Spacetime Geometry with Geometric Calculus

David Hestenes[1]
Department of Physics and Astronomy
Arizona State University, Tempe, Arizona 85287-1504

Geometric Calculus is developed for curved-space treatments of General Relativity and comparison with the flat-space gauge theory approach by Lasenby, Doran and Gull. Einstein's Principle of Equivalence is generalized to a gauge principle that provides the foundation for a new formulation of General Relativity as a Gauge Theory of Gravity on a curved spacetime manifold. Geometric Calculus provides mathematical tools that streamline the formulation and simplify calculations. The formalism automatically includes spinors so the Dirac equation is incorporated in a geometrically natural way.

I. Introduction

Using *spacetime algebra* [1, 2] in an essential way, Cambridge physicists Lasenby, Doran and Gull have created an impressive new *Gauge Theory of Gravity* (GTG) based on flat spacetime [3, 4]. In my opinion, GTG is a huge improvement over the standard tensor treatment of Einstein's theory of *General Relativity* (GR), both in conceptual clarity and in computational power [5]. However, as the prevailing preference among physicists is for a curved-space version of GR, a debate about the relative merits of flat-space and curved-space versions will no doubt be needed to change the minds of many. This paper aims to contribute to that debate by providing a conceptual and historical bridge between curved and flat space theories couched in the unifying language of *geometric algebra*.

This article sketches the extension of geometric algebra to a *geometric calculus* (GC) that includes the tools of differential geometry needed for a curved-space version of GR. My purpose is to demonstrate the unique geometrical insight and computational power that GC brings to GR, and to introduce mathematical tools that are ready for use in research and teaching [6]. I presume that the reader has some familiarity with standard treatments of GR as well as with geometric algebra as presented in any of the above references, so certain concepts, notations and results developed there are taken for granted here. Additional mathematical tools introduced herein are sufficient to treat any topic in GR with GC.

This article introduces three different formulations of GR in terms of a unified GC that integrates them into a system of alternative approaches. The first is a *coordinate based formulation* that facilitates translation to and from the standard tensor formulation of GR [7]. The second is a deeper *gauge theory*

[1]electronic mail:hestenes@asu.edu

formulation that is the main concern of this paper. The third is an *embedding formulation* that deserves mention but will not be elaborated here. Although our focus is on GR, it should be recognized that the mathematical tools of GC are applicable to any problem in differential geometry.

Recognition that GR should be formulated as a gauge theory has been a long time coming, and it is still relegated to a subtopic in most GR textbooks, in part because the standard covariant tensor formalism is not well suited to gauge theory. Still less is it recognized that there is a connection between gravitational gauge transformations and Einstein's Principle of Equivalence. Gauge theory is the one strong conceptual link between GR and quantum mechanics, if only because it is essential for incorporating the Dirac equation into GR. This is sufficient reason to bring gauge theory to the fore in the formulation of GR.

This article demonstrates that GC is conceptually and computationally ideal for a gauge theory approach to GR — *conceptually ideal*, because concepts of vector and spinor are integrated by the geometric product into its mathematical foundations — *computationally ideal*, because computations can be done without coordinates. Much of this article is devoted to demonstrating the efficiency of GC in computations. Appendix A gives a new treatment of integrability theorems with important aplications to motions and deformations of material bodies and fields in flat space as well as curved space.

On the foundational level, GC and gauge theory provide us with new conceptual resources for reexamining the physical interpretation of GR, in particular, the much-debated Principles of Relativity and Equivalence. The analysis leads to new views on the notions of Special and General Relativity as well as the relation of theory to measurement. The result is a new *Gauge Principle of Equivalence* to serve as the cornerstone for the GC formulation of GR. It is instructive to compare the GR formulation of gauge equivalence given herein with the apparently quite different formulation in GTG [5] to see how subtle is the difference between passive and active interpretations of equivalent transformations.

Finally, to facilitate detailed comparison of flat space and curved space formulations of differential geometry and GR with GC, the correspondence between basic quantities is summarized in Appendix B. The details are sufficient to prove equivalence of these alternative formulations, though no formal proof is given.

II. Spacetime Models

Every *real entity* has a definite location in space and time — this is the fundamental criterion for existence assumed by every scientific theory. In Einstein's Theory of Relativity, the spacetime of real physical entities is a 4-dimensional continuum modeled mathematically by a 4D differentiable manifold $\mathcal{M}^4$. As described in GA2, in the *Theory of Special Relativity* $\mathcal{M}^4$ is identified with a 4D *Minkowski vector space* $\mathcal{V}^4$. This makes it a *flat space model* of spacetime. In this model, spacetime points and vector fields are elements of the same vector space. The *Theory of General Relativity* (GR) employs a *curved space model* of spacetime, which places points and vector fields in different spaces. Our primary task is to describe how to do that with GC.

In the standard definition of a differentiable manifold coordinates play an essential role. Although GC enables a coordinate-free formulation, we begin with a coordinate-based definition of the spacetime manifold, because that provides the most direct connection to standard practice. Moreover, coordinates are often useful for representing symmetries in vector fields.

To be specific, let x be a generic point in the spacetime manifold $\mathcal{M}^4 = \{x\}$, and suppose that a patch of the manifold is parametrized by a set of coordinates $\{x^\mu; \mu = 0, 1, 2, 3\}$, as expressed by

$$x = x(x^0, x^1, x^2, x^3)\,. \tag{1}$$

The *coordinate frame* of tangent vectors $g_\mu = g_\mu(x)$ to the coordinate curves parametrized by the $\{x^\mu\}$ are then given by

$$g_\mu = \partial_\mu x = \frac{\partial x}{\partial x^\mu}\,. \tag{2}$$

At each point x the vectors $g_\mu(x)$ provide a basis for a vector space $\mathcal{V}^4(x)$ called the *tangent space to* $\mathcal{M}^4$ *at* x. The vectors in $\mathcal{V}^4(x)$ do not lie in $\mathcal{M}^4$. To visualize that, think of a 2D surface such as sphere $\mathcal{M}^2$ embedded in the 3D vector space $\mathcal{V}^3$. The tangent space $\mathcal{V}^2(x)$ at each point x on the surface is the 2D plane of vectors tangent to the surface at x [8].

At this point we part company with standard treatments of GR by presuming that the tangent vectors at each point x generate a Minkowski geometric algebra $\mathcal{G}_4(x) = \mathcal{G}(\mathcal{V}^4(x))$ called the *tangent algebra* at x. Consequently, the inner product of coordinate tangent vectors $g_\mu = g_\mu(x)$ generates the components $g_{\mu\nu} = g_{\mu\nu}(x)$ of the usual *metric tensor* in GR, that is,

$$g_\mu \cdot g_\nu = \tfrac{1}{2}(g_\mu g_\nu + g_\nu g_\mu) = g_{\mu\nu}\,. \tag{3}$$

Thus, all the rich structure of the *spacetime algebra* developed in [2] is inherited by the tangent algebras on the spacetime manifold $\mathcal{M}^4$. This defines a generalized *spacetime algebra* (STA) of multivector and spinor fields on the whole manifold.

Such fields are inherently geometrical, so they provide raw material for representing real physical entities as geometric objects. It remains to be seen if this material is sufficient for the purposes of physics. As demonstrated in the following sections, the STA of the spacetime manifold carries us a long way towards the ideal of inherently geometrical physics.

One great advantage of STA is that it enables coordinate-free formulation of multivector fields and field equations. To relate that to the coordinate-based formulation of standard tensor calculus, we return to our discussion of coordinates. The inverse mapping of (1) is a set of scalar-valued functions

$$x^\mu = x^\mu(x) \tag{4}$$

defined on the manifold $\mathcal{M}^4$. The gradient of these functions are vector fields

$$g^\mu = g^\mu(x) = \nabla x^\mu \,, \tag{5}$$

where $\nabla = \partial_x$ is the derivative with respect to the spacetime point x. It follows that

$$g_\mu \cdot g^\nu = \delta_\mu^\nu \qquad \text{or} \qquad g_\mu = g_{\mu\nu} g^\nu \,, \tag{6}$$

where the standard summation convention on repeated indices is used. Accordingly, we say that the *coordinate coframe* $\{g^\nu\}$ is "algebraically reciprocal" to the coordinate frame $\{g_\mu\}$.

