdot\nabla\,]\cdot A = (d^{k+1}x)\cdot(\nabla\wedge A)\,. \tag{6.6b}$$

In this case, therefore, the exterior derivative is equivalent to the curl.

An alternative form of the Fundamental Theorem called "Gauss's Theorem" is commonly used in physics. If L is a 3-form, its 3-vector argument can be written as the dual of a vector, and a *tensor field* $T(n) = T[n(x);\, x]$ can be defined by

$$T(n) = L(in)\,. \tag{6.7}$$

According to (6.2) we can write

$$d^4x = i\,|\,d^4x\,| \qquad \text{and} \qquad d^3x = in^{-1}\,|\,d^3x\,|\,, \tag{6.8}$$

where n is the *outward unit normal* defined by the relation $I_3 n = I_4 = i$. Substitution into (6.5) then gives *Gauss's Theorem:*

$$\int \grave{T}(\grave{\nabla})\,|\,d^4x\,| = \oint T(n^{-1})\,|\,d^3x\,|\,. \tag{6.9}$$

where $n^{-1} = \epsilon n$ with signature ϵ. Though $\grave{T}(\grave{\nabla})$ may be called the "divergence of the tensor T," it is not generally equivalent to the divergence as defined earlier for multivector fields. However, if L is scalar-valued as in (6.6a), then (6.7) implies that

$$T(n) = n\cdot a\,, \tag{6.10a}$$

where $a = a(x) = A(x)i$ is a vector field. In this case, we do have the divergence

$$\grave{T}(\grave{\nabla}) = \nabla\cdot a\,. \tag{6.10b}$$

Note that duality has changed the curl in (6.6b) into the divergence in (6.10b).

A *change of integration variables* in a directed integral is a transformation on a differential form by direct substitution. Thus, for the k-form defined in (6.1) we have

$$L'(d^k x') = L(d^k x)\,, \tag{6.11}$$

where

$$d^k x' = \underline{f}(d^k x) \qquad \text{or} \qquad d^k x = \underline{f}^{-1}(d^k x') \tag{6.12}$$

In other words, $L' = L\underline{f}^{-1}$ or, more explicitly,

$$L'(d^k x';\, x') = L[\,\underline{f}^{-1}(d^k x);\, f^{-1}(x)\,] = L(d^k x; x)\,. \tag{6.13}$$

The value of the integral of (6.11) is thus unaffected by the change of variables,

$$\int L'(d^k x') = \int L[\,\underline{f}^{-1}(d^k x')\,] = \int L(d^k x)\,. \tag{6.14}$$

The exterior derivative and hence the fundamental theorem are likewise unaffected. In other words,

$$dL' = dL\,. \tag{6.15}$$

This follows from

$$(d^k x')\cdot\nabla' = \underline{f}(d^k x)\cdot\bar{f}(\nabla) = \underline{f}[(d^k x)\cdot\nabla] \tag{6.16}$$

and

$$d\underline{f}(d^k x) = \grave{\underline{f}}\,[(d^k x)\cdot\grave{\nabla}] = 0\,. \tag{6.17}$$

Like (6.4), the last equation is a consequence of the integrability condition.

Part II. INDUCED GEOMETRY ON FLAT SPACETIME

7. Gauge Tensor and Gauge Invariance

We shall regard (Minkowski) spacetime as a mathematical device for representing the ordering of physical events. A *spacetime map* represents the ordering of particular events by points (vectors) in spacetime. No other physical property is attributed to spacetime itself. Instead, all other properties of physical entities are represented as fields on spacetime or as constructs from such fields.

A given ordering of events can be represented by a map in many different ways, just as the surface of the earth can be represented by Mercator projection, stereographic projection or many other equivalent maps. As the physical world is independent of the way we construct our maps, we seek a physical theory which is equally independent. This idea can be formulated as a general theoretical principle, which, with deference to Einstein, we dub

The Principle of General Invariance (PGI): *The equations of physics must be invariant under arbitrary smooth remappings of events onto spacetime.*

The precise mathematical implementation of this principle leads naturally to a geometric theory of gravitation, as we shall show.

A smooth remapping of spacetime is a diffeomorphism of spacetime onto itself. The mathematical apparatus for handling such transformations was set out in Section 5. We saw there that, via the chain rule, differentiation automatically induces transformations of vectors and covectors. For example, let $x = x(\tau)$ be a timelike curve representing a *particle history*. According to (5.2), the diffeomorphism (5.1) induces the transformation

$$\dot{x} = \frac{dx}{d\tau} \qquad \to \qquad \dot{x}' = \underline{f}(\dot{x}) \,. \tag{7.1}$$

Thus the description of particle velocity by $\dot{x}$ is "covariant" under spacetime diffeomorphisms. An "invariant" description by a *velocity* vector $v = v(x(\tau))$ can be achieved by introducing an invertible tensor field $\underline{h}$ such that

$$\dot{x} = \underline{h}(v), \tag{7.2}$$

while supposing that $\underline{h}$ undergoes the induced transformation

$$\underline{f}: \quad \underline{h} \qquad \to \qquad \underline{h}' = \underline{f}\,\underline{h} \,, \tag{7.3}$$

where it is understood that $\underline{h}$ is a function of x while, using (5.1), $\underline{h}'$ is expressed as a function of x'. Then (7.1) implies that

$$\dot{x}' = \underline{h}'(v) \tag{7.4}$$

where $v = v(\underline{f}^{-1}(x')) \equiv v(x')$ has the same value as in (7.2), but it is taken as a function of $x'(\tau)$ instead of $x(\tau)$. To distinguish the velocity representation $\dot{x}$ from

$$v = \underline{h}^{-1}(\dot{x}) \,, \tag{7.5}$$

let us refer to $\dot{x}$ as the *map velocity*. Otherwise the term *velocity* designates v. The normalization

$$v^2 = 1 \tag{7.6}$$

fixes the scale on the parameter τ, which can therefore be interpreted as *proper time*.

The transformation of $\underline{h}$ defined by (7.3) is called a *"position gauge transformation,"* and $\underline{h}$ (or $\bar{h}$) is called the *"gauge tensor"* or simply the *"gauge"* on spacetime. Actually, the term "gauge" is more appropriate here than elsewhere in physics, because $\underline{h}$ does indeed determine the "gauging" of a metric on spacetime. To see that, use (7.1) in (7.5) to derive the following expression for the *invariant line element* on a timelike particle history:

$$d\tau^2 = [\,\underline{h}^{-1}(dx)\,]^2 = dx \cdot g(dx) \,, \tag{7.7a}$$

where

$$g = \bar{h}^{-1}\,\underline{h}^{-1} \tag{7.7b}$$

is a symmetric *metric tensor*. Our formulation indicates the *gauge* as a more fundamental geometric entity than the metric, and this view will be confirmed by developments below. Einstein has taught us to interpret the metric tensor physically as a gravitational potential, making it easy and natural to transfer this interpretation to the gauge tensor. Some readers will recognize $\underline{h}$ as equivalent to a "vierbein field," which has been proposed before to represent gravitational fields on flat spacetime [17]. However, availability of the spacetime calculus makes all the difference in turning this idea into a practical reality.

To verify that (7.7a) is equivalent to the standard invariant line element in general relativity, coordinates can be introduced. Let $x = x(x^0, x^1, x^2, x^3)$ be a parametrization of the points, in some spacetime region, by an arbitrary set of coordinates $\{x^\mu\}$. Partial derivatives then give tangent vectors to the coordinate curves $\partial_\mu x$, which, in direct analogy to (7.2) and (7.5), determine a set of position gauge invariant vector fields $\{g_\mu\}$ according to the equation

$$\partial_\mu x = \frac{\partial x}{\partial x^\mu} = \underline{h}(g_\mu)\,, \tag{7.8}$$

or, equivalently, by

$$g_\mu = \underline{h}^{-1}(\partial_\mu x)\,. \tag{7.9}$$

The components for this coordinate system are then given by

$$g_{\mu\nu} = g_\mu \cdot g_\nu = (\partial_\mu x) \cdot g(\partial_\nu x)\,. \tag{7.10}$$

Therefore, with $dx = dx^\mu \partial_\mu x$, the line element (7.7a) can be put in the familiar form

$$d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu\,. \tag{7.11}$$

More about coordinates in Part III.

The adjoint of (7.3) gives us the transformation rule

$$\bar{h}' = \bar{h}\bar{f}\,. \tag{7.12}$$

Applying this to the vector derivative ∇ with its transformation rule (5.14), we can define a *position gauge invariant derivative* by

$$\bar{\nabla} \equiv \bar{h}(\nabla) = \bar{h}'(\nabla')\,. \tag{7.13}$$

Applied to a gauge invariant field $\varphi = \varphi(x)$ this gives the *position gauge invariant gradient*

$$\bar{\nabla}\varphi \equiv \bar{h}(\nabla\varphi) = \bar{h}'(\nabla'\varphi')\,. \tag{7.14}$$

where $\varphi' = \varphi(x') = \varphi(f^{-1}(x'))$. From the operator $\bar{\nabla}$ we obtain a *position gauge invariant directional derivative*

$$a \cdot \bar{\nabla} = a \cdot \bar{h}(\nabla) = (\underline{h}a) \cdot \nabla\,, \tag{7.15}$$

where a is a "free vector," which is to say that it can be regarded as constant or as transforming by substitution:

$$a' = a'(x') = a(f^{-1}(x')) = a(x) = a\,. \tag{7.16}$$

The transformation rules (7.3) and (7.12) now give us

$$\underline{h}'(a') \cdot \nabla' = [\underline{f}\,\underline{h}(a)] \cdot \nabla' = \underline{h}(a) \cdot \bar{f}(\nabla') = \underline{h}(a) \cdot \nabla,$$

which makes the invariance of (7.15) explicit.

For any field $F = F(x(\tau))$ defined on a particle history $x(\tau)$, the chain rule gives the operator relation

$$\frac{d}{d\tau} = \dot{x} \cdot \nabla = \underline{h}(v) \cdot \nabla = v \cdot \bar{\nabla}\,, \tag{7.17}$$

so that

$$\frac{dF}{d\tau} = v \cdot \bar{\nabla} F\,. \tag{7.18}$$

For $F(x) = x$, this gives

$$\frac{dx}{d\tau} = v \cdot \bar{\nabla} x = (\underline{h}v) \cdot \nabla x = \underline{h}v\,, \tag{7.19}$$

recovering (7.2).

Besides the gauge transformation (7.3) or (7.12), there is another kind of gauge transformation that leaves invariant the form of the above equations. From the discussion in Section 3, we know that the velocity normalization (7.6) is invariant under the Lorentz rotation

$$\overline{R}\,: \quad v = v(x) \qquad \rightarrow \qquad v' = \overline{R}\,v = \overline{R}\,(v(x);\, x)\,. \tag{7.20}$$

Here, however, v is a vector field and the rotation $\overline{R}$ varies smoothly from point to point but is otherwise arbitrary. Let us signify this local variability of $\overline{R}$ by calling it a *local Lorentz rotation*. Since $\overline{R} = \underline{R}^{-1}$, the relation

$$\dot{x} = \underline{h}(v) = \underline{h}'(v')$$

is left invariant by (7.20) if it is accompanied by the adjoint gauge change

$$\underline{h}' = \underline{h}\,\underline{R}\,. \tag{7.21}$$

This kind of gauge transformation leaves the metric tensor (7.7b) invariant, since

$$(\overline{R}\,\bar{h})^{-1}(\underline{h}\,\underline{R})^{-1} = \bar{h}^{-1}\underline{R}\overline{R}\,\underline{h}^{-1} = \bar{h}^{-1}\underline{h}^{-1} = g\,.$$

The gauge change (7.21) does not entail any remapping of spacetime points. Rather, via (7.20), it induces a change of direction at each spacetime point. We express this fact by referring to (7.21) as a *directional gauge change*. This name distinguishes it from the *positional gauge change* (7.3). Of course, the two kinds are combined in the most general *spacetime gauge change:*

$$\underline{h}' = \underline{f}\,\underline{h}\,\underline{R}\,. \tag{7.22}$$

The idea expressed by (7.2), that the tangent vector $\dot{x}$ to a particle history is not generally collinear with its velocity v, is unfamiliar to most physicists. It may be of some comfort, therefore, to note that if $\dot{x}$ is timelike, as is usual for a material particle except in an extreme gravitational field, then it can be *aligned* with v by a local Lorentz rotation along a segment of the particle history. In that case the gauge is set so that v is an eigenvector of the gauge tensor, and (7.3) reduces to

$$\dot{x} = \underline{h}(v) = \gamma v\,. \tag{7.23}$$

The eigenvalue γ evidently characterizes a gravitationally induced shift in the particle's proper time rate, as does the equation (7.7a) for the invariant line element.

In general, the noncollinearity of v and $\dot{x}$ in (7.2) signifies the presence of a gravitational field, so it cannot be transformed away globally. A major problem for gravitation theory is therefore to ascertain precisely the global constraints on gauge freedom induced by various physical assumptions. Just as positional gauge freedom has been theoretically formalized by the Principle of General Invariance (PGI), directional gauge freedom can be formalized by

> **The Principle of Local Relativity (PLR):** *The equations of physics must be covariant (i.e. form invariant) under local Lorentz rotations.*

This principle, first formulated with spacetime calculus in [4], can be interpreted physically as asserting that the relative directions of multivector fields representing physical properties at each spacetime point have an absolute physical significance, while the comparison of directions at different points generally does not. Equation (3.19) tells us that the geometric product is preserved by Lorentz rotations at a point, so that *relative directions at each spacetime point are invariant*. Obviously, the position dependence of local Lorentz rotations implies that this is not so for relative directions at different points — although solutions to the equations of physics may determine some physical significance for this so-called "distant parallelism," as in the case of flat spacetime.

An important consequence of local relativity is that, although the operator $\not\nabla$ defined by (7.13) is positionally gauge invariant, it is not directionally invariant. The gauge change (7.21) induces a transformation of $\not\nabla$ to

$$\not\nabla' = \bar{R}\not\nabla\,. \tag{7.24}$$

Except when operating on scalars, as in (7.14), $\not\nabla$ is not directionally covariant. To satisfy the Principle of Local Relativity a directionally covariant differential operator must be defined. That will be the main task in the next section.

Our two general gauge principles PGI and PLR are obviously closely related to Einstein's principles of general and special relativity. To make the relation definite, let us regard the PGI as a gauge formulation of Einstein's *Principle of General Covariance* (PGC). Recall that Einstein regarded the PGC as a cornerstone of his general theory, but that shortly after publication Kretschmann and Cartan argued that it is devoid of physical content. Einstein retreated but continued to assert that the PGC plays an essential heuristic role in his theory. Precisely what that role might be has remained obscure to this day. Mutually incompatible attempts to pin it down by various commentators range as far as outright rejection of Einstein's claim. All obscurity is removed, however by the restatement of the PGC as the PGI, with its precise formulation in terms of the spacetime calculus. We have seen that implementation of the PGI requires the existence of a gauge tensor and of the geometry it entails. This is a nontrivial heuristic consequence of great importance, for it is the main idea driving the construction of the whole gauge theory of gravitation in [1]. Thus it provides strong grounds for asserting that: Once again, Einstein's physical intuition is unerring — outstripping the analysis of his critics.

Another obscure point in Einstein's work which has been endlessly discussed, is the relation between his Special and General Theories. Gauge theory throws new light on this issue with its sharp distinctions between positional and directional gauge transformations. Since the latter are Lorentz rotations $\{\underline{R}\}$, the PLR can be regarded as a generalization of the *Principle of Special Relativity*, to which it reduces when each $\underline{R}$ is a constant. The justification for this reduction is that, in the absence of gravity, there is a preferred gauge $\underline{h} = 1$ and all points are geometrically equivalent. From (7.21), therefore, all other gauges are given by $\underline{h}' = \underline{R}$. This fact can be interpreted as an expression of the *isotropy* of the spacetime gauge. Homogeneity of the spacetime gauge is expressed by the existence of preferred position gauge changes, called *Poincaré transformations*, with the general form

$$x' = f(x) = \underline{R}x + c\,, \tag{7.25}$$

where $\underline{R}$ and c are constant. The differential of this transformation is

$$\underline{f} = \underline{R}\,. \tag{7.26}$$

Thus, positional and directional gauge changes merge in Special Relativity. Conversely, the Lorentz transformations of Special Relativity are generalized in two distinctly different ways by General Relativity. This fact is obscured in the usual covariant formulation in terms of tensor analysis.

8. Covariant Derivative and Parallel Transport

To satisfy the Principle of Local Relativity we construct a new differential operator which is covariant (i.e. form invariant) under directional gauge transformations. Generalizing (7.20) and using (3.22), the directional gauge change of an arbitrary Multivector field $A = A(x)$ can be written

$$\overline{R}\,(A) = RA\widetilde{R} \tag{8.1}$$

where $R = R(x)$ is a rotor field. The directional derivative of the rotor can be written in the form

$$a\cdot\nabla R = \tfrac{1}{2}R\Omega(a)\,, \tag{8.2}$$

where the quantity

$$\Omega(a) = 2\widetilde{R}\,a\cdot\nabla R = -2(a\cdot\nabla\widetilde{R})R = -\widetilde{\Omega}(a) \tag{8.3}$$

is a bivector field. Therefore, the directional derivative of $\overline{R}\,(A)$ can be written

$$a\cdot\grave{\nabla}\grave{\overline{R}}\,(A) = \overline{R}\,[\Omega(a)\times A] = [\,\overline{R}\,\Omega(a)\,]\times\overline{R}\,(A) \tag{8.4}$$

where $\times$ is the commutator product defined by (1.32). According to (1.36), the commutator product with $\Omega(a)$ is grade-preserving.

A *gauge covariant directional derivative* can now be defined in a standard way by introducing a so-called *gauge field* $\omega(a) = \omega(a;x)$ with the operator definition

$$a \cdot D \equiv a \cdot \check{\nabla} + \omega(a) \times \; . \tag{8.5}$$

Applied to any multivector field A this gives, of course,

$$a \cdot DA = a \cdot \check{\nabla} A + \omega(a) \times A \, . \tag{8.6}$$

The idea is to assign to $\omega(a)$ the gauge transformation rule

$$\omega'(a') = \overline{R}\,(\omega(a) - \Omega(a)) = R\omega(a)\widetilde{R} - 2(a \cdot \check{\nabla} R)\widetilde{R} \tag{8.7}$$

with $a' = \underline{R}a$, so that its last term cancels (8.4) and $a \cdot D$ satisfies the *gauge covariance condition*

$$a' \cdot D' \overline{R}\,(A) = \overline{R}\,(a \cdot DA) \tag{8.8}$$

with

$$a' \cdot D' = a' \cdot \check{\nabla}' + \omega'(a') \times = a \cdot \check{\nabla} + \omega'(\underline{R}a) \times \; . \tag{8.9}$$

Clearly the gauge field $\omega(a)$ must be bivector-valued to satisfy (8.7).

To distinguish $a \cdot D$ from the *directional derivative* defined by (7.15) we call it the *directional coderivative*. A vectorial *coderivative* D can then be defined by

$$D = \partial_a a \cdot D = \check{\nabla} + \partial_a \omega(a) \times \, , \tag{8.10}$$

where $\check{\nabla}$ is given by (7.13). In geometry the gauge field $\omega(a)$ is called a *linear connection* or *connexion*. The term "coderivative" can be regarded as short for *covariant derivative*.

Besides gauge invariance (8.8), the directional coderivative has a number of easily derived properties which are listed here for convenience:

(1) *Linearity:*

$$(a + b) \cdot D = a \cdot D + b \cdot D \, . \tag{8.11}$$

$$a \cdot D(A + B) = a \cdot DA + a \cdot DB \, . \tag{8.12}$$

(2) *Leibniz:* For geometric products,

$$a \cdot D(AB) = (a \cdot DA)B + A(a \cdot DB) \tag{8.13a}$$

and for tensor fields,

$$a \cdot D \underline{f}\, \underline{h} = a \cdot \grave{D} \grave{\underline{f}}\, \underline{h} + \underline{f} a \cdot D \underline{h} \, . \tag{8.13b}$$

(3) *Grade-preserving:*

$$a \cdot D \langle A \rangle_k = \langle a \cdot DA \rangle_k \, . \tag{8.14}$$

(4) *Gradient:* For the scalar field $\varphi = \varphi(x)$,

$$D\varphi = \dot{\nabla}\varphi = \bar{h}(\nabla\varphi)\,, \tag{8.15}$$

$$a\cdot D\varphi = a\cdot\dot{\nabla}\varphi = \underline{h}(a)\cdot\nabla\varphi \tag{8.16}$$

(5) *Codifferential.* The codifferential $\delta_a T$ of a tensor $T(a_1, a_2, \ldots, a_k)$ is defined by

$$\begin{aligned}\delta_a T(a_1,\ldots,a_k) &= a\cdot\grave{D}\grave{T}(a_1,\ldots,a_k)\\ &\equiv a\cdot DT(a_1,\ldots,a_k) - T(a\cdot Da_1, a_2,\ldots,a_k) - T(a_1, a\cdot Da_2,\ldots,a_k) - \ldots\end{aligned} \tag{8.17}$$

(6) *Curvature from commutator:*

$$[\delta_a,\delta_b]A = [a\cdot D, b\cdot D]A - [a,b]\cdot DA = R(a\wedge b)\times A\,, \tag{8.18}$$

where the *commutator* of codifferentials is defined by

$$[\,\delta_a,\delta_b\,] \equiv \delta_a\delta_b - \delta_b\delta_a\,, \tag{8.19a}$$

the *Lie bracket* is defined by

$$[\,a,b\,] \equiv a\cdot Db - b\cdot Da\,, \tag{8.19b}$$

and the *curvature tensor* $R(a\wedge b)$ is a bivector-valued field of a bivector variable $a\wedge b$ given by

$$\begin{aligned}R(a\wedge b) &\equiv \delta_a\omega(b) - \delta_b\omega(a) + \omega(a)\times\omega(b)\\ &= a\cdot\grave{\nabla}\grave{\omega}(b) - b\cdot\grave{\nabla}\grave{\omega}(a) + \omega(a)\times\omega(b) - \omega([a,b])\,.\end{aligned} \tag{8.20}$$

(7) *Spinor derivatives.* For a spinor field $\psi = \psi(x)$, the gauge rule (8.1) is replaced by

$$\psi \quad\to\quad R\psi \tag{8.21}$$

so, to achieve the condition

$$a\cdot D(R\psi) = R\,a\cdot D\psi \tag{8.22}$$

for gauge invariance, the definition (8.5) must be revised to

$$a\cdot D\psi = (a\cdot\dot{\nabla} + \tfrac{1}{2}\omega(a))\psi\,, \tag{8.23}$$

while (8.7) is retained.

If C is a constant multivector, then

$$A = \psi C\widetilde{\psi} \tag{8.24}$$

is a multivector field satisfying (8.1), and differentiation of A by the two different rules (8.6) and (8.23) give the same result. This justifies using the same symbol $a\cdot D$ in both cases. The difference can then be attributed to the different types of quantity being differentiated, the two types being distinguished by two different rules, (8.1) and (8.21), for gauge change. An alternative approach to spinor derivatives is adopted in [1] where (8.23) is regarded as defining a new differential operator, an operator which acts only on spinors — to be distinguished from the operator defined for multivector fields by (8.6). Spinor fields will not be considered further in this paper.

The equation for a particle history $x = x(\tau)$ generated by a timelike velocity $v = v(x(\tau))$ in the presence of an "ambient" gauge tensor $\underline{h}$ is

$$\dot{x} = \underline{h}(v)\,. \tag{8.25}$$

We have seen that, applied to any field $A = A(x(\tau))$, the derivative

$$\frac{d}{d\tau} = v\cdot\nabla\!\!\!\!/ = \dot{x}\cdot\nabla \tag{8.26}$$

is positionally gauge invariant. Now we can define a directionally gauge invariant *coderivative* by

$$\frac{\delta}{\delta\tau} = v\cdot D = \frac{d}{d\tau} + \omega(v)\times\ . \tag{8.27}$$

Obviously, all this applies to arbitrary differentiable curves in spacetime if the requirement that v be timelike is dropped.

The equation for a *geodesic* can now be written

$$\frac{\delta v}{\delta\tau} = v\cdot Dv = \dot{v} + \omega(v)\cdot v = 0 \tag{8.28}$$

With $\omega(v)$ specified by the "ambient geometry," this equation can be integrated for v and then (8.24) solved for the curve $x(\tau)$. The solution is facilitated by noting that (8.28) implies

$$v\cdot\dot{v} = \tfrac{1}{2}\,\frac{dv^2}{d\tau} = v\cdot\omega(v)\cdot v = \omega(v)\cdot(v\wedge v) = 0\,.$$

Therefore v^2 is constant on the curve and the value of v at any point on the curve can be obtained from a reference value v_0 by a Lorentz rotation. Thus,

$$v = \overline{R}\,v_0 = Rv_0\widetilde{R}\,, \tag{8.29}$$

where $R = R(x(\tau))$ is a rotor field on the curve satisfying the differential equation

$$\frac{\delta R}{\delta\tau} = v\cdot DR = \dot{R} + \tfrac{1}{2}\omega(v)R = 0\,. \tag{8.30}$$

Note that the rotor R submits to the spinor derivative (8.23); it can in fact be regarded as a special kind of spinor. A solution of equation (8.30) gives more than the velocity; it also determines the gravitational precession of a small rigid body, gyroscope or electron spin moving on the curve [18].

Indeed, the *parallel transport* of any fixed multivector A_0 to a field $A = A(x(\tau))$ defined on the whole curve is given by

$$A = \overline{R}\,A_0 = RA_0\widetilde{R}\,. \tag{8.31}$$

Of course, it satisfies the differential equation

$$\frac{\delta A}{\delta\tau} = \dot{A} + \omega(v) \times A = 0\,. \tag{8.32}$$

The similarity of this formulation for parallel transport to a directional gauge transformation is obvious.

With the spinor coderivative defined by (8.30) (and not necessarily vanishing), we can rewrite (8.27) as

$$\frac{\delta}{\delta\tau} = \underline{R}\,\frac{d}{d\tau}\,\overline{R}\,, \tag{8.33}$$

or equivalently, as

$$\overline{R}\,\frac{\delta}{\delta\tau} = \overline{R}\,v\cdot D = v\cdot \bar{\nabla}\overline{R} = \frac{d}{d\tau}\,\overline{R}\ . \tag{8.34}$$

This shows that a coderivative can be locally transformed to an ordinary derivative by a suitable gauge transformation.

The entire gauge covariant geometry on spacetime, including the curvature tensor (8.20), derives from the properties of the coderivative and its connexion $\omega(a)$. This Section has emphasized the directional coderivative ahead of the vector coderivative and considered only those properties which do not depend explicitly on the role of the gauge tensor. The next three sections derive those properties that do depend on the gauge and show that the vector coderivative demands a more prominent role. All of the results follow from (8.18), which embraces in one equation both the so-called “first and second fundamental forms” of Riemannian geometry.

9. From Gauge to Connexion

A Riemannian geometry on spacetime is completely determined by specifying a definite gauge tensor $\underline{h}$ or, equivalently, its adjoint $\bar{h}$. Although the connexion $\omega(a)$ was introduced in the last Section without reference to the gauge, it is in fact related to the gauge by “compatibility conditions” on the commutator of derivatives. In this section we determine the dependence of $\omega(a)$ on $\bar{h}$ and establish a number of results which are helpful in calculations. Definitions and results from previous sections will be used freely without further comment.

For a scalar field $\varphi = \varphi(x)$ the fundamental equation (8.18) reduces to

$$(a\cdot D\,b\cdot D - b\cdot D\,a\cdot D)\varphi = 0 \tag{9.1}$$

when a and b satisfy $[a,b] = 0$. The operator D, however, must be differentiated, because it is a function of $\bar{h}$, as is now made explicit. Differentiating by $\partial_a\partial_b$ we obtain

$$D\wedge D\varphi = D\wedge\bar{\nabla}\varphi = D\wedge\bar{h}(\nabla\varphi) = 0\,. \tag{9.2}$$

This is one of the fundamental equations determining the spacetime geometry. Its equivalent in the language of differential forms was dubbed the “first fundamental form” by Cartan. It could have been taken as a defining property for the vector coderivative in place of the scalar equation (9.1) for the directional coderivative. All results in this section and the next derive from it.

For $\varphi = a \cdot x$ and constant a, (9.2) becomes

$$D \wedge \bar{h}(a) = 0 . \tag{9.3}$$

By (8.10) this can be expanded to

$$\grave{\nabla} \wedge \grave{\bar{h}}(a) + \partial_b \wedge [\omega(b) \cdot \bar{h}(a)] = 0 . \tag{9.4}$$

This equation can be solved for $\omega(a)$. Define

$$H(a) \equiv -\bar{h}(\nabla \wedge \bar{h}^{-1}(a)), \tag{9.5}$$

and note that

$$H(a) = -\bar{h}(\grave{\nabla}) \wedge \bar{h}[\,\grave{\bar{h}}^{-1}(a)\,] = \grave{\nabla} \wedge \grave{\bar{h}}[\,\bar{h}^{-1}(a)\,] , \tag{9.6}$$

so (7.4) becomes

$$H(a) = \partial_b \wedge [\,a \cdot \omega(b)\,] . \tag{9.7}$$

Protraction of this equation gives the position gauge invariant quantity

$$H \equiv \partial_b \wedge \omega(b) = -\tfrac{1}{2}\partial_b \wedge H(b) . \tag{9.8}$$

It follows that

$$a \cdot H = \omega(a) - \partial_b \wedge [\,a \cdot \omega(b)\,] .$$

Adding this to (9.7), we obtain the desired result

$$\omega(a) = H(a) + a \cdot H . \tag{9.9}$$

This formula enables us to calculate $\omega(a)$ from a given $\bar{h}$ by first calculating $H(a)$ from (9.5) and then H from (9.8). The simple functional form (9.5) makes $H(a)$ seem almost as important as $\omega(a)$ itself; that impression will be reinforced as we learn more about it.

In the absence of gravity we can choose the gauge $\underline{h} = \underline{1}$, so $\omega(a)$ vanishes everywhere, and the coderivative reduces to the derivative. This is most easily proved from (9.5). Every other $\underline{h}$ can be generated from $\underline{1}$ by a transformation $f(x)$, so its adjoint $\bar{h}$ becomes a gradient and

$$\bar{h}^{-1}(a) = \bar{f}^{-1}(a) = \partial_x f^{-1}(x) \cdot a . \tag{9.10}$$

Inserted into (9.5), this makes $H(a)$ vanish by (2.15), so $\omega(a)$ vanishes by (9.9). Similarly, position gauge invariance of $H(a)$, and hence of $\omega(a)$, can be verified directly by substituting $\bar{h}' = \bar{h}\bar{f}$ into (9.5).

For the geodesic equation (8.28), we see from (9.8) that

$$\dot{v} = v \cdot \omega(v) = v \cdot H(v) , \tag{9.11}$$

so $H(v)$ determines the geodesic directly. Using (9.5) and (3.14) we can write

$$v \cdot H(v) = \underline{h}^{-1}(\dot{x}) \cdot H(v) = -\bar{h}[\dot{x} \cdot (\grave{\nabla} \wedge \grave{\bar{h}}^{-1}(v))] ,$$

and consequently (9.11) can be put in the form

$$\bar{h}^{-1}(\dot{v}) = -\dot{x}\cdot(\grave{\nabla}\wedge\grave{\bar{h}}^{-1}(v))\,.$$

Introducing the metric tensor g defined by (7.7b) and expanding the right hand side algebraically, we obtain

$$\frac{d}{d\tau}\,g(\dot{x}) = \tfrac{1}{2}\grave{\nabla}[\,\dot{x}\cdot\grave{g}(\dot{x})\,]\,. \tag{9.12}$$

Solved for $\ddot{x}$, this becomes

$$\ddot{x} = g^{-1}\big[\tfrac{1}{2}\grave{\nabla}\dot{x}\cdot\grave{g}(\dot{x}) - \dot{g}(\dot{x})\big]\,. \tag{9.13}$$

This may be recognized as the standard formulation for the geodesic equation in terms of the "Christoffel connection." Note how the metric tensor appears when the equation of motion is formulated in terms of a particular spacetime map.

For future use, it is convenient to set down the algebraic properties of the connexion systematically with the help of some definitions. Contraction of (9.9) gives the position gauge invariant

$$\Gamma \equiv \partial_a\cdot\omega(a) = \partial_a\cdot H(a)\,. \tag{9.14}$$

The *traction* (= contraction + protraction) of $\omega(a)$ can therefore be written

$$\partial_a\omega(a) = \Gamma + H\,, \tag{9.15}$$

where the right side can be obtained from $H(a)$ by (9.8) and (9.14).