This algebraic reciprocity facilitates decomposition of a vector field $a = a(x)$ into its *covariant components* $a_\mu = a \cdot g_\mu$ or its *contravariant components* $a^\mu = a \cdot g^\mu$; thus,

$$a = a^\mu g_\mu = a_\mu g^\mu \,, \tag{7}$$

Likewise, a bivector $F = F(x)$ has the expansion

$$F = \tfrac{1}{2} F^{\mu\nu} g_\mu \wedge g_\nu \,, \tag{8}$$

with its "scalar components" $F^{\mu\nu}$ given by

$$F^{\mu\nu} = g^\mu \cdot F \cdot g^\nu = g^\nu \cdot (g^\mu \cdot F) = (g^\nu \wedge g^\mu) \cdot F \,. \tag{9}$$

Similarly, the gradient operator can be defined in terms of partial derivatives by

$$\nabla = g^\mu \partial_\mu, \tag{10}$$

or vice-versa by

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = g_\mu \cdot \nabla \,. \tag{11}$$

The action of these operators on scalars is well defined, but differentiation of vectors on a curved manifold requires additional considerations, to which we now turn.

III. Coderivative and Curvature

On flat spacetime the vector derivative $\nabla = \partial_x$ is the only differential operator we need. For curved spacetime, we introduce the vector *coderivative* D as an intrinsic version of ∇. Operating on a scalar field $\phi = \phi(x)$, the two operators are equivalent:

$$D\phi = \nabla \phi \,. \tag{12}$$

Like the directional derivative $\partial_\mu = g_\mu \cdot \nabla$, the *directional coderivative* $D_\mu = g_\mu \cdot D$ is a "scalar differential operator" that maps vectors into vectors. Accordingly, we can write

$$D_\mu g_\nu = L^\alpha_{\mu\nu} g_\alpha \,, \tag{13}$$

which merely expresses the derivative as a linear combination of basis vectors. This defines the so-called *coefficients of connexion* $L^\alpha_{\mu\nu}$ for the frame $\{g_\nu\}$. By differentiating (5), we find the complementary equation

$$D_\mu g^\alpha = -L^\alpha_{\mu\nu} g^\nu \,, \tag{14}$$

When the coefficients of connexion are known functions, the coderivative of any multivector field is determined.

Thus, for any vector field $a = a^\nu g_\nu$ we have

$$D_\mu a = (D_\mu a^\nu) g_\nu + a^\nu (D_\mu g_\nu).$$

Then, since the a_ν are scalars, we get

$$D_\mu a = (\partial_\mu a^\alpha + a^\nu L^\alpha_{\mu\nu}) g_\alpha \,. \tag{15}$$

Note that the coefficient in parenthesis on the right is the standard expression for a "covariant derivative" in tensor calculus.

The derivative of any sum or product of multivector fields is easily computed by noting that D_μ is a scalar derivation, so it satisfies the usual Leibnitz and distributive rules of a derivative. In fact, those rules were used in computing the derivative in (15).

At last we are prepared to define the vector coderivative by

$$D = g^\mu D_\mu \,. \tag{16}$$

The "directional coderivative" with respect to any vector field $a = a(x)$ can now be defined by

$$a \cdot D = a^\mu D_\mu \,. \tag{17}$$

Both differential operators D and $a \cdot D$ are coordinate free. Though they have been defined with respect to coordinates, they can often be evaluated without reference to coordinates.

Since D is a vectorial differential operator, we can use the coordinate free algebraic operations of STA to manipulate it in precisely the same way we did with ∇ in [2]. Thus, the coderivative of any k-vector field $F = F(x)$ can be decomposed into a *codivergence* $D \cdot F$ and a *cocurl* $D \wedge F$, as expressed by

$$DF = D \cdot F + D \wedge F \,. \tag{18}$$

If F is an electromagnetic bivector field, we have the obvious generalization of Maxwell's equation to curved spacetime:

$$DF = J \,. \tag{19}$$

As done for the vector derivative in [2], this can be decomposed into the vector and trivector equations

$$D \cdot F = J\,, \tag{20}$$

$$D \wedge F = 0\,. \tag{21}$$

From the last equation it is tempting to conclude that $F = D \wedge A$, where A is a vector potential, but that depends on a property of D that remains to be proved.

To ascertain the geometric properties of the cocurl, we use (13) to obtain

$$D \wedge g^{\mu} = g^{\alpha} \wedge g^{\beta} L^{\mu}_{\alpha\beta}\,. \tag{22}$$

The quantity on the right side of this equation is called *torsion.* In the Riemannian geometry of GR torsion vanishes, so we leave the interesting consideration of nonzero torsion to another day. Considering the antisymmetry of the outer product on the right side of (22), we see that the torsion vanishes if and only if

$$L^{\mu}_{\alpha\beta} = L^{\mu}_{\beta\alpha}\,. \tag{23}$$

This can be related to the metric tensor by considering

$$D_{\mu} g_{\alpha\beta} = \partial_{\mu} g_{\alpha\beta} = (D_{\mu} g_{\alpha}) \cdot g_{\beta} + g_{\alpha} \cdot (D_{\mu} g_{\beta})\,,$$

whence

$$\partial_{\mu} g_{\alpha\beta} = g_{\alpha\nu} L^{\nu}_{\mu\beta} + g_{\beta\nu} L^{\nu}_{\mu\alpha}\,. \tag{24}$$

Combining three copies of this equation with permuted free indices, we solve for

$$L^{\mu}_{\alpha\beta} = \tfrac{1}{2} g^{\mu\nu} (\partial_{\alpha} g_{\beta\nu} + \partial_{\beta} g_{\alpha\nu} - \partial_{\nu} g_{\alpha\beta})\,. \tag{25}$$

This is the classical *Christoffel formula* for a *Riemannian connexion.*

To understand the geometric meaning of vanishing torsion, it is helpful to define a *torsion tensor*

$$T(a,b) \equiv a \cdot Db - b \cdot Da - [a,b]\,, \tag{26}$$

where $[\,a,b\,]$ is the *Lie bracket* of vector fields a and b defined by

$$[\,a,b\,] \equiv a \cdot \nabla b - b \cdot \nabla a\,. \tag{27}$$

For a coordinate frame the torsion tensor reduces to

$$T(g_{\mu}, g_{\nu}) = g_{\mu} \cdot D g_{\nu} - g_{\nu} \cdot D g_{\mu}\,, \tag{28}$$

because $[\,g_{\mu}, g_{\nu}\,] = (\partial_{\mu}\partial_{\nu} - \partial_{\nu}\partial_{\mu})x = 0$. From (28) we see that vanishing of the torsion tensor is equivalent to the symmetry condition (23) on the coefficients of connexion. Thus, from (26) we can conclude that vanishing torsion implies that

$$[\,a,b\,] = a\cdot Db - b\cdot Da \tag{29}$$

This relation between Lie bracket and coderivative plays an important role in our study of integrability in the Appendix.

To look at the significance of vanishing torsion from another angle, note that since g^μ is the gradient of a scalar coordinate function, the equation

$$D\wedge g^\mu = 0 \tag{30}$$

is equivalent to the following general property of the coderivative:

$$D\wedge D\phi = D\wedge\nabla\phi = 0\,, \tag{31}$$

where $\phi = \phi(x)$ is any scalar field. This is actually an *integrability condition* for scalar fields, as seen by considering

$$D\wedge D\phi = D\wedge g^\mu\partial_\mu\phi = g^\mu\wedge\nabla\partial_\mu\phi = g^\nu\wedge g^\mu\partial_\nu\partial_\mu\phi = 0\,, \tag{32}$$

whence

$$\partial_\nu\partial_\mu\phi = \partial_\mu\partial_\nu\phi\,. \tag{33}$$

This commutativity of partial derivatives is the classical condition for *integrability.*

To investigate the integrability of vector fields, we differentiate (14) to get

$$[D_\mu, D_\nu]g^\alpha = R^\alpha_{\mu\nu\beta}g^\beta\,, \tag{34}$$

where the operator commutator has the usual definition

$$[D_\mu, D_\nu] \equiv D_\mu D_\nu - D_\nu D_\mu\,, \tag{35}$$

and

$$R^\alpha_{\mu\nu\beta} = \partial_\mu L^\alpha_{\nu\beta} - \partial_\nu L^\alpha_{\mu\beta} + L^\alpha_{\nu\sigma}L^\sigma_{\mu\beta} - L^\alpha_{\mu\sigma}L^\sigma_{\nu\beta}\,, \tag{36}$$

is the usual tensor expression for the *Riemannian curvature* of the manifold. Vanishing of the curvature tensor is a necessary and sufficient condition for the manifold to be flat, in which case the coderivative reduces to the vector derivative of [2].