The dependence of $\omega(a)$ on Γ and H can be made explicit. The derivatives

$$\partial_a\cdot(a\wedge\Gamma) = 3\Gamma \qquad \text{and} \qquad \partial_a\wedge(a\cdot H) = 3H \tag{9.16}$$

enable us to define a "tractionless part" of $H(a)$:

$$H_0(a) \equiv H(a) - \tfrac{1}{3}a\wedge\Gamma + \tfrac{2}{3}a\cdot H\,, \tag{9.17}$$

so that

$$\partial_a H_0(a) = 0\,. \tag{9.18}$$

Therefore

$$\omega(a) = H_0(a) + \tfrac{1}{3}(a\wedge\Gamma + a\cdot H)\,. \tag{9.19}$$

Though this is a position gauge invariant decomposition of $\omega(a)$, the "less complete" decomposition (9.9) appears to be more useful.

Both $H(a)$ and $\omega(a)$ are bivector-valued linear functions of a vector variable, so they have *adjoints* $\underline{H}(B)$ and $\underline{\omega}(B)$ which are linear vector-valued functions of a bivector B. As usual, the adjoint can be defined by

$$H(a)\cdot B = a\cdot\underline{H}(B)\,, \tag{9.20}$$

so that

$$\underline{H}(B) = \partial_a H(a)\cdot B\,. \tag{9.21}$$

From (9.6) and (9.9) we find that

$$\underline{H}(a \wedge b) = \underline{h}^{-1}[b \cdot \grave{\nabla}\, \underline{\grave{h}}(a) - a \cdot \grave{\nabla}\, \underline{\grave{h}}(b)] = \omega(a) \cdot b - a \cdot \omega(b)\,. \tag{9.22}$$

This is the adjoint counterpart of (9.5). The counterpart of (9.9) is

$$\underline{\omega}\,(a \wedge b) = \underline{H}(a \wedge b) + H \cdot (a \wedge b)\,. \tag{9.23}$$

Note that (9.22) involves $\underline{h}$ while its counterpart (9.5) involves $\bar{h}$. Thus, the adjoints of the connection reflect adjoints of the gauge. The underbar notations $\underline{H}$ and $\underline{\omega}$ reflect this correspondence. Using the definitions (9.21) and (9.14), for the contraction of (9.23) we find

$$\partial_b \cdot \underline{\omega}\,(a \wedge b) = \partial_b \cdot \underline{H}(a \wedge b) = \Gamma \cdot a\,. \tag{9.24}$$

For the protraction, with the help of (9.7) and the definition (9.21), we find

$$\partial_b \wedge \underline{\omega}\,(a \wedge b) = H(a)\,. \tag{9.25}$$

This suggests that the adjoint of the connexion may have a deeper significance than the connexion itself, but that possibility will not be pursued in this paper.

10. Coderivatives and Integrals

The main task of this section is to ascertain the properties of the coderivative D that derive from the fundamental equation (9.2). In perfect analogy to (2.5), for an arbitrary multivector field $A = A(x)$, we can write

$$DA = D \cdot A + D \wedge A\,. \tag{10.1}$$

where $D \cdot A$ and $D \wedge A$ are respectively called the *codivergence* and the *cocurl*. Their properties will be analyzed separately, because their dependence on the gauge is different.

From the field A another field $A' = A'(x)$ is uniquely determined by writing

$$A = \bar{h}(A')\,. \tag{10.2}$$

Supposing, for the moment, that A is a vector field, direct differentiation using the Leibniz rule and (9.3) gives

$$D \wedge A = \bar{h}(\nabla \wedge A'\,)\,. \tag{10.3}$$

However, this result holds also for multivector fields of arbitrary grade, as is proved easily from the vector case by factoring A and A' into outer products of vectors and exploiting the outermorphism property of $\bar{h}$.

Employing the integrability condition (2.15), from (10.3) we obtain

$$D \wedge D \wedge A = \bar{h}(\nabla \wedge \nabla \wedge A') = 0\,. \tag{10.4}$$

Replacing the A in this equation by Ai and using the duality of inner and outer products to remove the i, we obtain

$$(D \wedge D) \cdot A = D \cdot (D \cdot A) = 0 \qquad \text{for} \qquad (\text{grade } A) \geq 2\,. \tag{10.5}$$

Since $D \wedge D = D \times D$ is algebraically a bivector operator, the identity (1.37) enables us to combine (10.4) and (10.5) into the single equation

$$D \wedge DA = (D \wedge D) \times A\,. \tag{10.6}$$

As will be seen in the next section, the value of the right side depends on the curvature tensor, but for a scalar field it vanishes and we recover (9.2).

The explicit dependence of the codivergence $D \cdot A$ on the gauge can be derived from (10.3) by exploiting duality. The derivation is essentially the same as that for (5.19) and yields the result

$$D \cdot A = h\underline{h}^{-1}[\nabla \cdot (h^{-1}\underline{h}(A))] = \overline{\nabla} \cdot A + \underline{\Gamma}(A)\,, \tag{10.7}$$

where,

$$\underline{\Gamma}(A) \equiv \partial_a \cdot \underline{H}_a(A) - (\overline{\nabla} \ln h) \cdot A\,, \qquad \text{and} \qquad h = \det \underline{h} = \det \bar{h}\,. \tag{10.8}$$

Some new notation has been introduced here, because it is useful in other contexts. It is motivated and explained in the next paragraph and is then used to derive more explicit expressions for $\underline{\Gamma}$.

The position gauge invariant *differential* $a \cdot \overline{\nabla}\,\underline{h}$ appears so frequently that it is convenient to have more compact notations for it; so write

$$\underline{h}_a \equiv a \cdot \overline{\nabla}\,\underline{h} = \underline{h}\underline{H}_a\,, \tag{10.9}$$

with

$$\underline{H}_a = \underline{h}^{-1}\underline{h}_a = -\underline{h}_a^{-1}\underline{h}\,. \tag{10.10}$$

The operator $\underline{H}_a$ has already appeared in the important formula (9.22). Here we generalize it. The problem is that, if (10.10) is defined for operations on vectors, it cannot be extended to multivectors by an outermorphism, because *the differential of an outermorphism is not the outermorphism of a differential*. For example, for $\underline{h}(b \wedge c) = \underline{h}(b) \wedge \underline{h}(c)$, we have

$$\underline{h}_a(b \wedge c) = \underline{h}_a(b) \wedge \underline{h}(c) + \underline{h}(b) \wedge \underline{h}_a(c)\,. \tag{10.11}$$

To make the right side an explicit function of $b \wedge c$ and then generalize to an arbitrary multivector A, we write

$$\underline{h}_a(A) = \big[\underline{h}_a\underline{h}^{-1}(c)\big] \wedge \big[\partial_c \cdot \underline{h}(A)\big] = \underline{H}_a(\partial_c) \wedge \big[\,\bar{h}(c) \cdot \underline{h}(A)\,\big]\,. \tag{10.12}$$

It may be checked that (10.11) is recovered for $A = b \wedge c$. The important point is that $\underline{h}_a(A)$ denotes the differential of the outermorphism $\underline{h}(A)$, so (10.9) and (10.10) can be regarded as applicable to outermorphisms, not just to the type 2-1 tensor $\underline{h}(b)$. The messy expression (10.12) is cleaned up by operating on it with the outermorphism $\underline{h}^{-1}$ and using (3.14), to get

$$\underline{H}_a(A) = \underline{H}_a(\partial_c) \wedge (c \cdot A) = \underline{H}_a(c) \wedge (\partial_c \cdot A)\,. \tag{10.13}$$

This expresses the differential of an outermorphism in terms of the differential for a vector $\underline{H}_a(c)$. A simple but important application of (10.13) with (10.10) is to the outermorphism $\underline{h}(i) = hi$. Thus,

$$\underline{H}_a(i) = h^{-1}h_a i = \underline{H}_a(\partial_c) \cdot ci\,,$$

whence

$$a \cdot \not\nabla (\ln h) = \partial_c \cdot \underline{H}_a(c) = \mathrm{Tr}\, \underline{H}_a\,. \tag{10.14}$$

The results and notations in this paragraph apply to any type 2-1 tensor field, so (10.14) is a general relation between the derivative of its determinant and the trace of its differential.

In the equation

$$DA = \not\nabla A + \partial_a \omega(a) \times A\,, \tag{10.15}$$

the term on the right gives an alternative way to evaluate the operator $\underline{\Gamma}(A)$ in (10.7). Just like $\underline{H}_a(A)$, the factor $\omega(a) \times A$ is a "derivation," so it can be evaluated in much the same way. Accordingly, we write

$$\omega(a) \times A = [\omega(a) \cdot \partial_b] \wedge (b \cdot A) = [\,\omega(a) \cdot b\,] \wedge (\partial_b \cdot A)\,. \tag{10.16}$$

Then

$$\begin{aligned}
\partial_a \cdot [\omega(a) \times A] &= [\,\partial_a \cdot \omega(a) \cdot \partial_b\,] \wedge (b \cdot A) - [\,\omega(a) \cdot b\,] \wedge [\,(\partial_a \wedge \partial_b) \cdot A\,] \\
&= \Gamma \cdot A - \tfrac{1}{2} \underline{H}(a \wedge b) \wedge [(\partial_a \wedge \partial_b) \cdot A] \\
&= \Gamma \cdot b\,(\partial_b \cdot A) - \underline{H}_a(b) \wedge [\,(\partial_a \wedge \partial_b) \cdot A\,] \\
&= (b \cdot \not\nabla \ln h)(\partial_b \cdot A) + \partial_a \underline{H}_a(A)\,,
\end{aligned}$$

where (9.22), (10.14) and (10.13) have been used, and it has been ascertained that

$$\Gamma = \partial_a \cdot H(a) = \overline{H}_a(\partial_a) - \partial_c \partial_a \cdot \underline{H}_c(a) = \grave{\bar{h}}\,(\grave{\nabla}) - \not\nabla \ln h\,, \tag{10.17}$$

where $\overline{H}_a$ is the vector-valued adjoint of $\underline{H}_a$.

Finally, we can express $\underline{\Gamma}(A)$ in two alternative explicit forms:

$$\begin{aligned}
\underline{\Gamma}(A) &= \Gamma \cdot A + \underline{H}(a \wedge b) \wedge [\,\tfrac{1}{2}(\partial_b \wedge \partial_a) \cdot A\,] \\
&= (\not\nabla \ln h) \cdot A + \partial_a \cdot [\,\underline{H}_a(\partial_b) \wedge (b \cdot A)\,]\,,
\end{aligned} \tag{10.18}$$

where it is understood that the last term vanishes for (grade A) < 2. The appearance of Γ in (10.18) motivates the notation $\underline{\Gamma}$. Indeed, for a *vector field* (10.7) reduces to

$$D \cdot A = h \nabla \cdot (h^{-1}\, \underline{h}(A)) = \not\nabla \cdot A + \Gamma \cdot A\,. \tag{10.19}$$

For a bivector field $F = F(x)$, (10.18) puts (10.7) in the form

$$D \cdot F = \not\nabla \cdot F + \underline{H}(F) + \Gamma \cdot F\,. \tag{10.20}$$

By duality, with $N = ai$, (10.19) yields

$$D \wedge N = \not\nabla \wedge N + \Gamma \wedge N\,. \tag{10.21}$$

Of course, this holds for any trivector N.

Similar general results can be obtained for the cocurl. From (10.16) with (9.7), we obtain

$$\partial_a \wedge [\,\omega(a) \times A\,] = [\,\partial_a \wedge (\omega(a) \cdot b)\,] \wedge (\partial_b \cdot A) = -H(b) \wedge (\partial_b \cdot A)\,. \tag{10.22}$$

Hence (10.15) gives

$$D \wedge A = \not\nabla \wedge A - H(b) \wedge (\partial_b \cdot A)\,. \tag{10.23}$$

For a vector field this reduces to

$$D \wedge A = \not\nabla \wedge A - H(A)\,. \tag{10.24}$$

This result has important consequences.

A vector field $a = a(x)$ is said to be *constant* (independent of x) if $b \cdot \not\nabla a = 0$ for all x and any vector b. It follows that $\not\nabla \wedge a = 0$, and (10.24) gives

$$H(a) = -D \wedge a\,. \tag{10.25}$$

Of course, this is not a gauge covariant quantity because the assumption that a is constant does not allow a to be a gauge covariant field. Combining (10.25) with (10.19) we have

$$Da = \Gamma \cdot a - H(a)\,. \tag{10.26}$$

Thus, the coderivative of a constant vector is not zero. Indeed, *"pure geometry"* is generated by differentiating constant vectors. Equation (10.26) shows the fundamental significance of $H(a)$ and Γ — though we know that Γ can be derived from $H(a)$. In fact, we could have adopted (10.25) as a *defining property* of the coderivative together with

$$D \wedge H(a) = -D \wedge D \wedge a = 0\,, \tag{10.27}$$

which, instead, we obtain here as a consequence of (9.5) and (10.14). Note that, in effect, the connexion was defined in the preceding section by

$$b \cdot D\,a = \omega(b) \cdot a\,. \tag{10.28}$$

Thus, $\omega(a)$ relates to the directional coderivative as $H(a)$ relates to the vector coderivative. If the operator D is regarded as more fundamental than $a \cdot D$, then $H(a)$ and (10.23) are more fundamental than $\omega(a)$ and (10.28). A reason for starting with $H(a)$ instead of $\omega(a)$ is the simplicity of the "integrability condition" (10.27). Of course, (9.9) makes it easy to translate between $H(a)$ and $\omega(a)$, so we have the benefit of both "points of view."

Another important consequence of (10.24) comes from applying it to a nonconstant vector field $v = v(x)$. Using the identity

$$v \cdot (D \wedge v) = v \cdot D\,v - \tfrac{1}{2} D v^2\,, \tag{10.29}$$

and assuming constant v^2, we obtain from (10.24)

$$v \cdot (D \wedge v) = v \cdot D\,v = v \cdot \not\nabla\, v - v \cdot H(v)\,. \tag{10.30}$$

This agrees with (8.28) and (9.11), so it can be used to find the "integral curves" of a given vector field.

The above results have important applications to integral theorems. Putting (10.19) into (6.9), we get the scalar version of *Gauss's Theorem* in the manifestly gauge invariant form

$$\int D\cdot A\, h^{-1}\,|\,d^4x\,| = \oint A\cdot \bar{h}(n^{-1})\, h^{-1}\,|\,d^3x\,|\,, \tag{10.31}$$

where A is a vector field. The explicit appearance of the gauge here comes from putting the directed measures (6.8) in the gauge covariant form

$$\underline{h}^{-1}(d^4x) = ih^{-1}\,|\,d^4x\,|\,, \tag{10.32}$$

$$\underline{h}^{-1}(d^3x) = \bar{h}^{-1}(in^{-1})\,|\,d^3x\,| = h^{-1}i\bar{h}(n^{-1})|\,d^3x\,|\,. \tag{10.33}$$

Of course, the gauge covariant form $\bar{h}(n^{-1})$ for the normal is necessary to make the value of the integral independent of the chosen spacetime map.

To get a version of Gauss's theorem for a general tensor field $T(a)$ which makes the role of the connexion explicit, use (10.17) to write

$$\grave{h}^{-1}\grave{T}\bar{h}(\grave{\nabla}) = h^{-1}\,[\,\grave{T}(\grave{\nabla}) + T(\Gamma)\,]\,. \tag{10.34}$$

By inserting this into (6.9), we get the *general Gauss's theorem* in the form

$$\int [\grave{T}(\grave{\nabla\!\!\!/}) + T(\Gamma)]\, h^{-1}\,|\,d^4x\,| = \oint T(\bar{h}n^{-1})\, h^{-1}\,|\,d^3x\,|\,. \tag{10.35}$$

This formula is position gauge invariant, but it is not rotation gauge covariant although it does admit a kind of "asymptotic rotation gauge invariance" where the rotor field is assigned a fixed value on the boundary. Evidently such a condition is essential for a global energy-momentum conservation law.

The formula (10.35) can be expressed in terms of the coderivative by using the relation

$$\grave{T}(\grave{\nabla\!\!\!/}) = \grave{T}(\grave{D}) - \omega(a)\times T(\partial_a)\,. \tag{10.36}$$

Remarkably, equation (10.35) is identical in form to a result derived in [3] for vector manifolds, although the analogue of Γ there is orthogonal to the tangent space of the manifold.