Using (30) we can recast the curvature equation (34) in terms of the coderivative:

$$D\wedge Dg^\alpha = \tfrac{1}{2}R^\alpha_{\mu\nu\beta}(g^\mu\wedge g^\nu)g^\beta\,. \tag{37}$$

This can be analyzed further in the following way:

$$D^2g^\alpha = (D\cdot D + D\wedge D)g^\alpha = D(D\cdot g^\alpha + D\wedge g^\alpha)\,. \tag{38}$$

Hence, using (30), we obtain

$$(D \wedge D)g^\alpha = D(D \cdot g^\alpha) - (D \cdot D)g^\alpha \,. \tag{39}$$

The right hand side of this equation has only a vector part; hence the trivector part of (37) vanishes to give us

$$D \wedge D \wedge g^\alpha = \tfrac{1}{2} R^\alpha_{\mu\nu\beta}(g^\mu \wedge g^\nu \wedge g^\beta) = 0 \,. \tag{40}$$

This is equivalent to the well known symmetry property of the curvature tensor:

$$R^\alpha_{\mu\nu\beta} + R^\alpha_{\beta\mu\nu} + R^\alpha_{\nu\beta\mu} = 0 \,. \tag{41}$$

However, its deep significance is that it implies

$$D \wedge D \wedge A = 0 \,. \tag{42}$$

for any k-vector field $A = A(x)$. This answers the question raised above about the existence of a vector potential for the electromagnetic field. It is a consequence of the condition (30) for vanishing torsion.

Equation (37) reduces to

$$D \wedge D g^\alpha = (D \wedge D) \cdot g^\alpha = R^\alpha_\beta g^\beta \,, \tag{43}$$

where

$$R^\alpha_\beta = R^\alpha_{\beta\mu\nu} g^{\mu\nu} \tag{44}$$

is the standard *Ricci tensor*. Comparing (43) with (39), we get the following provocative form for the Ricci tensor:

$$R(g^\alpha) \equiv R^\alpha_\beta g^\beta = D(D \cdot g^\alpha) - (D \cdot D)g^\alpha \,. \tag{45}$$

We return to this later.

IV. Gauge Principle of Equivalence

General Relativity is a theory of spacetime measurement. Any measurement of distance or direction in spacetime is a comparison of events with a standard, and for that purpose over an extended region a reference system is set up. In Special Relativity theory that purpose is met by *inertial reference frames* and encoded in the *Principle of Relativity*, which holds that the laws of physics (or measurements, if you will) are equivalent with respect to all inertial frames. A more precise formulation of this principle is that *the equations of physics are Lorentz invariant*, that is, invariant (or covariant) under Lorentz rotations.

In creating GR, Einstein struggled to find a suitable generalization of the Relativity Principle, and he formulated his conclusions in his *Principle of Equivalence*. However, the theoretical significance and physical meaning of the Equivalence Principle has remained intensely controversial to this day [9]. We are speaking here about the socalled “Strong Principle of Equivalence.” The “Weak Principle of Equivalence,” expressed by the equivalence of gravitational and inertial mass, is not problematic. The Strong Principle is vaguely described as equivalence of gravitational forces to accelerating systems. However, the tools of GC enable us to make a more general and precise formulation of the Principle that preserves the spirit if not the content of Einstein’s thinking.

Confusion about the Equivalence Principle can be traced to failure to make crucial distinctions between reference frames and coordinate systems [10]. At a single spacetime point a *reference frame* can be unambiguously defined as an orthonormal frame of vectors $\{\gamma_\mu\}$, which serve as a local standard for measurements of length and direction. This can be extended to a differentiable field of orthonormal vectors $\{\gamma_\mu = \gamma_\mu(x)\}$, which I call a *“fiducial frame”* or *fiducial frame field* to emphasize its role as a standard for measurement [11, 12]. It can be regarded as a generalization of “inertial frame” to curved spacetime, and visualized as a field of idealized rigid bodies at each point.

In contrast to the concept of a reference system as a fiducial frame field, a *coordinate system* is merely a means for labelling events, so it does not involve any spacetime geometry without additional assumptions. In Special Relativity, the terms “inertial coordinates” and “inertial frames” are often used interchangeably. Indeed, the standard choice of rectangular coordinates satisfies both coordinate and frame criteria for a reference system. However, this possibility is unique to flat spacetime. As can be proved with the mathematical apparatus developed below, on curved spacetime a fiducial frame cannot be identified with a coordinate frame, because it is a nonintegrable (or nonholonomic) system of vector fields. Vanishing of the curvature tensor is a necessary and sufficient condition for integrability of fiducial frames. Indeed, we shall see how to calculate the curvature tensor from inertial frames. The concept of “integrability” is so fundamental to differential geometry that an Appendix is devoted to further discussion of it.

With identification of fiducial frames as the appropriate generalization of inertial frames, the generalization of the Special Relativity Principle is now fairly obvious. We simply require equivalence of physics with respect to all fiducial frames. To mathematize this idea, we note that any given fiducial frame $\{\gamma_\mu\}$ is related to any other fiducial frame $\{\gamma'_\mu\}$ by a differentiable Lorentz rotation $\underline{R}$, which we know from [2] has the canonical form

$$\gamma'_\mu = \underline{R}(\gamma_\mu) = R\gamma_\mu\widetilde{R}\,, \tag{46}$$

where the underbar indicates that $\underline{R}$ is a linear operator, and $R = R(x)$ is a differentiable *rotor field* with the normalization

$$R\widetilde{R} = 1\,. \tag{47}$$

We can now formulate the

Gauge Principle of Equivalence (GPE): *The equations of physics are invariant under Lorentz rotations relating fiducial frames.*

In other words, with respect to fiducial frames all physical measurements are equivalent.

To justify its name we need to establish that the GPE is indeed a "gauge principle" and that it is a suitable generalization of Einstein's Principle of Equivalence. First, in contrast to the Special Relativity Principle that it generalizes, the GPE is indeed a gauge principle because in requires invariance under a *position dependent symmetry group*, the group of local Lorentz rotations (46). We show below that this is just what is needed to determine the form of gravitational interactions. Second, we note that the Lorentz rotation in (46) can be chosen to be a position dependent boost to a frame that is "accelerating" with respect to the initial frame, just as Einstein had contemplated in his version of the Equivalence Principle. Later we show how to generalize the local cancellation of apparent gravitational effects noted in his analysis.

We are now in position to conclude that Einstein's analysis was deficient in two respects: first, in overlooking the crucial distinction between reference frames and coordinate systems; second, in analysis that was too limited to ascertain the full gauge group. Still, we see here one more example of Einstein's astounding physical intuition in recognizing seeds of an important physical principle before it is given an adequate mathematical formulation.

The above analysis of reference frames and the Equivalence Principle suffices to motivate a reformulation of General Relativity with fiducial frames and the GPE at the foundation. First, some definitions and conventions are needed to streamline the formulation of basic formulas and theorems. The orthonormality of a fiducial frame $\{\gamma_\mu = \gamma_\mu(x)\}$ is conveniently expressed by

$$\gamma_\mu \cdot \gamma_\nu = \eta_\mu \delta_{\mu\nu} \,, \tag{48}$$

where $\eta_\mu = \gamma_\mu^2$ is the *signature indicator*. The reciprocal frame γ^μ is then simply given by

$$\gamma^\mu = \eta_\mu \gamma_\mu \,. \tag{49}$$

Of course, we assume that the fiducial frame is right-handed, so

$$i = \gamma_0\gamma_1\gamma_2\gamma_3 \tag{50}$$

where $i = i(x)$ is the *righthanded unit pseudoscalar* for the tangent space at x.

Any specified fiducial frame $\{\gamma_\mu\}$ is related to a specified coordinate frame $\{g_\mu\}$ by a differentiable linear transformation $\underline{h}$ called the *fiducial tensor*:

$$g_\mu = \underline{h}(\gamma_\mu) = h_\mu^\nu \gamma_\nu \,. \tag{51}$$

The matrix elements of the linear operator $\underline{h}$ are

$$h_\mu^\nu = \gamma^\nu \cdot \underline{h}(\gamma_\mu) = \gamma^\nu \cdot g_\mu = \bar{h}(\gamma^\nu) \cdot \gamma_\mu = g^\nu \cdot \gamma_\mu \,, \tag{52}$$

which shows that the adjoint of $\underline{h}$, denoted by $\bar{h}$, is

$$g^\nu = \bar{h}(\gamma^\nu) = h_\mu^\nu \gamma^\mu \,. \tag{53}$$

The fiducial tensor is related to the metric tensor by

$$g_{\mu\nu} = g_\mu \cdot g_\nu = \underline{h}(\gamma_\mu) \cdot \underline{h}(\gamma_\nu) = h_\mu^\alpha \eta_\alpha h_\nu^\alpha \,. \tag{54}$$

Alternatively, we can write

$$g_{\mu\nu} = \gamma_\mu \cdot \bar{h}\,\underline{h}(\gamma_\nu) = \gamma_\mu \cdot \underline{g}(\gamma_\nu)\,, \tag{55}$$

expressing the metric tensor as a symmetric linear transformation $\underline{g} = \bar{h}\,\underline{h} = \underline{h}\bar{h}$ on the fiducial frame. This shows that the metric tensor can be replaced by the fiducial tensor as a fundamental geometric object on spacetime. In the present formulation of GR, *the role of the fiducial tensor is to tie fiducial frames to the spacetime manifold by relating them to coordinate frames.*

We are now ready to investigate implications of the GPE. To achieve the gauge invariant equations required by the GPE, we need to define a gauge invariant derivative or, as we shall say, a *coderivative.* It turns out to be the same as the "coderivative" D defined in the last Section, but its physical significance is clarified, and its mathematical form is significantly improved. As before, it will be convenient to define the *directional coderivative* $D_\mu = g_\mu \cdot D$ first.

Since the fiducial frame $\{\gamma_\nu\}$ can only rotate under displacement, we know from [2] that its directional derivatives necessarily have the form

$$D_\mu \gamma_\nu = \omega_\mu \cdot \gamma_\nu \,, \tag{56}$$

where $\omega_\mu = \omega(g_\mu)$ is a bivector-valued "rotational velocity" for displacements in the g_μ direction. Let us call it the *fiducial connexion* for the frame $\{\gamma_\nu\}$. Generalizing this, we define action of the operator D_μ on an arbitrary multivector field $M = M(x)$ by

$$D_\mu M = \partial_\mu M + \omega_\mu \times M \,, \tag{57}$$

where the *commutator product* of A and B is defined by

$$A \times B = \tfrac{1}{2}(AB - BA)\,, \tag{58}$$

and it is assumed that

$$\partial_\mu \gamma_\nu = 0\,, \tag{59}$$

so the *the partial derivative* $\partial_\mu = g_\mu \cdot \nabla$ *operates only on scalar components of* M *relative to the fiducial basis.*

To manifest the relation of definition (57) to our previous definition, we apply it to coordinate frame vectors $g_\nu = h_\nu^\alpha \gamma_\alpha$ and compare with (13) to get

$$D_\mu g_\nu = L_{\mu\nu}^\alpha g_\alpha = (\partial_\mu h_\nu^\alpha)\gamma_\alpha + h_\nu^\alpha \omega_\mu \cdot \gamma_\alpha \,. \tag{60}$$

This equation establishes equivalence of the connexion for a coordinate frame to the connexion for a fiducial frame, but we have no more need for the coordinate connexion except to relate to literature that uses it.

Now the GPE requires invariance of D_μ under "change of gauge" to a different fiducial frame, as specified by equation (46). To ascertain necessary and sufficient conditions for gauge invariance, we differentiate (46) to get

$$\begin{aligned} D_\mu \gamma'_\nu &= (\partial_\mu R)\gamma_\nu \widetilde{R} + R\gamma_\nu \partial_\mu \widetilde{R} + R\omega_\mu \times \gamma_\nu \widetilde{R} \\ &= [(\partial_\mu R)\widetilde{R} + \tfrac{1}{2} R\omega_\mu \widetilde{R}] \times (R\gamma_\nu \widetilde{R})\,, \end{aligned} \tag{61}$$

where we have used $(\partial_\mu R)\widetilde{R} = -R\partial_\mu \widetilde{R}$, which follows from differentiating $R\widetilde{R} = 1$. It follows that

$$D_\mu \gamma'_\nu = \omega'_\mu \times \gamma'_\nu \tag{62}$$

provided that

$$\omega'_\mu = R\omega_\mu \widetilde{R} + 2(\partial_\mu R)\widetilde{R}\,. \tag{63}$$

In other words, the directional coderivative D_μ is invariant under a change of fiducial frame, as specified by the local Lorentz rotation (46), provided the change of fiducial connection is given by equation (63).

This completes our definition of the coderivative to satisfy the GPE. The definition refers to a coordinate frame only to exploit the well understood properties of partial derivatives. That inessential reference is eliminated in the following definition of the directional coderivative $a \cdot D$ with respect to an arbitrary vector field $a = a(x) = a^\mu g_\mu$:

$$a \cdot DM = a \cdot \nabla M + \omega(a) \times M\,, \tag{64}$$

where $\omega(a) = a^\mu \omega_\mu$ is the connexion for any chosen fiducial frame $\{\gamma_\mu\}$, and $a \cdot \nabla$ is the directional derivative of any scalar coefficients with respect to that frame.

We are now mathematically equipped for a deeper analysis of Einstein's *Strong Principle of Equivalence* (SPE). Without attempting to parse its many alternative formulations, we adopt the following formulation of the SPE: At any spacetime point x there exists an inertial (i.e. fiducial) reference frame in which the gravitational force vanishes. The nub of Einstein's idea is that the gravitational force can be cancelled by a suitable acceleration of the reference frame. Mathematically, this means that there exists a fiducial frame for which the connexion vanishes. In other words, the rotor field in the equation (46) for change of frame can be chosen to make $\omega'_\mu = 0$ in (63), so that

$$\omega_\mu = -2\widetilde{R}\partial_\mu R\,. \tag{65}$$

Read this as asserting that the gravitational force on the left is cancelled by acceleration of the reference frame on the right. A simple counting of degrees of freedom is sufficient to show that this condition can be satisfied at a single point. However, if it is satisfied in a finite neighborhood of that point, then, as established in the next Section, the curvature tensor vanishes and the manifold must be flat. Even so, the condition (65) can be imposed along any curve in spacetime. Indeed, in Section VII we impose it along timelike curves to get an equation of motion for a test body. Therefore, a more precise forumlation of the SPE is the following: *Along any spacetime curve there exists an inertial (i.e. fiducial) reference frame in which the gravitational force vanishes.*

In the present formulation of GR based on the GPE, the SPE is a theorem rather than a defining principle of the theory [9]. Evidently the SPE played a heuristic role in Einstein's thinking that helped him identify the gravitational force with a Riemannian connexion, but it is time to replace it with the deeper GPE. The necessity for this conclusion comes from recognizing that, to have physical content, any proposed relativity group must be a symmetry group of the theory. Thus the GPE expresses equivalence of observers (represented by fiducial frames) under local Lorentz rotations, and the gauge invariant coderivative is the theoretical consequence of this symmetry. Some such symmetry of observers seems to have been at the back of Einstein's mind, but the SPE is insufficient to designate a full symmetry group.

Now let us turn to more practical matters about how to perform calculations in GR. We have introduced the full *gauge invariant coderivative* by defining it in terms of directional derivatives with $D = g^\mu D_\mu$. However, that was merely for convenience, and it is worth noting that the operator D can be regarded as more fundamental than D_μ, as illustrated by the following important theorem:

$$\omega(\gamma_\mu) = \tfrac{1}{2}(\gamma_\alpha \wedge D \wedge \gamma^\alpha) \cdot \gamma_\mu - D \wedge \gamma_\mu \,. \tag{66}$$

This formula shows explicitly how to calculate a fiducial connexion from the cocurl of the frame vectors. We shall see later that this is a practical method for calculating the curvature tensor.