To express the Generalized Stokes' Theorem (6.5) as an explicit function of the gauge covariant measure $\underline{h}^{-1}(d^kx)$, we define a gauge invariant k-form $M(\langle\, A\,\rangle_k)$ by writing (6.1) in the form

$$M = L(d^kx) = M(\,\underline{h}^{-1}d^kx)\,. \tag{10.37}$$

Since

$$\underline{h}^{-1}((d^{k+1}x)\cdot\nabla) = [\,\underline{h}^{-1}d^{k+1}x\,]\cdot\bar{h}(\nabla) = [\,\underline{h}^{-1}d^{k+1}x\,]\cdot\nabla\!\!\!/\,, \tag{10.38}$$

the exterior differential (6.3) can be written

$$dM = \grave{M}[\,\grave{\underline{h}}^{-1}((d^{k+1}x)\cdot\grave{\nabla\!\!\!/})\,] + M[\,\underline{h}^{-1}((d^{k+1}x)\cdot\Gamma_0)\,]\,, \tag{10.39}$$

where

$$\Gamma_0 = \grave{\bar{h}}(\grave{\nabla}) = \Gamma + \nabla\!\!\!/\ln h\,. \tag{10.40}$$

Consequently *Stokes' Theorem* (6.5) retains the same form as before:

$$\int dM = \oint M\,. \tag{10.41}$$

When M is scalar-valued there exists a k-vector field $A = \bar{h}(A')$ so that

$$M = (\underline{h}^{-1} d^k x) \cdot A = (d^k x) \cdot A'\,. \tag{10.42a}$$

Moreover, by virtue of (10.3), the exterior differential (10.39) simplifies to

$$dM = (\underline{h}^{-1} d^{k+1} x) \cdot (D \wedge A) = (d^{k+1} x) \cdot (\nabla \wedge A')\,. \tag{10.42b}$$

This is exactly the result (6.6a,b) obtained before gauge invariance was incorporated in the formulation. It shows that the cocurl is equivalent to a curl in Stokes' Theorem for scalar-valued forms, but the role of the gauge should be noted.

The above results enable us to make electrodynamics positionally and directionally gauge invariant, and thus to incorporate the influence of gravity on the formation and propagation of electromagnetic fields. The generalization is achieved simply by replacing the vector derivative by the coderivative in (2.4). Thus, we obtain the gauge covariant form of *Maxwell's equation*

$$DF = J\,. \tag{10.43}$$

This separates into trivector and vector parts:

$$D \wedge F = 0\,, \tag{10.44a}$$

$$D \cdot F = J\,. \tag{10.44b}$$

By virtue of (10.19), the latter implies local charge conservation:

$$D \cdot J = h \nabla \cdot (h^{-1}\, \underline{h} J) = 0\,. \tag{10.45}$$

Gauss's theorem (10.31) puts it in the integral form

$$\oint J \cdot \bar{h}(n^{-1})\, |\, d^3 x \,| = 0\,. \tag{10.46}$$

By virtue of (10.4), (10.44a) allows us to derive F from a vector potential:

$$F = D \wedge A\,. \tag{10.47}$$

If we adopt the "Lorentz condition"

$$D \cdot A = 0\,, \tag{10.48}$$

then $F = DA$, and Maxwell's equation (10.42) can be put in the form

$$D^2 A = D \cdot DA + R(A) = J\,, \tag{10.49}$$

where, as will be seen in the next section,

$$R(A) = D \wedge DA \tag{10.50}$$

is the Ricci tensor.

One interpretation of (10.2) is as describing the effect of the gravitational potential $\bar{h}$ acting on a flat space vector potential A' to produce the "physical field" A. On the other hand, it describes the relation of a map dependent representation A' to the position gauge invariant representation A, and its only use has been in deriving properties of D from properties of ∇. The gravitational field can be regarded as entering entirely through D rather than partially through A.

Evidently $D \cdot D$ is a generalization of the d'Alembertian, so let us call it the *co-d'Alembertian*. This operator can be expressed as an explicit function of the vector derivatives with the help of (10.7). For example, the *scalar wave equation* can be written

$$D \cdot D\varphi = [\,\not\nabla + \Gamma\,] \cdot \not\nabla \varphi = [\,g^{-1}(\nabla) + \Gamma_g\,] \cdot \nabla \varphi = 0\,. \tag{10.51}$$

where g is the metric tensor and

$$\Gamma_g \equiv \underline{h}(\Gamma) + \grave{\underline{h}}(\grave{\not\nabla}) = \grave{g}^{-1}(\grave{\nabla}) + g^{-1}(\nabla \ln h)\,. \tag{10.52}$$

Again we see the appearance of the metric tensor when an equation is put in covariant form.

11. Curvature

This section analyzes and summarizes properties of the (Riemann) *curvature* $R(a \wedge b)$. As defined by (8.20), it is a tensor of type 4-2. It can also be regarded as a linear bivector-valued function of a bivector variable B, as determined by

$$R(B) \equiv \tfrac{1}{2} B \cdot (\partial_b \wedge \partial_a)\, R(a \wedge b)\,. \tag{11.1}$$

The curvature tensor is related to the directional coderivative by (8.18), which, for vector fields a and b satisfying $[a, b] = 0$ reduces to

$$[a \cdot D, b \cdot D]A = R(a \wedge b) \times A \tag{11.2}$$

The existence of such fields is established in Section 13. They are employed here to simplify the derivation of properties of the curvature tensor from the coderivative, after which the results hold for arbitrary fields.

As was noted in Section 9, (11.2) amounts to the "first fundamental equation" of Riemannian geometry when A is a scalar field. When A is not a scalar, it becomes the "second fundamental equation." The fact that the right side of (11.2) is a linear function of A shows that the equation does not depend on whether A is constant; its content depends only on the grade of A. Recall from (10.26) that the coderivative of a constant vector is not zero.

Results of the preceding section which were derived from the "first fundamental equation" imply restrictions on the curvature tensor. These can be readily expressed by reformulating (11.2) as a condition on the vector coderivative D. For a vector c the commutator product is equivalent to the inner product and (11.2) becomes

$$[a \cdot D, b \cdot D]c = R(a \wedge b) \cdot c \,. \tag{11.3}$$

To reformulate this as a condition on the vector coderivative, we simply eliminate the variables a and b by traction. Protraction of (11.3) gives

$$\partial_b \wedge [\, a \cdot D, b \cdot D\,]c = \partial_b \wedge [\, R(a \wedge b) \cdot c\,] = R(c \wedge a) + c \cdot [\, \partial_b \wedge R(a \wedge b)\,] \,. \tag{11.4}$$

Another protraction together with

$$D \wedge D = \tfrac{1}{2}(\partial_b \wedge \partial_a)[\, a \cdot D, b \cdot D\,] \tag{11.5}$$

gives

$$D \wedge D \wedge c = [\, \partial_b \wedge \partial_a \wedge R(a \wedge b)\,] \cdot c + \partial_a \wedge R(a \wedge c) \,. \tag{11.6}$$

The left side of this equation vanishes by (10.4), and since the terms on the right have different functional dependence on the free variable c, they must vanish separately. Therefore

$$\partial_a \wedge R(a \wedge b) = 0 \,. \tag{11.7}$$

This constraint on the curvature tensor is called the *Ricci identity*.

The requirement (11.7) that the curvature tensor is *protractionless* has an especially important consequence. The identity

$$\partial_b \wedge [B \cdot (\partial_a \wedge R(a \wedge b))] = \partial_b \wedge \partial_a B \cdot R(a \wedge b) - B \cdot (\partial_b \wedge \partial_a) R(a \wedge b) \tag{11.8}$$

vanishes on the left side because of (11.7), and the right side then implies that

$$A \cdot R(B) = R(A) \cdot B \,. \tag{11.9}$$

Thus, the curvature is a *symmetric* bivector function. This symmetry can be used to recast (11.7) in the equivalent form

$$e \cdot R\big((a \wedge b \wedge c) \cdot \partial_e\big) = 0 \,. \tag{11.10}$$

On expanding the inner product in its argument, it becomes

$$a \cdot R(b \wedge c) + b \cdot R(c \wedge a) + c \cdot R(a \wedge b) = 0 \,, \tag{11.11}$$

which is closer to the usual tensorial form for the Ricci identity.

Contraction of the curvature tensor defines the *Ricci tensor*

$$R(a) \equiv \partial_b \cdot R(b \wedge a) \,. \tag{11.12}$$

The Ricci identity (11.7) implies that we can write

$$\partial_b \cdot R(b \wedge a) = \partial_b R(b \wedge a)\,, \tag{11.13}$$

and also that the Ricci tensor is protractionless:

$$\partial_a \wedge R(a) = 0\,. \tag{11.14}$$

This implies the symmetry

$$a \cdot R(b) = R(a) \cdot b\,. \tag{11.15}$$

An alternative expression for the Ricci tensor is obtained by operating on (11.3) with (11.5) and establishing the identity

$$\tfrac{1}{2}(\partial_a \wedge \partial_b) \cdot [\, R(a \wedge b) \cdot c\,] = R(c)\,. \tag{11.16}$$

The result is

$$D \wedge D\, a = (D \wedge D) \cdot a = R(a)\,. \tag{11.17}$$

This could be adopted as a definition of the Ricci tensor directly in terms of the coderivative.

Equation (11.17) shows the fundamental role of the operator $D \wedge D$, but operating with it on a vector gives only the Ricci tensor. To get the full curvature tensor from $D \wedge D$, one must operate on a bivector. To that end, we take $A = a \wedge b$ in (11.2) and use (10.6) to put it in the form

$$D \wedge D(a \wedge b) = D \wedge D \times (a \wedge b) = \tfrac{1}{2}(\partial_d \wedge \partial_c) \times [\, R(c \wedge d) \times (a \wedge b)\,]\,.$$

Although the commutator products make it possible to use the Jacobi identity, a fair amount of algebra is involved in reducing the right side of this equation. The result is

$$D \wedge D(a \wedge b) = R(a) \wedge b + a \wedge R(b) - 2R(a \wedge b)\,, \tag{11.18}$$

or equivalently

$$2R(a \wedge b) = (D \wedge Da) \wedge b + a \wedge (D \wedge Db) - D \wedge D(a \wedge b)\,. \tag{11.19}$$

This is the desired expression of the curvature tensor in terms of $D \wedge D$.

Contraction of the Ricci tensor defines the *scalar curvature*

$$R \equiv \partial_a R(a) = \partial_a \cdot R(a)\,, \tag{11.20}$$

where (10.15) has been used. Since $R(a \wedge b)$, $R(a)$, and R can be distinguished by their arguments, there is no danger of confusion from using the same symbol R for each.

Besides the Ricci identity, there is one further general constraint on the curvature tensor, which can be derived as follows. The commutators of directional coderivatives satisfy the Jacobi identity

$$[\,a \cdot D, [\,b \cdot D, c \cdot D]] + [\,b \cdot D, [\,c \cdot D, a \cdot D]] + [\,c \cdot D, [\,a \cdot D, b \cdot D]] = 0\,. \tag{11.21}$$

By operating with this on an arbitrary nonscalar A and using (11.2), we can translate it into a condition of the curvature tensor that is known as the *Bianchi identity:*

$$a \cdot DR(b \wedge c) + b \cdot DR(c \wedge a) + c \cdot DR(a \wedge b) = 0\,. \tag{11.22}$$

Like the Ricci identity (11.10), this can be expressed more compactly as

$$\grave{R}[\,(a \wedge b \wedge c) \cdot \grave{D}\,] = 0\,. \tag{11.23}$$

"Dotting" it with a free bivector B, we obtain

$$\grave{R}[\,(a \wedge b \wedge c) \cdot \grave{D} \cdot B = (a \wedge b \wedge c) \cdot (D \wedge R(B))\,.$$

Therefore the Bianchi identity can be expressed in the compact form

$$\grave{D} \wedge \grave{R}(a \wedge b) = 0\,. \tag{11.24}$$

This condition on the curvature tensor is the source of general conservation laws in General Relativity.

Contraction of (11.24) with ∂_a gives

$$\grave{R}(\grave{D} \wedge b) - D \wedge R(b) = 0\,. \tag{11.25}$$

A second contraction yields

$$\grave{G}(\grave{D}) = \grave{R}(\grave{D}) - \tfrac{1}{2}DR = 0\,, \tag{11.26}$$

where

$$G(a) \equiv R(a) - \tfrac{1}{2}aR \tag{11.27}$$

is the Einstein tensor.

In General Relativity, for a given *energy-momentum tensor* $T(a)$, the spacetime geometry is determined by *Einstein's equation*

$$G(a) = \kappa T(a)\,. \tag{11.28}$$

The contracted Bianchi identity (11.27) implies the generalized *energy-momentum conservation* law

$$\grave{T}(\grave{D}) = 0\,. \tag{11.29}$$

This is not a conservation law in the usual sense, because, as (10.36) explicitly shows, it is not generally a perfect divergence and so is not convertible to a surface integral by Gauss's theorem (10.35).

To solve Einstein's equation (11.28) for a given energy-momentum tensor, Einstein's tensor $G(a)$ must be expressed so as to make (11.28) a differential equation for the gauge. A direct expression for $G(a)$ in terms of the gauge and its derivatives is very complicated and its structure is not very transparent. A simpler approach has been developed in [1,2]. First a gauge is chosen to make the gauge $\bar{h}$ and its derivatives as simple as possible, in particular by incorporating symmetries of the given physical situation. Then (9.5) and

(9.9) can be used to translate these simplifications into conditions on the functional form of the connexion $\omega(a)$. Note that introduction of $\underline{H}(a\wedge b)$ from (9.22) allows us to express the curvature tensor (8.20) in the form

$$R(a\wedge b) = a\cdot\not\nabla\omega(b) - b\cdot\not\nabla\omega(a) + \omega(a)\times\omega(b) - \omega(\underline{H}(a\wedge b))\,, \tag{11.30}$$

a relation which holds for arbitrary vector fields a and b. Contraction of (11.30) gives the Ricci tensor and the Einstein tensor. Then Einstein's equation and the contracted Bianchi identity become differential equations for the undetermined functional parameters in the connexion.

Let us consider an alternative approach. Using (11.17), we can put Einstein's equation (11.28) in the form.

$$D\wedge Da = \kappa(T(a) + \tfrac{1}{2}a\,\mathrm{Tr}\,T)\,. \tag{11.31}$$

Now the gradient

$$a = D\varphi = \not\nabla\varphi = \bar{h}(\nabla\varphi) \tag{11.32}$$

is a function of the gauge tensor, and let us suppose that φ can be chosen to satisfy the "gauge condition"

$$D\cdot a = D\cdot D\varphi = 0\,. \tag{11.33}$$

Then

$$Da = D\cdot a + D\wedge a = 0\,, \tag{11.34}$$

so that

$$D^2 a = D\cdot Da + D\wedge Da = 0\,. \tag{11.35}$$

That enables us to put (11.31) in the form

$$D\cdot Da = -\kappa(T(a) + \tfrac{1}{2}a\,\mathrm{Tr}\,T)\,. \tag{11.36}$$

This appears to be a simplification in the form of Einstein's equation. Indeed, in the linear approximation it reduces immediately to the wave equation

$$\nabla^2 a = \nabla^2\bar{h}(\nabla\varphi) = -\kappa(T(a) + \tfrac{1}{2}a\,\mathrm{Tr}\,T) \tag{11.37}$$

for the gauge $\bar{h}$. The formulation (11.36) for Einstein's equation was first derived in [4], but it has never been studied to see if its apparent simplicity leads to any practical advantages. Indeed, there may be some difficulty in satisfying the gauge condition.

We close this section with comments on the algebraic classification of curvature tensors. The curvature tensor can be put in the form

$$R(a\wedge b) = W(a\wedge b) + \tfrac{1}{2}[\,a\wedge R(b) + b\wedge R(a)\,] - \tfrac{1}{6}Ra\wedge b\,, \tag{11.38}$$

which implicitly defines the Weyl tensor $W(a\wedge b)$ as its "tractionless part"; thus,

$$\partial_a W(a\wedge b) = \partial_a\cdot W(a\wedge b) + \partial_a\wedge W(a\wedge b) = 0\,. \tag{11.39}$$

The complete *Petrov classification* of Weyl tensors is worked out in language of spacetime calculus in [3] and adapted in [1]. Together with the classification of symmetric tensors given in Section 4 and applied to the Ricci tensor, this gives a complete algebraic classification of Riemann curvature tensors for spacetime geometry. Further details will not be presented here.

Part III. SPACETIME FLOWS

12. Gauge Covariant Flows and Flow Derivatives

This section extends the theory of spacetime transformations in Section 5 to one-parameter families called *flows*. Although flows appear in a variety of physical contexts, we will be most interested in using them to describe the "physical flow (or motion)" through spacetime of a material body or some other physical entity such as electromagnetic radiation. The main problem will be to define a suitable "flow derivative" to describe how quantities change along a flow.

The method developed here should be of considerable interest to physicists and mathematicians, because it enables a complete treatment of transformation group geometry on flat manifolds. In other words, it provides the foundation for a *gauge theory of transformation groups*. Applications will not be pursued here. Rather, we construct the tools of Lie group theory within the context of gauge geometry so they are ready to be applied within the gauge theory of gravitation.