We can prove theorem (66) by solving the frame coderivative equations (56) for the connexion. First, we contract those equations to get

$$D \wedge \gamma_\nu = g^\mu \wedge [\omega(g_\mu) \cdot \gamma_\nu] = \gamma^\mu \wedge [\omega(\gamma_\mu) \cdot \gamma_\nu] \,,$$

and we note that

$$[\gamma^\mu \wedge \omega(\gamma_\mu)] \cdot \gamma_\nu = \gamma^\mu \wedge [\omega(\gamma_\mu) \cdot \gamma_\nu] + \omega(\gamma_\nu) \,.$$

Hence

$$\omega(\gamma_\nu) = -D \wedge \gamma_\nu + [\gamma^\mu \wedge \omega(\gamma_\mu)] \cdot \gamma_\nu \,. \tag{67}$$

To express the last term on the right hand side of this equation in terms of the cocurl, we return to (56) and observe that

$$(\omega_\mu \cdot \gamma_\nu)\gamma^\nu = (\omega_\mu \cdot \gamma_\nu) \wedge \gamma^\nu = 2\omega_\mu = (D_\mu \gamma_\nu)\gamma^\nu \,,$$

whence

$$2g^\mu \wedge \omega_\mu = 2\gamma^\mu \wedge \omega(\gamma_\mu) = (D \wedge \gamma_\mu) \wedge \gamma^\mu \,.$$

Inserting this into (67), we get the formula (66) as desired.

Finally, it may be noted that the integrability condition (30) for g^μ enables us to calculate the fiducial cocurl from the fiducial tensor. Writing $\gamma^\mu = \bar{h}^{-1}(g^\mu) = k^\mu_\nu g^\nu$, we find

$$D \wedge \gamma_\mu = \eta_\mu(\nabla k^\mu_\nu) \wedge g^\nu\,. \tag{68}$$

V. Gravitational Curvature and Field Equations

We have seen in (34) that the curvature tensor derives from the commutator of coderivatives. From the fiducial definition of the coderivative (57), we easily derive a more transparent and useful result: For any multivector field $M = M(x)$ we have

$$[D_\mu, D_\nu]M = \omega_{\mu\nu} \times M\,, \tag{69}$$

where

$$\omega_{\mu\nu} \equiv \partial_\mu\omega_\nu - \partial_\nu\omega_\mu + \omega_\mu \times \omega_\nu = R(g_\mu \wedge g_\nu) \tag{70}$$

is the curvature tensor evaluated on the bivector $g_\mu \wedge g_\nu$. It must be remembered that the partial derivatives here are given by

$$\partial_\mu\omega_\nu = \tfrac{1}{2}(\partial_\mu\omega^{\alpha\beta}_\nu)\gamma_\alpha \wedge \gamma_\beta\,, \tag{71}$$

where the scalar coefficients are $\omega^{\alpha\beta}_\nu = \gamma^\alpha \cdot \omega_\nu \cdot \gamma^\beta = \omega_\nu \cdot (\gamma^\beta \wedge \gamma^\alpha)$.

At this point it is worth noting that if the fiducial connection is derivable from a rotor field, as specified by the equation $\omega_\mu = -2\widetilde{R}\partial_\mu R$ from (65), then the curvature tensor (70) vanishes, as is easily proved by direct substitution. Thus, this is a sufficient condition for vanishing curvature. It is probably also a necessary condition for vanishing curvature, but I have not proved that.

The rest of this Section is devoted to summarizing and analyzing properties of the curvature tensor using the coordinate-free techniques of GC to demonstrate its advantages. For vector fields $a = a^\mu g_\mu$ and $b = b^\nu g_\nu$ the fundamental equation (69) can be put in the form

$$[a \cdot D, b \cdot D]M = R(a \wedge b) \times M\,, \tag{72}$$

provided $[\,a, b\,] = 0$. Vanishing of the Lie bracket is assumed here merely to avoid inessential complications.

Equation (72) shows that *curvature is a linear bivector-valued function of a bivector variable* that is defined in the tangent algebra at each spacetime point. Thus, for an arbitrary bivector field $B = B(x)$ we can write

$$R(B) \equiv \tfrac{1}{2}B \cdot (\partial_b \wedge \partial_a)\, R(a \wedge b) = \tfrac{1}{2}B^{\nu\mu}R(g_\mu \wedge g_\nu)\,, \tag{73}$$

where ∂_a is the usual vector derivative operating on the tangent space instead of the manifold, and $B^{\nu\mu} = B \cdot (g^\mu \wedge g^\nu)$. Note that this use of the vector derivative supplants decomposition into basis vectors and summation over indices, a technique that has been developed into a general method for basis-free formulation and manipulation of tensor algebra [12]. To that end, it is helpful to introduce the terminology *traction, contraction* and *protraction*, respectively, for the tensorial operations

$$\partial_a R(a \wedge b) = g^\mu R(g_\mu \wedge b) = \gamma^\mu R(\gamma_\mu \wedge b)\,, \tag{74}$$

$$\partial_a \cdot R(a \wedge b) = g^\mu \cdot R(g_\mu \wedge b) = \gamma^\mu \cdot R(\gamma_\mu \wedge b)\,,$$

$$\partial_a \wedge R(a \wedge b) = g^\mu \wedge R(g_\mu \wedge b) = \gamma^\mu \wedge R(\gamma_\mu \wedge b)\,.$$

that are employed below. These relations are easily proved by decomposing the vector derivative with respect to any basis and using the linearity of $R(a \wedge b)$ as in (73). Of course, the replacement of vector derivatives by basis vectors and sums over indices in (74) is necessary to relate the following coordinate-free relations to the component forms of standard tensor analysis.

To reformulate (72) as a condition on the vector coderivative D, note that for a vector field $c = c(x)$ the commutator product is equivalent to the inner product and (72) becomes

$$[a \cdot D, b \cdot D]c = R(a \wedge b) \cdot c\,. \tag{75}$$

To reformulate this as a condition on the vector coderivative, we simply eliminate the variables a and b by traction. Protraction of (75) gives

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

Another protraction together with

$$D \wedge D = \tfrac{1}{2}(\partial_b \wedge \partial_a)[\, a \cdot D, b \cdot D\,] \tag{76}$$

gives

$$D \wedge D \wedge c = [\, \partial_b \wedge \partial_a \wedge R(a \wedge b)\,] \cdot c + \partial_a \wedge R(a \wedge c)\,. \tag{77}$$

According to (42) the left side of this equation vanishes as a consequence of vanishing torsion, and, because the terms on the right have different functional dependence on the free variable c, they must vanish separately. Therefore

$$\partial_a \wedge R(a \wedge b) = 0\,. \tag{78}$$

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

The requirement (78) that the curvature tensor is *protractionless* has an especially important consequence. The identity

$$\partial_b \wedge [B \cdot (\partial_a \wedge R(a \wedge b))] = \partial_b \wedge \partial_a B \cdot R(a \wedge b) - B \cdot (\partial_b \wedge \partial_a) R(a \wedge b) \tag{79}$$

vanishes on the left side because of (78), and the right side then implies that

$$A \cdot R(B) = R(A) \cdot B\,. \tag{80}$$

Thus, the curvature is a *symmetric* bivector function. This symmetry can be used to recast (78) in the equivalent form

$$R\big((a \wedge b \wedge c) \cdot \partial_e\big) \cdot e = 0\,. \tag{81}$$

On expanding the inner product in its argument, it becomes

$$R(a \wedge b) \cdot c + R(c \wedge a) \cdot b + R(b \wedge c) \cdot a = 0\,, \tag{82}$$

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

As noted in (44), contraction of the curvature tensor defines the *Ricci tensor*

$$R(a) \equiv \partial_b \cdot R(b \wedge a)\,. \tag{83}$$

The Ricci identity (78) implies that we can write

$$\partial_b \cdot R(b \wedge a) = \partial_b R(b \wedge a)\,, \tag{84}$$

and also that the Ricci tensor is protractionless:

$$\partial_a \wedge R(a) = 0\,. \tag{85}$$

This implies the symmetry

$$a \cdot R(b) = R(a) \cdot b\,. \tag{86}$$

An alternative expression for the Ricci tensor is obtained by operating on (75) with (76) and establishing the identity

$$\tfrac{1}{2}(\partial_a \wedge \partial_b) \cdot [\, R(a \wedge b) \cdot c\,] = R(c)\,. \tag{87}$$

The result is, in agreement with (43),

$$D \wedge D\, a = (D \wedge D) \cdot a = R(a)\,. \tag{88}$$

This could be adopted as a definition of the Ricci tensor directly in terms of the coderivative without reference to the curvature tensor. That might lead to a more efficient formulation of the gravitational field equations introduced below.

Equation (88) 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 $M = a \wedge b$ in (72) and 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 product has the useful "distributive property" $A \times [B \times C] = A \times B + A \times C$, a fair amount of algebra is needed to reduce the right side of this equation. The result is

$$D \wedge D(a \wedge b) = R(a) \wedge b + a \wedge R(b) - 2R(a \wedge b)\,, \tag{89}$$

or equivalently

$$2R(a \wedge b) = (D \wedge Da) \wedge b + a \wedge (D \wedge Db) - D \wedge D(a \wedge b)\,. \tag{90}$$

This *differential identity* is the desired expression for the curvature tensor in terms of $D \wedge D$.

Contraction of the Ricci tensor defines the *scalar curvature*

$$R \equiv \partial_a R(a) = \partial_a \cdot R(a)\,. \tag{91}$$

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 that can be derived as follows. The commutators of directional coderivatives satisfy the Jacobi identity

$$[a \cdot D, [\,b \cdot D, c \cdot D\,]] + [\,b \cdot D, [\,c \cdot D, a \cdot D\,]] + [\,c \cdot D, [\,a \cdot D, b \cdot D\,]] = 0\,. \tag{92}$$

By operating with this on an arbitrary nonscalar multivector M and using (72), we can translate it into a condition on the curvature tensor that is known as the *Bianchi identity:*

$$a \cdot DR(b \wedge c) + b \cdot DR(c \wedge a) + c \cdot DR(a \wedge b) = 0\,. \tag{93}$$

Like the Ricci identity (81), this can be expressed more compactly as

$$\grave{R}[(a \wedge b \wedge c) \cdot \grave{D}\,] = 0\,, \tag{94}$$

where the accent serves to indicate that D differentiates the tensor R but not its tensor arguments. "Dotting" by 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

$$\grave{D} \wedge \grave{R}(a \wedge b) = 0\,. \tag{95}$$

This condition on the curvature tensor is the source of general conservation laws in General Relativity.

Contraction of (95) with ∂_a gives

$$\grave{R}(\grave{D} \wedge b) - D \wedge R(b) = 0\,. \tag{96}$$

A second contraction yields the differential identity

$$\grave{G}(\grave{D}) = \grave{R}(\grave{D}) - \tfrac{1}{2}DR = 0\,, \tag{97}$$

where

$$G(a) \equiv R(a) - \tfrac{1}{2}aR \tag{98}$$

is the *Einstein tensor.*

In General Relativity, for a given *energy-momentum tensor* $T(a)$, the spacetime geometry is determined by *Einstein's equation*

$$G(a) = \kappa T(a)\,, \tag{99}$$

where κ is a constant. The contracted Bianchi identity (97) implies the generalized *energy-momentum conservation* law

$$\grave{T}(\grave{D}) = 0\,. \tag{100}$$

As is well known, this is not a conservation law in the usual sense, because it is not a perfect divergence and so is not convertible to a surface integral by Gauss's theorem.

To solve Einstein's equation (99) for a given energy-momentum tensor, Einstein's tensor $G(a)$ must be expressed in a form that makes (99) a differential equation that describes the dynamics of spacetime geometry. A direct expression for $G(a)$ in terms of a fiducial connexion and its derivatives is very complicated and its structure is not very transparent. Let us consider an alternative approach. Using (88), we can put Einstein's equation (99) in the form.

$$D \wedge Da = \kappa(T(a) + \tfrac{1}{2}a\,\mathrm{Tr}\,T)\,, \tag{101}$$

where $\mathrm{Tr}\,T = \partial_a T(a)$.

As already noted in connection with equation (38), we can express this in alternative forms with the identity

$$D^2 a = D \wedge Da + D \cdot Da = D(D \cdot a) + D(D \wedge a)\,. \tag{102}$$

The last term vanishes if the vector field a is a gradient,

$$a = D\varphi = \nabla\varphi\,, \tag{103}$$

in which case, (101) can be put in the form

$$D \cdot Da - D(D \cdot a) = -\kappa(T(a) + \tfrac{1}{2}a\,\mathrm{Tr}\,T)\,. \tag{104}$$

This appears to be a simplification in the form of Einstein's equation, and it can be further simplified by adopting the "gauge condition" $D \cdot a = 0$. Indeed, in the linear approximation its left hand side reduces immediately to the usual d'Alembertian wave operator. This formulation of Einstein's equation was first derived in ref. [1], but it has never been studied further to see if its apparent simplicity leads to any practical advantages.

VI. Curvature Calculations

Equations (68), (66), (71), and (70) provide us with an efficient method for calculating curvature from the fiducial tensor in the following sequence of steps

$$h^\mu_\nu \quad \to \quad D \wedge \gamma_\mu \quad \to \quad \omega_\mu \quad \to \quad \omega_{\mu\nu} \,. \tag{105}$$

Conventional curvature calculations begin by specifying the metric tensor as a function of coordinates by writing the "line element"

$$dx^2 = dx \cdot dx = g_{\mu\nu} dx^\mu dx^\nu = h^\alpha_\mu \eta_\alpha h^\alpha_\nu dx^\mu dx^\nu \,. \tag{106}$$

Of course, we can take the same starting point for calculations with the fiducial tensor.

For orthogonal coordinates, we can choose the gauge so the γ_μ are eigenvalues of the fiducial tensor, and one has the simplifications

$$h^\mu_\nu = h_\nu \delta^\mu_\nu \tag{107}$$

$$\gamma_\mu = \eta_\mu \gamma^\mu = \eta_\mu h_\mu \nabla x^\mu = h^{-1}_\mu g_\mu \tag{108}$$

$$\omega_\mu = -h_\mu \nabla \wedge \gamma_\mu = \gamma_\mu \wedge \nabla h_\mu = \gamma_\mu \wedge \gamma^\alpha h^{-1}_\alpha \partial_\alpha h_\mu \tag{109}$$

$$\partial_\mu \omega_\nu = \eta_\alpha \partial_\mu (h^{-1}_\alpha \partial_\alpha h_\nu)] \gamma_\nu \wedge \gamma^\alpha \,. \tag{110}$$

As an important example, we calculate the curvature tensor from the Schwarzschild line element

$$\begin{aligned} dx^2 &= e^{2\Phi} dt^2 - e^{2\lambda} dr^2 - r^2 (d\theta^2 + \sin^2\theta d\phi^2) \\ &= h_t^2 dt^2 - h_r^2 dr^2 - h_\theta^2 d\theta^2 + h_\phi^2 d\phi^2, \end{aligned} \tag{111}$$

where $\Phi = \Phi(r,t)$ and $\lambda = \lambda(r,t)$ are scalar functions independent of θ and ϕ. Comparing (111) with (108), one can immediately write down

$$\begin{aligned} \gamma_t &= e^\Phi \nabla t & h_t &= e^\Phi \\ \gamma_r &= -e^\lambda \nabla r & h_r &= e^\lambda \\ \gamma_\theta &= -r \nabla \theta & h_\theta &= r \\ \gamma_\phi &= -r \sin\theta \nabla \phi & h_\phi &= r \sin\theta \end{aligned} \tag{112}$$

Use of the same symbols t, r, θ, ϕ for indices, coordinates and independent variables should not cause confusion, as the distinction is clear from the context.