Let $v = v(x)$ be a vector defined on some region of spacetime, possibly on the whole of spacetime, or possibly on some k-dimensional submanifold. A curve $x(\tau)$ is said to be an *integral curve* of the vector field $v(x)$ if

$$\frac{dx(\tau)}{d\tau} = \underline{h} v(x(\tau)) \,. \tag{12.1}$$

As explained in Section 7, the gauge tensor $\underline{h}$ appears in this equation to make it gauge covariant. According to a fundamental theorem in the theory of differential equations, for nonvanishing $\underline{h} v$ equation (12.1) has the unique solution

$$x(\tau) = f(x, \tau) \tag{12.2}$$

for a given initial value $x(0) = f(x, 0)$. Here x is any convenient vector parametrization for the region of interest. The 1-parameter family of transformations $f(x, \tau)$ describes a *congruence* of curves, with a single integral curve through each point of the region. This congruence is called the *flow generated by* v.

Sometimes it is convenient to identify an *arbitrary* x in (12.2) with an initial value for the flow. That choice will be indicated with the subscript notation

$$f_\tau(x) = f(x, \tau) \qquad \text{so that} \qquad f_\tau(f(x, t)) = f(x, t + \tau) \,. \tag{12.3}$$

The function $f_\tau = f_\tau(x)$ can be called the *relative flow* to distinguish it from the *flow* (12.2), though the difference is usually obvious in context. The relative flows have the following properties of a "transformation group":

(a) *Composition:* $$f_\tau \circ f_t = f_{t+\tau}. \tag{12.4a}$$

(b) *Associativity:* $$(f_\tau \circ f_t) \circ f_s = f_\tau \circ (f_t \circ f_s), \tag{12.4b}$$

(c) *Identity:* $$f_0(x) = x, \tag{12.4c}$$

(d) *Inverse:* $$f_\tau^{-1} = f_{-\tau}. \tag{12.4d}$$

Generally the inverse is only a "local inverse," which is to say that, if the parameter τ is restricted to some finite open interval, then an inverse may not exist for "large" values of τ. This is sometimes expressed by saying that the relative flows compose a *pseudogroup.*

Given a vector field $u = u(x)$ defined on the same region as v, the differential $\underline{f}_\tau$ of the flow generated by v determines a gauge covariant transformation of $u(x)$ from any chosen point x to another point along the flow, as defined by

$$\underline{f}_\tau : \quad \underline{h}u(x) \qquad \to \qquad \underline{f}_\tau \, \underline{h}(u(x)) \equiv \underline{h}u'(\,f_\tau(x))\,. \tag{12.5}$$

This transformation is described by saying that $u(x)$ is *dragged along* or *transported* by the flow. In this way the vector $u(x)$ at a given point x is extended to a vector field $u'(\,f_\tau(x))$ defined on the whole integral curve of v through that point.

The vector field $u = u(x)$ is said to be an *invariant of the flow* generated by v if its value at every point along the flow is equal to its "dragged along" value, so that

$$u(\,f_\tau(x)) = \underline{h}^{-1}\,\underline{f}_\tau\,\underline{h}(u(x)) \equiv u'(\,f_\tau(x))\,. \tag{12.6}$$

To measure the deviation from this invariance we define the *flow derivative*, more commonly known as the *Lie derivative* and denoted by $\mathcal{L}_v$. The definition can be given in the equivalent forms

$$\begin{aligned}\mathcal{L}_v u \equiv \underline{h}^{-1}\Big\{ \lim_{\tau\to 0} \frac{1}{\tau}\left[\,\underline{h}u(\,f_\tau(x)) - \underline{f}_\tau\,\underline{h}(u(x))\,\right]\Big\} &= \underline{h}^{-1}\Big\{ \lim_{\tau\to 0} \frac{1}{\tau}\left[\,\underline{f}_\tau^{-1}\,\underline{h}u(\,f_\tau(x)) - \underline{h}u(x)\,\right]\Big\} \\ &= \underline{h}^{-1}\frac{d}{d\tau}\,\underline{f}_\tau^{-1}\,\underline{h}u\Big|_{\tau=0} = \underline{h}^{-1}\underline{f}\,\frac{d}{d\tau}\left[\,\underline{f}^{-1}\,\underline{h}u\,\right].\end{aligned} \tag{12.7}$$

The last form has the advantage of applying to the arbitrary parametrization of points by (12.2) and so holds for any value of τ. The second form simplifies evaluation of the derivative. With the help of (7.17), (9.22) and (8.5). we find

$$\begin{aligned}\mathcal{L}_v u = \underline{h}^{-1}[(\underline{h}v)\cdot\nabla(\underline{h}u) - (\underline{h}u)\cdot\nabla(\underline{h}v)] &= \underline{h}^{-1}[(v\cdot\dot{\nabla}(\underline{h}u) - u\cdot\dot{\nabla}(\underline{h}v)] \\ = v\cdot\dot{\nabla}u - u\cdot\dot{\nabla}v + \underline{H}(u\wedge v) &= v\cdot Du - u\cdot Dv\,.\end{aligned} \tag{12.8}$$

Recalling the definition (8.19b) for the *Lie bracket* and introducing the notation

$$\underline{v}(u) \equiv u\cdot Dv \tag{12.9}$$

for the *codifferential* of a vector field v, we can write

$$\mathcal{L}_v u = [v, u] = \underline{v}(u) - \underline{u}(v)\,. \tag{12.10}$$

The advantage of the codifferential notation (12.9) will appear when its outermorphisms are examined. The Lie derivative can also be expressed in terms of the codivergence using the identity

$$[\,v,u\,] = D\cdot(v\wedge u) - uD\cdot v + vD\cdot u\,. \tag{12.11}$$

The definition (12.7) reduces to the conventional definition of the Lie derivative for constant $\underline{h}$. Ordinarily the same definition is used for curved and flat manifolds. Here, the incorporation of $\underline{h}$ into the definition is an alternative to the usual generalization to curved manifolds. With (12.7) we have, for the first time, a gauge-invariant formulation of the Lie derivative on a flat space. The essential role of the gauge tensor in achieving gauge invariance is a new idea for Lie theory. Among its implications is the likelihood that the theory of manifolds with Lie group structure can be reduced to the study of gauge tensor fields on flat space. In gravitation theory, it is a sharp tool for studying gauge-invariant flows and conserved quantities.

The Lie derivative definition (12.7) generalizes immediately to any k-vector field $A = A(x)$ by interpreting $\underline{f}$ and $\underline{h}$ as outermorphisms. Thus,

$$\mathcal{L}_v A = \underline{h}^{-1}\frac{d}{d\tau}\,\underline{f}_\tau{}^{-1}\,\underline{h}A\Big|_{\tau=0} = \underline{h}^{-1}\underline{f}\frac{d}{d\tau}\,[\,\underline{f}^{-1}\,\underline{h}(A)\,]\,, \tag{12.12}$$

which evaluates to

$$\mathcal{L}_v A = v\cdot DA - \grave{v}\wedge(\grave{D}\cdot A) \equiv [\,v,A\,]\,, \tag{12.13}$$

defining a "generalized Lie-bracket" $[v,A]$. A further generalization of the bracket to arbitrary fields is treated in [3]. Equation (12.11) likewise generalizes to

$$[\,v,A\,] = D\cdot(v\wedge A) - AD\cdot v + v\wedge(D\cdot A)\,. \tag{12.14}$$

For a bivector field $A = a\wedge b$, we find

$$[\,v,a\wedge b\,] = [\,v,a\,]\wedge b + a\wedge[\,v,b\,]\,. \tag{12.15}$$

This generalizes to the "Leibniz property"

$$\mathcal{L}_v(A\wedge B) = (\mathcal{L}_v A)\wedge B + A\wedge(\mathcal{L}_v B)\,. \tag{12.16}$$

Therefore the Lie derivative is a derivation with respect to the outer product. This is easily proved directly from the definition (12.12).

Although the formula (12.13) does not apply to a scalar field, the definition (12.12) does hold, since outermorphisms do not alter scalars. Thus for a scalar field $\varphi = \varphi(x)$ we find

$$\mathcal{L}_v\varphi = v\cdot D\varphi = v\cdot\nabla\!\!\!\!/\,\varphi\,. \tag{12.17}$$

As noted in Section 5, every transformation of spacetime induces two complementary kinds of transformations, contravariant and covariant. The vector $u = u(x)$ is said to be *contravariant*, to associate it with the contravariant transformation (12.5). Similarly a *covariant vector or covector* $w = w(x)$ is associated with the *covariant* transformation

$$\bar{f}_\tau{}^{-1}:\quad \bar{h}^{-1}w(x)\quad\rightarrow\quad \bar{f}_\tau^{-1}\bar{h}^{-1}(w(x)) \equiv \bar{h}^{-1}w'(\underline{f}_\tau(x))\,, \tag{12.18}$$

where, the adjoint $\bar{f}_\tau$ has been defined by (5.5). Accordingly, the Lie derivative of w is defined by

$$\mathcal{L}_v w \equiv \bar{h}\,\frac{d}{d\tau}\,\bar{f}_\tau \bar{h}^{-1} w\,\Big|_{\tau=0} = \bar{h}\bar{f}^{-1}\frac{d}{d\tau}\,\bar{f}\,\bar{h}^{-1} w\,. \tag{12.19}$$

Evaluating this derivative with the help of (5.5), (7.17) and (7.13), we obtain

$$\begin{aligned}\mathcal{L}_v w &= \bar{h}[\grave{\nabla}(\grave{\underline{h}}v)\cdot(\bar{h}^{-1}w) + v\cdot\nabla\!\!\!/\,(\bar{h}^{-1}w)]\\ &= \bar{h}[(\grave{\nabla}\grave{v}\cdot w) + v\cdot\nabla\!\!\!/\,w + \bar{h}[v\cdot\grave{\nabla}\,\grave{\bar{h}}^{-1}(w) - \grave{\nabla}v\cdot(\bar{h}\,\grave{\bar{h}}^{-1}w)]\\ &= \grave{\nabla\!\!\!/}v\cdot\grave{w} + v\cdot\nabla\!\!\!/\,w + v\cdot\bar{h}[\grave{\nabla}\wedge\grave{\bar{h}}^{-1}(w)\,].\end{aligned} \tag{12.20}$$

Finally, using (9.5) and (9.7), this can be put in the form

$$\mathcal{L}_v w = v\cdot\nabla\!\!\!/\,w + \grave{\nabla\!\!\!/}\grave{v}\cdot w + H(w)\cdot v = v\cdot Dw + \grave{D}\grave{v}\cdot w = \underline{w}(v) + \bar{v}(w)\,, \tag{12.21}$$

where the adjoint of the codifferential (12.9) is defined by

$$\bar{v}(w) \equiv \grave{D}\grave{v}\cdot w\,. \tag{12.22}$$

Recall that the accent is used to distinguish quantities that are differentiated from those that are not. In the form (12.21), the Lie derivative of a covector is manifestly gauge covariant. Note that, although the Lie derivatives of vector v and covector w have different forms, their coderivatives do not. In other words, contravariant and covariant vectors are distinguished by their Lie derivatives, but not by their coderivatives.

As in the contravariant case, the covariant transformation rule (12.17) and the Lie derivative definition (12.18) generalize without change from covector field $w = w(x)$ to comultivector field $W = W(x)$ simply by interpreting operators as outermorphisms. Likewise, it is easily proved from the definition that the Leibniz property (12.16) also holds for comultivectors. By factoring W into covectors and applying the Leibniz rule, (12.20) can be used to express its Lie derivative in terms of its coderivative:

$$\mathcal{L}_v W = v\cdot DW - \grave{D}(\grave{v}\cdot W)\,. \tag{12.23}$$

Considering the identity

$$v\cdot(D\wedge W) = v\cdot DW + \grave{D}\wedge(v\cdot\grave{W})\,, \tag{12.24}$$

we observe that if $D\wedge W = 0$, (12.22) can be put in the form

$$\mathcal{L}_v W = D\wedge(v\cdot W)\,. \tag{12.25}$$

This is relevant for Lie derivatives of the electromagnetic field.

The Lie derivative of the inner product $u\cdot w$ of a vector with a covector also satisfies the Leibniz rule. Thus, from (12.10) and (12.21) we find

$$\mathcal{L}_v(u\cdot w) = (\mathcal{L}_v u)\cdot w + u\cdot(\mathcal{L}_v w) = v\cdot D(u\cdot w) = v\cdot\nabla\!\!\!/\,(u\cdot w)\,, \tag{12.26}$$

which, for $\varphi = u \cdot v$, agrees with the formula (12.17) for the Lie derivative of a scalar. In fact, starting from (12.17) and taking $\mathcal{L}_v u$ as given by (12.10), (12.16) could be used instead of (12.19) to define $\mathcal{L}_v w$, with (12.21) as the result. The Leibniz rule (12.26) generalizes in the obvious way to the inner product of a multivector with a comultivector. However, it does not in general apply to the inner product of two vectors or two covectors, as is shown below.

For fields defined on a 4-dimensional region, the properties of covariant fields can be derived from contravariant fields by duality. Introduce first

$$\rho \equiv \det \underline{f}^{-1}\underline{h} = h \det \underline{f}^{-1} \tag{12.27}$$

as an (inverse) *measure of flow volume*. Then note that the definition (12.12) gives

$$\mathcal{L}_v i = i \frac{1}{\rho}\frac{d\rho}{d\tau}\,, \tag{12.28}$$

while (12.13) yields

$$\mathcal{L}_v i = -iD \cdot v\,. \tag{12.29}$$

Therefore

$$D \cdot v = -\frac{1}{\rho}\frac{d\rho}{d\tau} = v \cdot \bar{\nabla} \ln \rho^{-1} = v \cdot D \ln \rho^{-1} \tag{12.30}$$

can be identified as the *volume expansion rate*. This implies the familiar "conservation law"

$$D \cdot (\rho v) = 0\,, \tag{12.31}$$

and supplies it with a geometrico-physical interpretation.

In the preceding paragraph the pseudoscalar i is taken to be contravariant. Accordingly, any contravariant field $A = A(x)$ can be expressed as the dual of a covariant field $W = W(x)$; that is,

$$A = Wi = W \cdot i\,. \tag{12.32}$$

Inserting this in formula (12.13) for the Lie derivative of A and extracting the i, we obtain

$$\mathcal{L}_v W - W(D \cdot v) = v \cdot DW - \grave{v} \cdot (\grave{D} \wedge W)\,,$$

which is indeed equivalent to the formula (12.23) for the Lie derivative of W. As well as providing an alternative approach, this result establishes the overall consistency of our various definitions.

To complete our catalogue of properties of the Lie derivative, we note that, for any multivector fields A, B and scalar field φ,

$$\mathcal{L}_v(A + B) = \mathcal{L}_v A + \mathcal{L}_v B\,, \tag{12.33}$$

$$\mathcal{L}_v(\varphi A) = (v \cdot \bar{\nabla}\varphi)A + \varphi \mathcal{L}_v A\,. \tag{12.34}$$

Finally, the Lie derivative of a tensor field $T(a, b)$ is defined by

$$\mathcal{L}_v T(a, b) \equiv \grave{\mathcal{L}}_v \grave{T}(a, b) = \mathcal{L}_v[\,T(a, b)\,] - T(\mathcal{L}_v a, b) - T(a, \mathcal{L}_v b)\,, \tag{12.35}$$

where the part in square brackets is evaluated as the Lie derivative of a single multivector. This definition preserves tensor type, that is, the rank, grade and degree of a tensor. The arguments a, b are left unchanged because their derivatives are removed by subtraction, as in the analogous definition for the codifferential of a tensor.

For a given algebraic form of $T(a, b)$, the definition (12.35) will give different Lie derivatives, according to the Lie transformation rules for its arguments a and b. For example, for $T(u, w) \equiv u \cdot w$ where u is contravariant and w is covariant, insertion of (12.26) into (12.35) gives

$$\mathcal{L}_v T(u, w) = 0 . \tag{12.36}$$

In general, any tensor T is said to be an *invariant of the flow* if

$$\mathcal{L}_v T = 0 . \tag{12.37}$$

As another example, consider a tensor $T(w) \equiv u \cdot w$ constructed from the same two contravariant and covariant vectors. From the definition (12.35), and (12.26),

$$\mathcal{L}_v T(w) = v \cdot D[\, u \cdot w \,] - u \cdot (\mathcal{L}_v w) = (\mathcal{L}_v u) \cdot w , \tag{12.38}$$

so, in this case, $\mathcal{L}_v T$ is equivalent to $\mathcal{L}_v u$.

Now consider the *metric tensor*

$$G(a, b) \equiv a \cdot b , \tag{12.39}$$

where a and b are *both* contravariant vectors. This is equivalent to the earlier definition (7.7b) for the metric tensor, but here the tensor is expressed in terms of position gauge invariant vectors instead of position gauge dependent ones. For the flow derivative of the metric tensor, we find

$$\mathcal{L}_v G(a, b) = a \cdot [\, \underline{v} + \bar{v} \,](b) = a \cdot (b \cdot D v) + b \cdot (a \cdot D v) . \tag{12.40}$$

This vanishes only in special circumstances, that is, only for special gauge tensors and particular vector fields. A vector field $k = k(x)$ for which $\mathcal{L}_k G = 0$ is called a *Killing vector*. According to (12.40), this can occur only if k satisfies *Killing's equation*

$$a \cdot (b \cdot Dk) + b \cdot (a \cdot Dk) = 0 \tag{12.41a}$$

or equivalently,

$$\underline{k} = -\bar{k} . \tag{12.41b}$$

Following Sobczyk [19] and Eisenhart [20], we catalogue properties of Killing vectors in the language of spacetime calculus.

According to (12.41b), the differential $\underline{k}$ is a skewsymmetric tensor; therefore it is completely determined by its protraction, the cocurl of k. Differentiating (12.41) by ∂_b and then by ∂_a, we obtain

$$a \cdot Dk = a \cdot \Omega = -\grave{D} a \cdot \grave{k} , \tag{12.42a}$$

or equivalently,

$$\underline{k}(a) = a \cdot \Omega = -\bar{k}(a) , \tag{12.42b}$$

where

$$\Omega = \tfrac{1}{2}\partial_a \wedge \underline{k}(a) = \tfrac{1}{2} D \wedge k\,. \tag{12.43}$$

Indeed, we obtain the stronger condition

$$Dk = \partial_a \underline{k}(a) = 2\Omega\,, \tag{12.44}$$

since (12.43) implies that

$$D \cdot k = \mathrm{Tr}\;\underline{k} = 0\,. \tag{12.45}$$

The result is that the coderivative $a \cdot Dk$ is completely determined by the cocurl $D \wedge k$. Later it will be seen that, for timelike flows, the bivector Ω can be identified with *vorticity* and that Killing's equation (12.41) is the condition that the *strain rate* vanishes; in other words, the strain is constant along the flow. This interpretation can be generalized to apply to any flow.

The existence of a Killing vector indicates a *symmetry* of the spacetime geometry. The physical significance of Killing vectors comes from their association of a conservation law with each symmetry. This is established by the following general argument. For the energy momentum tensor $T(a)$ of an isolated physical system, the "generalized conservation law"

$$\grave{T}(\grave{D}) = 0 \tag{12.46}$$

does not translate to a standard *integral conservation law*, because the left side is not a strict divergence. However, if there exists a Killing vector k, then the *energy-momentum flux* along the flow of k

$$P \equiv T(k) \tag{12.47}$$

satisfies

$$D \cdot P = 0\,. \tag{12.48}$$

This translates to an integral conservation law for P by Gauss's theorem (10.31). The proof of (12.48) rests on the fact that $T(a)$ is a symmetric tensor. Thus,

$$D \cdot P = D \cdot T(k) = \grave{T}(\grave{D}) \cdot k + T(D) \cdot k\,.$$

While the first term on the right vanishes by (12.46), the last term vanishes by Killing's equation.

For a uniform (or constant) gauge on spacetime, the curvature tensor vanishes and every constant vector is a Killing vector. Then Γ vanishes in Gauss's theorem (10.35), and (12.46) integrates directly to the integral energy-momentum conservation law

$$\oint T(n^{-1})\;|\,d^3x\,| = 0\,. \tag{12.49}$$

In this case every *constant* bivector Ω determines a Killing vector field, since every vector field a satisfies $a = a \cdot Dx = a \cdot \partial x$ whence (12.42a) integrates to

$$k(x) = x \cdot \Omega + c\,, \tag{12.50}$$

where c is an integration constant depending on the choice of origin. The integral curves of (12.50) are

$$x(\tau) = f_\tau(x) = R\,x_0\widetilde{R} + c\tau\,, \tag{12.51}$$

where $R = R(\tau)$ is the rotor

$$R = e^{\frac{1}{2}\Omega\tau}\,. \tag{12.52}$$

This is a spacetime-filling congruence parametrized by the initial position x_0. It is a one-parameter family of Poincaré transformations parametrized by τ. It can be used to model certain physical systems having constant angular momentum.

For any Killing vector k,

$$v\cdot k = \text{constant} \tag{12.53}$$

along any geodesic generated by a vector field v. The proof is direct:

$$v\cdot D(v\cdot k) = (v\cdot Dv)\cdot k + v\cdot(v\cdot Dk) = v\cdot(v\cdot\Omega) = 0\,.$$

If k itself generates geodesics, then

$$a\cdot Dk^2 = 2k\cdot(a\cdot Dk) = 2k\cdot(a\cdot\Omega) = a\cdot(k\cdot Dk) = 0\,,$$

which implies that

$$k^2 = \text{constant} \tag{12.54}$$

throughout the region, and the flow is called a *translation* by Eisenhart [15]. This generalization of the flat-space concept of "translation" is justified by the following facts: for translations k_1, k_2,

$$k_1\cdot k_2 = \text{constant}\,, \tag{12.55a}$$

so that the angle between k_1 and k_2 is constant. Also, for constant α and β, the vector field

$$k_3 = \alpha k_1 + \beta k_2 \tag{12.55b}$$

is itself a translation, since

$$Dk_3^2 = D(\alpha^2 k_1^2 + \alpha\beta k_1\cdot k_2 + \beta^2 k_2^2) = 0\,.$$

An argument against identifying every geodesic k with a translation is that the relation

$$k\cdot Dk = k\cdot\Omega = 0$$

allows a nonzero Ω, which may impart a "vorticity" to the congruence. It might be better, therefore, to identify translations with geodesic Killing vectors satisfying

$$Dk = D\wedge k = 0\,. \tag{12.56}$$

It was found in Section 11 that such vector fields have an especially simple relation to the curvature tensor.

The existence of a Killing vector k puts restrictions on the curvature tensor. Applying the "curvature equation" (8.18) to k, we obtain

$$R(a \wedge b) \cdot k = a \cdot \underline{\Omega}(b) - b \cdot \underline{\Omega}(a)\,, \tag{12.57}$$

where

$$\underline{\Omega}(a) \equiv a \cdot D\Omega \tag{12.58}$$

is the codifferential of Ω. Dotting (12.57) with c and using the symmetry of the curvature tensor we obtain

$$(a \wedge b) \cdot R(k \wedge c) = a \cdot \underline{\Omega}(b) \cdot c - b \cdot \underline{\Omega}(a) \cdot c\,.$$

Differentiation with $\frac{1}{2}\partial_b \wedge \partial_a$ yields

$$R(k \wedge c) = \partial_a \wedge [\underline{\Omega}(a) \cdot c] = -\grave{D} \wedge (\grave{\Omega} \cdot c) = -c \cdot DD\Omega\,,$$

where the last equality was obtained by using $D \wedge \Omega = D \wedge D \wedge k = 0$ in (12.24). Thus we have

$$R(a \wedge k) = \underline{\Omega}(a)\,. \tag{12.59}$$

Contraction gives the Ricci tensor

$$R(k) = D \cdot \Omega = D\Omega = D^2 k\,. \tag{12.60}$$

More generally, from (12.59) we find that

$$R(k \wedge a) \cdot b = a \cdot \grave{D}(b \cdot \grave{W}) = a \cdot \grave{D}\underline{\grave{k}}(b)\,, \tag{12.61}$$

which is the second codifferential of k.

Finally, we prove that if k_1 and k_2 are Killing vectors, then (in an obvious notation)

$$k_3 = [\,k_1, k_2\,] = k_1 \cdot \Omega_2 + k_2 \cdot \Omega_1, \tag{12.62}$$

where k_3 is a Killing vector with

$$\Omega_3 = \Omega_1 \times \Omega_2 + 2R(k_1 \wedge k_2)\,. \tag{12.63}$$

The proof is by straightforward differentiation followed by use of the Jacobi identity and (12.61). Thus,

$$\begin{aligned} a \cdot Dk_3 &= (a \cdot \Omega_1) \cdot \Omega_2 - (a \cdot \Omega_2) \cdot \Omega_1 + k_1 \cdot \underline{\Omega}_1(a) - k_2 \cdot \underline{\Omega}_2(a) \\ &= a \cdot (\Omega_1 \times \Omega_2) + k_1 \cdot R(a \wedge k_2) - k_2 \cdot R(a \wedge k_1)\,, \end{aligned}$$

where the last two terms are both equal to $a \cdot R(k_1 \wedge k_2)$. The significance of (12.62) is that the Killing vector fields for a given geometry form a Lie algebra, with the Lie bracket as product. For flat spacetime the curvature tensor vanishes, and we obtain the Lie algebra of the Poincaré group. Sobczyk [19] uses geometric algebra to treat the next simplest case, the conformal group.

13. Integrability and Coordinates

This section surveys the main integrability theorems relating multivector fields to curves and surfaces. These theorems are fundamental in the mathematical analysis of fields and field equations in physics, so it is important to further applications to have them formulated in the language of geometric calculus.

Continuing the argument at the beginning of the preceding section, let g_s be the flow generated by the vector field $u = u(x)$, while, as before, f_τ is the flow generated by $v = v(x)$. The question is: When are these flows "layered" into surfaces? The answer is given by the following theorem:

$$\underline{f}_\tau \, \underline{g}_s = \underline{g}_s \, \underline{f}_\tau \qquad \text{iff} \qquad [\, v, u\,] = 0\,. \tag{13.1}$$

In other words, the differentials of two flows commute if and only if the Lie bracket of their generating fields vanishes. This means that the integral curves of u are "preserved" by f_τ, while the integral curves of v are preserved by g_s. More specifically, (13.1) implies that the two parameter function

$$x(s, \tau) \equiv f_\tau \circ g_s(x_0) = g_s \circ f_\tau(x_0) \tag{13.2}$$

describes a 2-dimensional surface passing through a given point x_0. The integral curves of $x(s, \tau)$ sweep out a surface parametrized by "coordinates" s and τ. At each point the tangent vectors to the coordinate curves are given by

$$\partial_s x(s, \tau) = \underline{h} u(x(s, \tau)), \qquad \partial_\tau x(s, \tau) = \underline{h} v(x(s, \tau))\,. \tag{13.3}$$

Using this in (12.8) we find that

$$[\, v, u\,] = \underline{h}^{-1}[\, (\partial_\tau x) \cdot \nabla(\partial_s x) - (\partial_s x) \cdot \nabla(\partial_\tau x)\,] = \underline{h}^{-1}[\, (\partial_\tau \partial_s - \partial_s \partial_\tau) x(s, \tau)\,] = 0\,. \tag{13.4}$$

The vanishing of the Lie bracket is therefore a necessary and sufficient condition for the commutivity of partial derivatives in a (local) parametrization of surfaces swept out by integral curves of u and v.

The bivector $K' = K'(x(s, \tau)) \equiv (\partial_\tau x) \wedge (\partial_s x)$ is everywhere tangent to the surface. It determines a directed area element

$$d^2 x = K' d\tau ds = (d\tau \, \partial_\tau x) \wedge (ds \, \partial_s x)\,. \tag{13.5}$$

This can be used to express a directed integral over the surface as an iterated integral with respect to the scalar parameters. The gauge-covariant area element for the surface is

$$\underline{h}^{-1}(d^2 x) = K d\tau ds = (v d\tau) \wedge (u ds), \tag{13.6}$$

where the bivector

$$K = \underline{h}^{-1}(K') = v \wedge u\,. \tag{13.7}$$

is a "gauge-covariant tangent" to the surface. We rely on context to distinguish between the two kinds of tangent K' and K. The surface parametrized by (13.2) is said to be an *integral surface* of K or of K'.

The bivector field $K = K(x)$ is well-defined throughout the region of interest. Through each point there passes a unique integral surface of K, while the entire region is "filled" with such surfaces. Each surface is called a *leaf* or *folium* of K, and the region is said to be *foliated* by the leaves of K. The foliation of a region by the leaves of a bivector field obviously generalizes the foliation (filling) of a region by a congruence of integral curves (leaves) generated by a vector field.

We aim to generalize the concept of an integral curve to a k-dimensional *integral surface* generated by a k-vector field. To be tangent to a surface a k-vector $K = K(x)$ must be *simple* or *decomposable.* This means that there must exist vector fields $v_i = v_i(x)$ such that

$$K = v_1 \wedge v_2 \wedge \ldots \wedge v_k \,. \tag{13.8}$$

This decomposition is only *local*, however. It may not be possible to find a set of such vector fields covering the whole surface. For example, a 2-sphere has a non-vanishing tangent bivector, but for any two vector fields tangent to it, their outer product $v_1 \wedge v_2$ must vanish at some point. This expresses the fact that a 2-sphere cannot be completely covered by a single coordinate system.

Unlike a vector field, which always has integral curves, a k-vector field may not have an integral surface. If it does, it is said to be *integrable.* The general criterion for integrability of a simple k-vector field is called the *Frobenius Integrability Theorem.* We shall present it in several forms, each of which offers its own insight. To facilitate comparison with standard treatments [19], it is helpful to adapt some of the nomenclature from the language of differential forms.

A k-vector field W is said to be *closed* if $D\wedge W = 0$ and *exact* if there is a field A such that $W = D \wedge A$. According to the *Poincaré Lemma*, if k is differentiable in a simply-connected region, then it is closed if and only if it is exact. More generally, subject to the same conditions, every multivector field $M = M(x)$ has (nonunique) "multivector potentials" A and B such that

$$M = D \wedge A + D \cdot B \,. \tag{13.9}$$

This generalizes the well known *Helmholtz Theorem* of vector analysis [3].

A scalar field $\lambda = \lambda(x)$ is said to be an *integrating factor* for a field $W = W(x)$ if

$$D \wedge (\lambda W) = 0 \,. \tag{13.10}$$

A field which has an integrating factor is also said to be *integrable;* however, as we shall see, this notion of integrability is dual to the one adopted above. Introducing adjectives to distinguish the two complementary kinds of integrability when necessary, we may say that K in (13.8) is *directly integrable*, while W in (13.10) is *normally integrable.* The reason for saying "normally" appears below.

Now we are ready to state and discuss the *Frobenius theorem:*

A simple k-vector field $K = K(x)$ is integrable if and only if any of the following four conditions is satisfied:

(1) *For every vector field $v = v(x)$ satisfying $v \wedge K = 0$,*

$$\mathcal{L}_v K = [\, v, K\,] = 0 \,. \tag{13.11a}$$

(2) *If $u = u(x)$ is also a vector field satisfying $u \wedge K = 0$, then*

$$[\, v, u\,] \wedge K = 0 \,. \tag{13.11b}$$

(3) *The dual of K has an integrating factor so that*

$$D \wedge (\lambda K i) = 0\,. \tag{13.12a}$$

(4) *If $w = w(x)$ is a vector field satisfying $w \cdot K = 0$*, then

$$K \cdot (D \wedge w) = 0\,. \tag{13.12b}$$

The first two of these integrability critera are "direct versions" of the usual Frobenius theorem, while the last two are "dual versions." Accordingly, we discuss them in pairs.

The direct integrability conditions (13.11a) and (13.11b) are essentially the same. The equation $v \wedge K = 0$ can be interpreted as "*v is contained in K*." Then (13.11b) can be given the reading: "if u and v are contained in K, then so is their Lie bracket." Alternatively, it might be better to interpret $v \wedge K = 0$ as "v is tangent to K," because if K is integrable, v is indeed tangent to its integral surfaces. Then (13.11b) can be interpreted as the statement "The tangent vector fields on an integral surface are closed under composition by the Lie bracket." This assertion can be elaborated by noting that the vectors v_i in (13.8) form a complete set of vectors "contained in K," and therefore the Lie bracket closure condition (13.11b) implies that

$$[\, v_i, v_j \,] = \alpha_{ij}^{m}\, v_m\,, \tag{13.13}$$

where the α_{ij}^{m} are scalar-valued functions with $i, j, m = 1, 2, \ldots, k$ (and summation on the repeated index is understood). Equation (13.13) is a "classical form" for the integrability condition. Its equivalence to (13.11a) is easily proved by inserting (13.8) into (13.11a) and using properties of the Lie derivative.

To prove that (13.13) or (13.11b) are sufficient conditions for integrability, one can use them to construct an integral surface through any point x_0, using flows through x_0 to construct a set of coordinate curves for the surface according to the argument at the beginning of this section. To summarize the main idea: the various forms of the integrability condition ensure that Lie transport of "vector fields *in* K" along "integral curves *in* K" remain tangent to integral surfaces of K.

It should be remembered that we are dealing with *position gauge-invariant tangents* K and v_i. As explained before, the ordinary *position gauge-dependent tangents* $\underline{h}^{-1}(K)$ and $\underline{h}^{-1}(v_i)$ are *contravariant* under gauge transformations and Lie transport.