Using (112) in (109), one easily obtains

$$\begin{aligned} \omega_t &= -\gamma_t \wedge \gamma_r \Phi_r e^{\Phi-\lambda} \\ \omega_r &= \gamma_r \wedge \gamma_t \lambda_t e^{\lambda-\Phi} \\ \omega_\theta &= -\gamma_\theta \wedge \gamma_r e^{-\lambda} \\ \omega_\phi &= -\gamma_\phi \wedge (\gamma_r e^{-\lambda} \sin\theta + \gamma_\theta \cos\theta) \,, \end{aligned} \tag{113}$$

where $\Phi_r \equiv \partial_r \Phi$ and $\lambda_r \equiv \partial_r \lambda$.

The simple computation of ω_μ in (113) may be compared with the corresponding computation from the Friedman metric by Misner, Thorne and Wheeler [13]. They advocate skillful guessing with differential forms, while, in the sequence (105) we merely apply equation (66), which has no counterpart in their method. Indeed, application of (66) is easier as well as more straightforward than guessing, because the effort needed to check each guess has been expended once and for all in the derivation of (66).

In should be noted that the "wedges" in (113) are actually unnecessary because the γ_μ are orthogonal. This greatly simplifies evaluation of the commutators $\omega_\mu \times \omega_\nu$, and one easily determines that the nonvanishing commutators have the values

$$\begin{aligned}
\omega_t \times \omega_\theta &= \gamma_t \wedge \gamma_\theta \Phi_r e^{\Phi - 2\lambda} \\
\omega_t \times \omega_\phi &= \gamma_t \wedge \gamma_\phi \Phi_r e^{\Phi - 2\lambda} \sin\theta \\
\omega_r \times \omega_\theta &= \gamma_\theta \wedge \gamma_t \lambda_t e^{-\Phi} \\
\omega_r \times \omega_\phi &= \gamma_\phi \wedge \gamma_t \lambda_t e^{-\Phi} \sin\theta \\
\omega_\theta \times \omega_\phi &= \gamma_\phi \wedge (\gamma_r e^{-\lambda} \cos\theta - \gamma_\theta e^{-2\lambda} \sin\theta)\,.
\end{aligned} \tag{114}$$

Using (110) to compute the relevant fiducial derivatives from (113), one finds that the nonvanishing terms are

$$\begin{aligned}
\partial_r \omega_t &= -\gamma_t \wedge \gamma_r (\Phi_{rr} + \Phi_r^2 - \Phi_r \lambda_r) e^{\Phi - \lambda} \\
\partial_t \omega_r &= \gamma_r \wedge \gamma_t (\lambda_{tt} + \lambda_t^2 - \Phi_t \lambda_t) e^{\lambda - \Phi} \\
\partial_t \omega_\theta &= \gamma_\theta \wedge \gamma_r \lambda_t e^{-\lambda} \\
\partial_t \omega_\phi &= \gamma_\phi \wedge \gamma_r \lambda_t e^{-\lambda} \sin\theta \\
\partial_r \omega_\theta &= \gamma_\phi \wedge \gamma_r \lambda_r e^{-\lambda} \\
\partial_r \omega_\phi &= \gamma_\phi \wedge \gamma_r \lambda_r e^{-\lambda} \sin\theta \\
\partial_\theta \omega_\phi &= (\gamma_r e^{-\lambda} \cos\theta - \gamma_\theta \sin\theta) \wedge \gamma_\phi\,.
\end{aligned} \tag{115}$$

Finally, one obtains the curvature bivectors by inserting (114) and (115) into (70). I display the result to show both the coordinate components and the fiducial components $\hat{\omega}_{\mu\nu} = R(\gamma_\mu \wedge \gamma_\nu)$:

$$\begin{aligned}
\omega_{tr} &= e^{\Phi+\lambda}\hat{\omega}_{tr} \\
&= e^{\Phi+\lambda}\gamma_r \wedge \gamma_t[(\lambda_{tt} + \lambda_t^2 - \Phi_t\lambda_t)e^{-2\Phi} - (\Phi_{rr} + \Phi_r^2 - \Phi_r\lambda_r)e^{-2\lambda}] \\
\omega_{t\theta} &= e^{\Phi} r\hat{\omega}_{t\theta} \\
&= e^{\Phi} r\gamma_\theta \wedge [\gamma_r\lambda_t r^{-1}e^{-\Phi-\lambda} - \gamma_t\Phi_r r^{-1}e^{-2\lambda}] \\
\omega_{t\phi} &= e^{\Phi} r\sin\theta\hat{\omega}_{t\theta} \\
&= e^{\Phi} r\sin\theta\gamma_\phi \wedge [\gamma_r\lambda_t r^{-1}e^{-\Phi-\lambda} - \gamma_t\Phi_r r^{-1}e^{-2\lambda}] \\
\omega_{r\theta} &= e^{\lambda} r\hat{\omega}_{r\theta} \\
&= e^{\lambda} r\gamma_\theta \wedge [\gamma_r\lambda_r r^{-1}e^{-2\lambda} + \gamma_t\lambda_t r^{-1}e^{-\Phi-\lambda}] \\
\omega_{r\phi} &= e^{\lambda} r\sin\theta\hat{\omega}_{r\phi} \\
&= e^{\lambda} r\sin\theta\gamma_\phi \wedge [\gamma_r\lambda_r r^{-1}e^{-2\lambda} + \gamma_t\lambda_t r^{-1}e^{-\Phi-\lambda}] \\
\omega_{\theta\phi} &= r^2\sin\theta\hat{\omega}_{\theta\phi} \\
&= r^2\sin\theta\gamma_\theta \wedge \gamma_\phi r^{-2}(e^{-2\lambda} - 1)\,.
\end{aligned} \tag{116}$$

Of course, if we had so desired, we could have computed any one of these bivectors independently. Also, many of the coefficients in (116) are equivalent because of the curvature symmetry property $A \cdot R(B) = B \cdot R(A)$. This redundancy provides a check on the computations.

The orthogonality of the γ_μ makes it especially easy to contract (116) to get the Ricci tensor

$$\begin{aligned}
R(\gamma_t) = -2\gamma_r\lambda_t r^{-1}e^{-\Phi-\lambda} + \gamma_t[(\lambda_{tt} + \lambda_t^2 - \Phi_t\lambda_t)e^{-2\Phi} \\
(\Phi_{rr} + \Phi_r^2 - \Phi_r\lambda_r + 2\Phi_r r^{-1})e^{-2\lambda}]\,,
\end{aligned} \tag{117}$$

$$\begin{aligned}
R(\gamma_r) = -2\gamma_t\lambda_t r^{-1}e^{-\Phi-\lambda} + \gamma_r[-(\lambda_{tt} + \lambda_t^2 - \Phi_t\lambda_t)e^{-2\Phi} \\
+ (\Phi_{rr} + \Phi_r^2 - \Phi_r\lambda_r - 2\lambda_r r^{-1})e^{-2\lambda}]\,,
\end{aligned} \tag{118}$$

$$R(\gamma_\theta) = \gamma_\theta[(\lambda_r - \Phi_r)r^{-1}e^{-2\lambda} + (1 - e^{-2\lambda})r^{-2}] = \gamma_\phi\gamma_\theta R(\gamma_\phi)\,. \tag{119}$$

Setting $R(\gamma_\mu) = 0$ to find the "free space" gravitational field according to Einstein's theory, we see immediately that (117) and (118) imply that $\Phi_t = \lambda_t = 0$ and $\Phi = -\lambda$, so (119) reduces to the equation

$$2r\lambda_r + e^{2\lambda} - 1 = 0\,. \tag{120}$$

This integrates to the famous Schwarzschild solution

$$e^{-2\lambda} = 1 - \frac{2\kappa}{r} = e^{2\Phi}\,, \tag{121}$$

where κ is a constant.

Finally, we can substitute (121) into (113) to get ω_μ as an explicit function of the coordinates. The result can be used to describe the motion of a test body, but we take a more general approach in the next Section.

VII. Gravitational Motion and Precession

The spinor equations of motion for classical particles and rigid bodies set forth in [2] are now easily generalized to include gravitational interactions. This gives us a general method for evaluating gravitational effects on the motion and precession of a spacecraft or satellite, and thus a means for testing gravitational theory.