Turn now to the dual versions of the Frobenius theorem (13.12a,b). The condition $w \cdot K = 0$ on the vector w in (13.12b) means that w is *normal* to the integral surfaces of K. The $(n-k)$-vector

$$\hat{K} \equiv K i \tag{13.14}$$

is properly regarded as *the* normal to the integral surfaces, because the condition $w \cdot K = 0$ implies that any normal vector w is "contained in $\hat{K}$." (Of course $n = 4$ for spacetime, but it costs nothing to keep the dimension unspecified for the sake of generality.) Again it should be recognized that $\hat{K}$ is the *position gauge-invariant normal.* The corresponding position gauge-dependent normal is $\underline{h}^{-1}(\hat{K})$, and it is covariant under position gauge transformations and Lie transport.

When $k = n-1$, so that K is a pseudovector, then the integral surfaces are *hypersurfaces* and $\hat{K}$ is a vector, so that every other normal vector w is proportional to it. The condition (13.12b) can then be written $(iw) \cdot (D \wedge w) = 0$ or, equivalently,

$$w \wedge D \wedge w = 0\,. \tag{13.15}$$

We can write $w = \hat{K}$, so that (13.12a) becomes

$$D \wedge (\lambda w) = D(\lambda) \wedge w + \lambda D \wedge w = 0\,, \tag{13.16}$$

which is equivalent to (13.15). It is now clear that (13.10) and (13.12a) are called "normal integrability conditions" because they are conditions on the normals of integral surfaces. The normal integrability condition (13.15) was first formulated for vector fields on Euclidean 3-space by Kelvin in 1851, possibly the first published example of an integrability condition.

In the general case, $\hat{K}$ is simple because K is simple, so locally $\hat{K}$ can be decomposed into $r = n - k$ vector fields $w_i = w_i(x)$:

$$\hat{K} = w_1 \wedge w_2 \wedge \ldots \wedge w_r\,. \tag{13.17}$$

Such a set of linearly independent vector fields is called a *Pfaff system* of rank r. A set of vector fields becomes a Pfaff system by requiring *normal integrability* rather than direct integrability. In terms of the Pfaff system, the integrability condition (13.12b) can be written

$$w_1 \wedge w_2 \wedge \ldots \wedge w_r \wedge D \wedge w_i = 0 \tag{13.18}$$

for $i = 1, 2, \ldots, r$. This is an obvious generalization of (13.15).

By virtue of the Poincaré Lemma, the integrability condition (13.12a) implies that $\hat{K}$ is locally exact, so that

$$\lambda\hat{K} = D \wedge A\,, \tag{13.19}$$

where A is an $(r-1)$-vector field. If $\hat{K} = w$ is a vector, this becomes

$$\lambda w = D\varphi = \nabla\varphi\,, \tag{13.20}$$

where $\varphi = \varphi(x)$ is a scalar field. The equation

$$\varphi(x) = \mu \tag{13.21}$$

describes a 1-parameter family of hypersurfaces, the leaves of the pseudovector field $K = K(x)$.

In the general case, we say that a scalar field φ is a *first integral* of K if

$$K \cdot D\varphi = K \cdot \nabla\varphi = 0\,, \tag{13.22}$$

in other words, if $D\varphi = \nabla\varphi$ is normal to K. A set of $r = n - k$ first integrals $\varphi^i = \varphi^i(x)$ is said to be *maximal* if their gradients are linearly independent. Then

$$(D\varphi^1) \wedge (D\varphi^2) \wedge \ldots \wedge (D\varphi^r) = \lambda\hat{K} \tag{13.23}$$

satisfies the integrability condition (13.12a), and a specific A for (13.19) can easily be written down in r different ways. A maximal set of first integrals characterizes each integral surface of K as the intersection of r hypersurfaces. The foliation of K is an r-parameter family of $(k = n - r)$-dimensional surfaces.

Next we turn to a general treatment of frames and coordinates, both as a practical means to implement integrability conditions, and to clarify points of potential confusion. A set of vector fields $\{e_\mu = e_\mu(x);\ \mu = 0,1,2,3\}$ is said to be a *frame* for a spacetime region if the pseudoscalar field $e = e(x)$, defined by

$$e \equiv e_0 \wedge e_1 \wedge e_2 \wedge e_3, \tag{13.24}$$

does not vanish at any point of the region. A frame $\{e^\mu\}$ reciprocal to the frame $\{e_\mu\}$ is determined by the set of equations

$$e^\mu \cdot e_\nu = \delta^\mu_\nu\,, \tag{13.25}$$

where $\mu,\nu = 0,1,2,3$. These equations can be explicitly solved for the reciprocal vectors e^μ, with the result

$$e^\mu = (-1)^\mu (e_0 \wedge \ldots \wedge \check{e}_\mu \wedge \ldots \wedge e_3)e^{-1}\,, \tag{13.26}$$

where $\check{e}_\mu$ indicates that e_μ is omitted from the product. Moreover,

$$e^{-1} = \frac{e}{e^2} = e^0 \wedge e^1 \wedge e^2 \wedge e^3\,. \tag{13.27}$$

To interrelate derivatives of the frames $\{e_\mu\}$ and $\{e^\mu\}$, we consider

$$(e_\mu \wedge e_\nu)\cdot(D\wedge e^\alpha) = e_\mu \cdot (e_\mu \cdot De^\alpha) - e_\nu \cdot (e_\mu \cdot De^\alpha) = [\,e_\mu, e_\nu\,]\cdot e^\alpha\,, \tag{13.28}$$

where the last step involves differentiating (13.25). Solving for the Lie bracket, we obtain

$$[\,e_\mu, e_\nu\,] = e_\alpha (e_\mu \wedge e_\nu)\cdot(D\wedge e^\alpha) \tag{13.29}$$

This is the integrability condition (13.13), and it gives an explicit expression for the scalar coefficients on the right side of (13.13). Alternatively, (13.28) can be solved for

$$D\wedge e^\alpha = \tfrac{1}{2} e^\mu \wedge e^\nu [\,e_\mu, e_\nu\,]\cdot e^\alpha\,. \tag{13.30}$$

Since the frames $\{e_\mu\}$ and $\{e^\mu\}$ are dually related, the cocurls of the first should be related to codivergences of the second. To derive the relation, note that $e = \pm|\,e\,|i$, where the plus sign means that e has the same orientation as the unit pseudoscalar i. Since i is constant, the duality relation (1.12a) gives

$$|\,e\,|D\cdot(|\,e\,|^{-1}e_\mu) = [\,D\wedge(e_\mu e^{-1})\,]e\,. \tag{13.31}$$

On the other hand, using (13.27) we obtain

$$D\wedge(e_\mu e^{-1}) = (D\wedge e^\nu)\wedge(e_\nu \wedge e_\mu e^{-1}) = (D\wedge e^\nu)\cdot(e_\nu\wedge e_\mu)e^{-1}\,.$$

Inserting this in (13.31), we obtain the desired "duality relations"

$$|\,e\,|\,D\cdot(|\,e\,|^{-1}e_\mu) = (D\wedge e^\nu)\cdot(e_\nu\wedge e_\mu)\,. \tag{13.32}$$

Using (13.28) to express the right side of this expression in terms of Lie brackets, we find that

$$e^\nu\cdot[\,e_\nu, e_\mu\,] = D\cdot e_\mu - e^\nu\cdot(e_\mu\cdot De_\nu)\,, \tag{13.33}$$

whence

$$e_\mu \cdot D \ln |\, e \,| = e^\nu \cdot (e_\mu \cdot D e_\nu) = -e_\nu \cdot (e_\mu \cdot D e^\nu)\,. \tag{13.34}$$

This completes our collection of "differential identities" for arbitrary frames.

A frame $\{e_\mu\}$ is said to be *holonomic* if

$$[\, e_\nu, e_\mu \,] = 0 \tag{13.35}$$

for all its vectors. By virtue of (13.29) or (13.40), this is equivalent to the condition

$$D \wedge e^\mu = 0\,. \tag{13.36}$$

In standard approaches to differential geometry on curved manifolds, coordinate frames are holonomic. It is therefore important to understand why this is not the case here.

As introduced in (7.8), a *coordinate frame* is defined by

$$e_\mu = \partial_\mu x = \underline{h}\, g_\mu\,, \tag{13.37}$$

and its reciprocal frame is given by

$$e^\mu = \partial x^\mu\,, \tag{13.38}$$

where the $x^\mu = x^\mu(x)$ are scalar-valued *coordinate functions*. Besides being not holonomic, a coordinate frame is not gauge covariant. Nevertheless, it possesses all general properties of frames derived above. Also, it is associated with a gauge covariant frame $\{g_\mu\}$. The reciprocal of that frame is given by

$$g^\mu = \bar{h} e^\mu = \bar{h}(\nabla x^\mu) = \bar{\nabla} x^\mu = D x^\mu\,, \tag{13.39}$$

from which it follows immediately that

$$D \wedge g^\mu = 0 \qquad \Longleftrightarrow \qquad [\, g_\nu, g_\mu \,] = 0\,. \tag{13.40}$$

Thus, the gauge-covariant frame is holonomic.

Why not use gauge-covariant frames exclusively? One answer is that coordinate frames can simplify calculations by choosing a gauge that takes advantages of symmetries in a given problem. This is demonstrated conclusively in [2], where the coordinate frame is chosen to exploit spherical symmetry. The best practice is to coordinate the use of both types of frame.

The treatment of coordinate frames is a good place to compare curved-space with flat-space formulations of Riemannian geometry, because coordinates are used in both. The main difference is that the frames $\{e_\mu\}$ and $\{g_\nu\}$ are not distinguished in the curved-space approach, in which

$$\partial_\mu x = g_\mu\,. \tag{13.41}$$

This is perfectly explicit in [5]. Thus, it is $\{g_\nu\}$ in (13.37) that corresponds to a coordinate frame in curved space. It appears, therefore, that the operator $\underline{h}$ in (13.37) describes the effect of projecting a curved-space frame $\{g_\nu\}$ into a flat-space frame $\{e_\mu\}$, and that gauge invariance ensures that the result is independent of how the projection is made. It seems likely that a proof of the equivalence of curved-space to flat-space formulations can be made along these lines.

14. Flow Dynamics and Deformations

This section extends the treatment of flows in the preceding sections towards physical applications. The aim is to show how flows can be used to describe the motions and deformations of material bodies and fields. The general equation of motion for physical flows includes both gravitational and non-gravitational effects. This is extended to a general equation for deformation dynamics.

The location of a *material filament* at a given instant is described by a spacelike curve $x(\lambda) = g_\lambda(x_0)$. As the filament flows through spacetime, it generates a timelike surface

$$x = x(\tau, \lambda) = f_\tau \circ g_\lambda(x_0) \,. \tag{14.1}$$

The particles of the filament are parametrized by λ, and at any *proper time* τ, the tangent to the filament is

$$\partial_\lambda x = \underline{h}(n) \,, \tag{14.2}$$

where the gauge-covariant *tangent* $n = n(x)$ satisfies $n^2 < 0$. The velocity field of particles in the filament is given by

$$\partial_\tau x = \underline{h}(v) \,, \tag{14.3}$$

where the gauge-covariant *velocity* $v = v(x)$ is normalized to $v^2 = 1$. As explained in the preceding section the vector fields $v(x)$ and $n(x)$ generate the flow (14.1) if and only if

$$\mathcal{L}_v(n) = [\, v, n \,] = v \cdot D n - n \cdot D v = 0 \,. \tag{14.4}$$

Consequently, the coderivative (8.27) of n is given by

$$\frac{\delta n}{\delta \tau} = v \cdot Dn = n \cdot Dv = \underline{v}(n) \,. \tag{14.5}$$

This describes the rate of change in n along the flow, including (as explained below) the deformation rate of the filament.

To get an equation of motion for n, consider the second coderivative

$$\frac{\delta^2 n}{\delta \tau^2} = (v \cdot D)^2 n = v \cdot D(n \cdot Dv) = n \cdot D(v \cdot Dv) + [\, v \cdot D, n \cdot D \,] v \,. \tag{14.6}$$

According to (8.18),

$$[\, v \cdot D, n \cdot D \,] v = R(v \wedge n) \cdot v \,, \tag{14.7}$$

where $R(v \wedge n)$ is the curvature tensor. For a material point on the filament subject to a net non-gravitational force F, the equation of motion is

$$\frac{\delta v}{\delta \tau} = v \cdot Dv = F \,, \tag{14.8}$$

and the differential of the force along the filament is

$$\underline{F}(n) \equiv n \cdot DF = n \cdot D(v \cdot Dv)\,. \tag{14.9}$$

Consequently, (14.6) gives us the equation of motion

$$\frac{\delta^2 n}{\delta\tau^2} = (v \cdot D)^2 n = R(v \wedge n) \cdot v + \underline{F}(n)\,. \tag{14.10}$$

This equation measures the relative acceleration of neighboring points in a filament. For a geodesic flow F vanishes, and (14.10) reduces to the equation for *geodesic deviation*. The term $R(v \wedge n) \cdot v$ is a gravitational *tidal force*, while $\underline{F}(n)$ is a non-gravitational *tidal force*.

The right side of (14.10) is a linear function of n, so it is convenient to introduce the operator notation

$$\underline{R}_v(n) \equiv R(v \wedge n) \cdot v\,, \tag{14.11}$$

allowing (14.10) to be written in the more compact form

$$\frac{\delta^2 n}{\delta\tau^2} = (\underline{R}_v + \underline{F})n\,. \tag{14.12}$$

The linear operators on the right sides of (14.5) and (14.12) suggest the following generalization.

Introduce a tensor field $\underline{U}$ defined on the congruence of particle histories such that

$$v = \underline{U}(\gamma_0), \tag{14.13a}$$

$$n = \underline{U}(\gamma_\perp) \tag{14.13b}$$

where γ_0 and $\gamma_\perp$ are constant reference vectors. The tensor $\underline{U}$ is determined by specifying its equations of motion:

$$\dot{\underline{U}} \equiv \frac{\delta \underline{U}}{\delta\tau} = \underline{v}\underline{U}\,, \tag{14.14}$$

$$\ddot{\underline{U}} = (\underline{R}_v + \underline{F})\underline{U}\,. \tag{14.15}$$

The tensor $\underline{U}$ has been defined so that the coderivatives in (14.5) and (14.12) are replaced by ordinary derivatives. These equations can obviously be applied to any material medium, not just the filament that we started out with.

We can adopt (14.14) and (14.15) subject to (14.13a) as defining a *deformation tensor* $\underline{U}$ for any material medium. Then (14.14) describes the *kinematics* of deformation, while (14.15) describes the *kinetics* of deformation. Kinetics is concerned with the interactions producing the deformation while kinematics is not. Together they describe the *dynamics* of deformation.

Equation (14.14) conforms to the interpretation of $\underline{v}$ as the *deformation rate* tensor. To derive an equation of motion for the deformation rate, first differentiate (14.14) to get the kinematic equation

$$\ddot{\underline{U}} = (\dot{\underline{v}} + \underline{v}^2)\underline{U}\,. \tag{14.16}$$

Then eliminate $\ddot{\underline{U}}\underline{U}^{-1}$ from (14.16) and (14.15) to get the desired kinetic equation

$$\dot{\underline{v}} + \underline{v}^2 = \underline{R}_v + \underline{F}\,. \tag{14.17}$$

This equation is analyzed shortly.

To solve the kinematic equation (14.14), we must specify the *initial conditions*

$$\underline{U}_0 = \underline{U}\Big|_{\tau=0}\,. \tag{14.18}$$

We could take $\underline{U}_0 = 1$, but generally it is better to take γ_0 and $\gamma_\perp$ as generic *reference directions* characterizing an *undistorted reference configuration* for the medium or body, with

$$\gamma_0^2 = 1, \qquad \gamma_\perp^2 = -1\,, \qquad \gamma_0 \cdot \gamma_\perp = 0\,. \tag{14.19}$$

Then

$$n_0 = \underline{U}_0\gamma_\perp, \qquad v_0 = \underline{U}_0\gamma_0\,. \tag{14.20}$$

In some applications it may be convenient to interpret the tensor $\underline{U}_0$ as describing a "prestressed" medium. As in section 12, we can write

$$\underline{U} = \underline{U}_\tau \underline{U}_0\,, \tag{14.21}$$

where $\underline{U}_{\tau=0} = \underline{1}$, so $\underline{U}_\tau$ has the properties of a transformation pseudogroup.

The deformation tensor $\underline{U}$ is closely related to the Lie transport defined by (12.5). It describes the combined effects of gravitational and "mechanical" stresses, but without separating contributions from the gauge $\underline{h}$ and the integrable transformation differential $\underline{f}$.

Now suppose that the deformation tensor admits to the "polar decomposition"

$$\underline{U} = \underline{R}\,\underline{S} = \underline{S}'\underline{R}, \tag{14.22}$$

where $\underline{R}$ is a Lorentz rotation and $\underline{S} = \overline{S}$ is a symmetric tensor. Anticipating its physical interpretation, let us call $\underline{S}$ the *stretch tensor*. Then $\underline{S} - \underline{1}$ can be identified as the *strain tensor*, measuring deviation from an "undeformed" reference shape. On physical grounds it seems reasonable to suppose that

$$\underline{S}\gamma_0 = \gamma_0 \qquad \Longleftrightarrow \qquad \underline{S}'v = v\,, \tag{14.23}$$

but this may not be possible in general. Nevertheless, (14.23) implies that

$$\underline{U}\gamma_0 = \underline{R}\gamma_0 = v\,. \tag{14.24}$$

Moreover, according to (3.34), $\underline{R}$ can be decomposed into

$$\underline{R} = \underline{V}\,\underline{Q} = \underline{Q}', \tag{14.25}$$

where $\underline{V}$ is the boost of γ_0 to v and $\underline{Q}$ is a spatial rotation satisfying

$$\underline{Q}\gamma_0 = \gamma_0 \qquad \Longleftrightarrow \qquad \underline{Q}'v = v\,. \tag{14.26}$$

The operator $\dot{\underline{Q}}Q \equiv \underline{W}$ can be interpreted physically as a vorticity tensor (see below for a somewhat different approach). Accordingly, it might be better to identify $\underline{Q}\underline{S}$ as the deformation rather than its boosted version

$$\underline{U} = \underline{V}\underline{Q}\underline{S} \,. \tag{14.27}$$

The time evolution of $\underline{Q}\underline{S}$ can be interpreted directly as describing a deforming body fixed in one place, or, as viewed by a "comoving observer" flowing with the body. On the other hand, the equations of motion (14.14) and (14.15) are very much simpler than corresponding equations for $\underline{Q}\underline{S}$. For that reason it may be better to work with $\underline{U}$ and introduce the decomposition (14.27) only at the end to interpret results.

Before analysis of the deformation tensor $\underline{U}$, an analysis of the deformation rate $\underline{v}$ is called for. In any case, $\underline{v}$ is often of greater physical interest than $\underline{U}$. The analysis of $\underline{v}$ begins with the decomposition

$$\underline{v} = \underline{v}_{\parallel} + \underline{v}_{\perp} \,, \tag{14.28a}$$

$$\underline{v}_{\parallel}(n) \equiv \underline{v}(v)\, v \cdot n \,, \tag{14.28b}$$

$$\underline{v}_{\perp} \equiv \underline{v}\underline{I}_v \,, \tag{14.28c}$$

where $\underline{I}_v$ is a projection operator defined by

$$\underline{I}_v(n) = v(v \wedge n) = n - v\, v \cdot n \,. \tag{14.29}$$

The tensor $\underline{v}_{\parallel}$ is called the *acceleration tensor* because $\underline{v}(v) = v \cdot Dv$ is the gauge-covariant acceleration. The tensor $\underline{v}_{\perp}$ has, perhaps, a better claim to the name "deformation tensor" than the whole tensor $\underline{v}$. It submits to the further invariant decomposition (13.13). Thus,

$$\underline{v}_{\perp} = \underline{\theta} + \underline{W} = \underline{\sigma} + \tfrac{1}{3}\theta\underline{I}_v + \underline{W} \,, \tag{14.30}$$

where

$$\underline{\theta} \equiv \tfrac{1}{2}(\underline{v} + \bar{v})\underline{I}_v = \bar{\theta} \tag{14.31}$$

is the (symmetric) *strain rate* tensor, and

$$\underline{W} \equiv \tfrac{1}{2}(\underline{v} - \bar{v})\underline{I}_v = -\overline{W} \tag{14.32}$$

is the (skewsymmetric) *vorticity* tensor. The trace of $\underline{\theta}$ is the volume *expansion rate*

$$\theta \equiv \operatorname{Tr}\underline{\theta} = \operatorname{Tr}\underline{v} = D \cdot v \,. \tag{14.33}$$

The traceless part of $\underline{\theta}$ is the *shear rate* tensor

$$\underline{\sigma} \equiv \underline{\theta} - \tfrac{1}{3}\theta\underline{I}_v \,. \tag{14.34}$$

The interpretation of θ in (14.33) has already been justified by (12.30).

Note that the flow velocity v is a Killing vector if and only if the strain rate tensor $\underline{\theta}$ vanishes everywhere along the flow. This follows immediately from (12.41b) and (14.31).

The flow velocity cocurl $D\wedge v$ determines the flow vorticity completely, but it implies nothing about the strain rate. To see this, introduce the *vorticity bivector*

$$W \equiv \underline{I}_v(D\wedge v) = v(v\wedge D\wedge v) = iwv = v\cdot(iw) \tag{14.35}$$

and the *vorticity vector*

$$w \equiv -i(v\wedge D\wedge v) = Wvi = (W\wedge v)i\,. \tag{14.36}$$

These satisfy the obvious orthogonality conditions $v\cdot W = 0$ and $v\cdot w = 0$. From (14.35),

$$n\cdot W = \underline{I}_v[(\underline{I}_v n)\cdot Dv - \grave{D}\grave{v}\cdot(\underline{I}_v n)] = \underline{v}(\underline{I}_v n) - \bar{v}(\underline{I}_v n)\,.$$

Therefore, the vorticity tensor is related to the vorticity bivector by

$$\underline{W}(n) = n\cdot W = i(n\wedge v\wedge \bar{v}(v))\,. \tag{14.37}$$

Inversely,

$$W = \tfrac{1}{2}\partial_a\wedge \underline{W}(a)\,. \tag{14.38}$$

Noting that $D\wedge v = v[v\cdot(D\wedge v) + v\wedge D\wedge v)]$, recalling (10.29), and inserting (14.35), we obtain

$$D\wedge v = v\underline{v}(v) + W = v(\underline{v}(v) + iw)\,. \tag{14.39}$$

This is the desired relation between velocity cocurl and vorticity.

From the general kinetic equation (14.17) for $\underline{v}$, coupled equations of motion for $\underline{W}$, $\underline{\theta}$, $\underline{\sigma}$ and θ can be derived. Hawking and Ellis [21] took this approach, but they did not emphasise its general significance. The present formulation in terms of spacetime calculus makes it more attractive and perhaps even of practical value. We use the operator $\underline{I}_v$ to project out the tensor components orthogonal to v from (14.17). First, using (14.28c) and (14.30), we obtain

$$\underline{v}^2 = \underline{v}_\perp^2 = (\underline{\theta} + \underline{W})^2 = \underline{\theta}^2 + \underline{\theta}\,\underline{W} + \underline{W}\,\underline{\theta} + \underline{W}^2\,. \tag{14.40}$$

Also,

$$\underline{I}_v\dot{\underline{v}}\underline{I}_v = \underline{I}_v(\dot{\underline{\theta}} + \dot{\underline{W}})\underline{I}_v\,. \tag{14.41}$$

Therefore (14.17) projects to

$$\underline{I}_v\{\dot{\underline{\theta}} + \dot{\underline{W}} + (\underline{\theta} + \underline{W})^2\}\underline{I}_v = \underline{I}_v\{\underline{R}_v + \underline{F}\}\underline{I}_v\,. \tag{14.42}$$

It will be convenient to omit explicit mention of the projection operators in the rest of the argument. Accordingly, (14.42) can be decomposed into a symmetric part

$$\dot{\underline{\theta}} = \underline{R}_v + \underline{F}_+ - \underline{\theta}^2 - \underline{W}^2 \tag{14.43}$$

and a skewsymmetric part

$$\dot{\underline{W}} + \underline{\theta}\,\underline{W} + \underline{W}\,\underline{\theta} = \underline{F}_-\,, \tag{14.44}$$

where the obvious decomposition $\underline{F} = \underline{F}_+ + \underline{F}_-$ has been introduced.

The first thing to notice about (14.44) is that curvature affects the vorticity only indirectly through its effect on $\underline{\theta}$ in (14.43). More insight is obtained by taking the trace of (14.43), noting that

$$\mathrm{Tr}\ \underline{R}_v = \partial_a \cdot R(v \wedge a) \cdot v = -v \cdot R(v) \tag{14.45}$$

$$\mathrm{Tr}\ \underline{W}^{\,2} = W \cdot (\partial_a \wedge (a \cdot W)) = 2W = -2|\,W\,|^2 \geq 0\,, \tag{14.46}$$

$$\mathrm{Tr}\ \underline{\theta}^2 = \mathrm{Tr}\,(\underline{\sigma} - \tfrac{1}{3}\theta \underline{I}_v)^2 = \mathrm{Tr}\ \underline{\sigma}^2 - \tfrac{1}{3}\theta^2\,. \tag{14.47}$$

Hence the trace of (14.43) becomes

$$\dot{\theta} = -v \cdot R(v) + 2|\,W\,|^2 - \mathrm{Tr}\,\underline{\sigma}^2 - \tfrac{1}{3}\theta^2 + \mathrm{Tr}\,\underline{F}\,. \tag{14.48}$$

This is known as the *Raychaudhuri equation* [11,16]. Hawking and Ellis use it to prove singularity theorems in General Relativity. They note that, for a perfect fluid with mass density μ and pressure p,

$$v \cdot R(v) = 4\pi(\mu + 3p) > 0\,, \tag{14.49}$$

so the first term on the right side of (14.48) induces contraction, as do the third and fourth terms. The second term induces expansion, like a centrifugal force.

The equation (14.44) for vorticity evolution can be simplified by introducing, as an "integrating factor," a symmetric tensor $\underline{S} = \overline{S}$ satisfying

$$\dot{\underline{S}} = \underline{\theta}\,\underline{S} = \underline{S}\,\underline{\theta}\,. \tag{14.50}$$

Accordingly, multiplication of (14.44) on the left and right by $\underline{S}$ yields the simplified equation

$$\frac{d}{d\tau}\,(\underline{S}\,\underline{W}\,\underline{S}) = \underline{S}\,\underline{F}_-\,\underline{S}\,. \tag{14.51}$$

Of course, the $\underline{S}$ here is the stretch tensor introduced in (14.22).

Alternatively, we can derive an equation of motion for the vorticity bivector by taking the protraction of (14.44). With (14.38), the first term on the left gives $\dot{W} = \frac{1}{2}\partial_a \wedge \dot{\underline{W}}\,(a)$, while the right side becomes

$$F_- = \tfrac{1}{2}\partial_a \wedge \underline{F}_-(a). \tag{14.52}$$

To evaluate the protraction of the other two terms, note that W is a simple bivector, so it can be written as the product of vectors: $W = w_1 \wedge w_2$. Using this and introducing $\dot{\underline{S}}$ with (14.50), we obtain

$$\begin{aligned} \tfrac{1}{2}\partial_a \wedge [\,\underline{\theta}\,\underline{W}\,(a) + \underline{W}\,\underline{\theta}\,(a)\,] &= (\underline{\theta}\,w_1) \wedge w_2 + w_1 \wedge (\underline{\theta}\,w_2) = (W \cdot \partial_a) \wedge \underline{\theta}\,(a) \\ &= \underline{S}^{-1}[\,\underline{S}(w_1) \wedge \dot{\underline{S}}(w_2) + \dot{\underline{S}}(w_1) \wedge \underline{S}(w_2)\,] = \underline{S}^{-1}\dot{\underline{S}}(W)\,. \end{aligned} \tag{14.53}$$

This is recognized as the derivative of the outermorphism $\underline{S}(W)$. Assembling the various terms, we obtain the vorticity equation

$$\dot{W} + \underline{S}^{-1}\dot{\underline{S}}(W) = F_-\,. \tag{14.54}$$

Equivalently,

$$\frac{d}{d\tau}\,[\,\underline{S}W\,] = \underline{S}(F_-)\,. \tag{14.55}$$

Of course, this is equivalent to (14.51), but is simpler.

As noted by Hawking and Ellis, for a perfect fluid we have

$$F_- = -\frac{1}{\mu + p}\frac{dp}{d\tau}W. \tag{14.56}$$

In this case (14.55) integrates to *vorticity conservation:*

$$\lambda \underline{S}(W) = \text{constant}\,, \tag{14.57}$$

where

$$\ln \lambda = \int \frac{dp}{\mu + p}\,. \tag{14.58}$$

When $p \neq 0$ there is a relativistic effect: compression does work on the fluid, thereby increasing mass and inertia, so vorticity increases less under compression than otherwise.

Throughout this section the flow velocity v has been taken as timelike. However, the whole treatment is readily adapted to lightlike flows [21].

Acknowledgement: I am indebted to Steve Gull, Chris Doran and Anthony Lasenby for keeping me informed about their work in progress before publication and to Anthony Garrett for improving the form of the manuscript.

References

[1] A. Lasenby, C. Doran & S. Gull. Gravity, Gauge Theories and Geometric Algebra. *Phil. Trans. R. Soc. Lond. A.* (1997).

[2] A. Lasenby, C. Doran & S. Gull. Astrophysical and Cosmological Consequences of a Gauge Theory of Gravity. In N. Sanchez and A. Zichichi (ed.), *Current Topics in Astrofundamental Physics*: *Erice, 1994*, p.359, World Scientific, Singapore (1995).

[3] D. Hestenes and G. Sobczyk. *CLIFFORD ALGEBRA to GEOMETRIC CALCULUS, A Unified Language for Mathematics and Physics*, G. Reidel Publ. Co., Dordrecht/Boston (1984), paperback (1985). Third printing 1992.

[4] D. Hestenes. *Space-Time Algebra.* Gordon & Breach, New York, (1966).

[5] D. Hestenes. Curvature Calculations with Spacetime Algebra. *Int. J. Theo. Phys.*, **25**, 581–88 (1986).

[6] D. Hestenes. The Design of Linear Algebra and Geometry. *Acta Applicandae Mathematicae* **23**, 65–93 (1991).

[7] D. Hestenes. Differential Forms in Geometric Calculus. In F. Brackx *et al.* (eds), *Clifford Algebras and their Applications in Mathematical Physics.* Kluwer: Dordrecht/Boston (1993). p. 269–285.

[8] D. Hestenes. *New Foundations for Classical Mechanics*. G. Reidel Publ. Co., Dordrecht/Boston (1985), paperback (1987). Third printing with corrections (1993).

[9] A. Lasenby, C. Doran, Y. Dabrowski, and A. Challinor. Rotating astrophysical systems and a gauge theory approach to gravity. In N. Sanchez and A. Zichichi, editors, *Current Topics in Astrofundamental Physics, Erice 1996*, page 380. World Scientific, Singapore, 1997.

[10] C. Doran, A. Lasenby, and S. Gull. Physics of rotating cylindrical strings. *Phys. Rev. D*, 54(10):6021,1996.

[11] C. Doran, A. Lasenby, A. Challinor, and S. Gull. Effects of spin-torsion in gauge theory gravity. *J. Math. Phys.*, To appear, 1997.

[12] C. Doran. Integral equations and Kerr-Schild fields I. Spherically-symmetric fields. Submitted to: *Class. Quantum Grav.*, 1997.

[13] C. Doran, A. Lasenby, and S. Gull. Integral equations and Kerr-Schild fields II. The Kerr solution. Submitted to: *Class. Quantum Grav.*, 1997.

[14] Y. Choquet-Bruhat, C. De Witte-Morette & M. Dillard-Bleick. *Analysis, Manifolds and Physics*. North-Holland, New York (1977).

[15] H. Poincaré. *Science and Hypothesis*. (English trans. 1905). Dover, New York (1951).

[16] D. Kramer, H. Stephani, E. Herlt & M. MacCallum. E. Schmutzer (Ed.). *Exact Solutions of Einstein's Field Equations*. Cambridge U. Press, Cambridge (1980).

[17] W. R. Davis. The Role of Tetrad and Flat-Metric Fields in the Framework of the General Theory of Relativity. *Nuovo Cimento* **43B**, 2816–20 (1966).

[18] D. Hestenes. Spinor Approach to Gravitational Motion and Precession. *Int. J. Theo. Phys.*, **25**, 589–98 (1986).

[19] G. Sobczyk. Killing Vectors and Embedding of Exact Solutions in General Relativity. In J. Chisholm & A. Common (eds.), *Clifford Algebras and their Applications in Mathematical Physics.* D. Reidel, Dordrecht/Boston, p. 227–244 (1986).

[20] L. Eisenhart, *Continuous Groups of Transformations*. Dover, New York (1961).

[21] S. Hawking and G. Ellis, *The Large Scale Structure of Space-time.* Cambridge U. Press, Cambridge (1973).