We begin with the timelike worldline $x = x(\tau)$ of a material particle with velocity $v = v(\tau) = dx/d\tau \equiv \dot{x}$, where, as usual, $d\tau = |\,dx\,| = |\,(dx)^2\,|^{\frac{1}{2}}$, so $v^2 = 1$. We attach to this curve a (comoving orthonormal frame) or *mobile* $\{e_\mu = e_\mu(x(\tau)) = e_\mu(\tau); \mu = 0,1,2,3\}$. The mobile is tied to the velocity by requiring $v = e_0$. Rotation of the mobile with respect to a given fiducial frame $\{\gamma_\mu\}$ is described by

$$e_\mu = R\,\gamma_\mu \widetilde{R}\,, \tag{122}$$

where $R = R(x(\tau))$ is a unimodular rotor and $\{\gamma_\mu\}$ is any convenient fiducial frame. As noted in [2], the spinor can be used to model the motion of a small rigid body or a particle with intrinsic spin. In GR it is especially useful for modeling gravitational effects on gyroscopic precession.

In accordance with (64), the coderivative of the mobile is

$$v\cdot De_\mu = \dot{e}_\mu + \omega(v)\cdot e_\mu\,, \tag{123}$$

where $\{\mu = 0,1,2,3\}$, $\omega(v)$ is the fiducial connection for the frame $\{\gamma_\mu\}$, and $\dot{e}_\mu = v\cdot\nabla e_\mu$. Note that (123) is equivalent to the spinor equation

$$v\cdot DR = \big(\frac{d}{d\tau} + \tfrac{1}{2}\omega(v)\big)R\,, \tag{124}$$

The coderivative (123) includes a gauge invariant description of gravitational forces on the mobile. As explained in [2], effects of any nongravitational forces can be incorporated by writing

$$v\cdot De_\mu = \dot{e}_\mu + \omega(v)\cdot e_\mu = \Omega\cdot e_\mu\,, \tag{125}$$

where $\Omega = \Omega(x)$ is a bivector field acting on the mobile; for example, $\Omega = (e/m)F$ for an electron with mass m and charge e in a field $F = F(x)$. The four equations (125) include the equation of motion

$$\frac{dv}{d\tau} = (\Omega - \omega(v))\cdot v \tag{126}$$

for the particle, and are equivalent to the single rotor equation

$$\frac{dR}{d\tau} = \tfrac{1}{2}(\Omega - \omega(v))R\,. \tag{127}$$

For $\Omega = 0$ the particle equation becomes the equation for a geodesic, and the rotor equation describes parallel transfer of the mobile along the geodesic.

It is noteworthy that gravitational forces are described by the bivector $\omega(v)$ while nongravitational forces are described by specifying Ω. This suggests a *general superposition principle for interactions*, namely, that independent forces make independent contributions to the rotational velocity of a mobile. It generalizes the *principle of superposition of forces* familiar from classical mechanics. It considerably simplifies the analysis of complex problems by allowing us to determine the rotational velocities due to different interactions and add the results. This superposition principle may be derivable from a deeper gauge theory of particle interactions, but that remains to be seen.

To facilitate physical interpretation and comparison with the literature and experimental results, we need to express our results in terms of *relative variables* with respect to the specified fiducial reference frame. This is best done by the general method of *spacetime split* laid out in [2]. We recall some of its essential features for present application.

At every spacetime point x the fiducial timelike vector $\gamma_0 = \gamma_0(x)$ determines a *split* of the tangent algebra into space and time components. The particle velocity is split into space and time components by

$$v\gamma_0 = \beta(1 + \mathbf{v})\,, \tag{128}$$

where

$$\beta = v \cdot \gamma_0 = \frac{dt}{d\tau} = \left(1 - \mathbf{v}^2\right)^{-\frac{1}{2}} \tag{129}$$

is the "time dilation" factor, and

$$\mathbf{v} = \frac{v \wedge \gamma_0}{v \cdot \gamma_0} = \frac{d\mathbf{x}}{dt} = \frac{dx^k}{dt}\boldsymbol{\sigma}_k \tag{130}$$

is the *relative velocity* in the fiducial reference system, with the timelike bivectors

$$\boldsymbol{\sigma}_k \equiv \gamma_k \wedge \gamma_0 = \gamma_k\gamma_0 = \gamma_0\gamma^k \tag{131}$$

(for $k = 1, 2, 3$) composing a basis for relative vectors.

Recall the simple split of the electromagnetic bivector F into electric and magnetic parts:

$$F = \mathbf{E} + i\mathbf{B}\,. \tag{132}$$

Similarly, the "total rotational velocity" in equations (126) and (127) can be written in the split form:

$$\Omega - \omega(v) = \mathbf{e} + i\mathbf{b}\,, \tag{133}$$

which implicitly defines relative vectors $\mathbf{e}$ and $\mathbf{b}$. With this definition, a split of the particle equation of motion (126) yields a relative equation of motion in the familiar "Lorentz form"

$$\frac{d(\beta\mathbf{v})}{dt} = \mathbf{e} + \mathbf{v} \times \mathbf{b}\,. \tag{134}$$

This makes it easy to compare with the Newton gravitational force, for which $\mathbf{b} = 0$ and $\mathbf{e}$ is given by the inverse square law. An effective perturbation method for calculating deviations from classical Kepler motion is developed in my mechanics book [14]. However, as shown in [2], the proper equation of motion (126) is often easier to solve than the relative equation (134).

To separate the description of rigid body precession from translational motion, we follow [2] in making a split of the rotor R into the rotor product

$$R = LU\,, \tag{135}$$

where L is defined by

$$L\gamma_0\widetilde{L} = L^2\gamma_0 = v\,, \tag{136}$$

so U satisfies

$$U\gamma_0\widetilde{U} = \gamma_0\,. \tag{137}$$

These equations specify the unique factorization of the mobile Lorentz rotation (122) into a rotation in the instantaneous fiducial rest frame specified by U followed by a boost specified by L.

In the the instantaneous fiducial rest frame of γ_0, the axes of the rigid body are represented by the relative vectors

$$\mathbf{e}_k = \widetilde{L}(e_k e_0)L = U\boldsymbol{\sigma}_k\widetilde{U}\,. \tag{138}$$

The rotational motion can be described by the three vector equations

$$\frac{d\mathbf{e}_k}{dt} = \mathbf{\Omega}\times\mathbf{e}_k \tag{139}$$

or, better, by the single rotor equation

$$\frac{dU}{dt} = -\tfrac{1}{2}i\mathbf{\Omega}U\,. \tag{140}$$

We are most interested here in the rotational velocity $\mathbf{\Omega}$. The algebraic problem of solving for $\mathbf{\Omega}$ in terms of $\mathbf{e}$ and $\mathbf{b}$ is solved in [2], with the result [15]

$$\mathbf{\Omega} = -\beta^{-1}\mathbf{b} + (1+\beta)^{-1}\mathbf{v}\times\mathbf{e}\,. \tag{141}$$

This is a combination of relativistic Larmor and Thomas precessions.

Considering gravitational effects alone and using (128), the rotational velocity (133) takes the form

$$-\omega(v) = \mathbf{e} + i\mathbf{b} = -v\cdot\gamma^\mu\omega(\gamma_\mu) = -\beta[\omega(\gamma_0) + \mathbf{v}\cdot\boldsymbol{\sigma}_k\omega(\gamma_k)] \tag{142}$$

The problem remains to evaluate $\omega(\gamma_\mu)$ in terms of a given fiducial (or metric) tensor. The solution is given by the general formula (66), which reduces the problem to calculating $D\wedge\gamma_\mu$.

To get a sensitive test of GR, we consider a fiducial tensor that is somewhat more general than the Schwarzschild one. Specifically, we assume that the fiducial frame is given by

$$\begin{aligned}\gamma_0 &= e^{\Phi}\nabla x^0 = \gamma^0 \\ \gamma_k &= -h_k \nabla x^0 + e^{\lambda}\nabla x^k = -\gamma^k\,,\end{aligned} \tag{143}$$

where Φ, λ and the h_k are scalar functions describing the gravitational field. From (143) we can read off the components of the fiducial tensor:

$$h_0^0 = e^{\Phi}, \qquad h_k^0 = 0, \qquad h_0^k = h_k, \qquad h_j^k = e^{\lambda}\delta_j^k\,, \tag{144}$$

for $j, k = 1, 2, 3$.

The coordinate frame $g_\mu = h_\mu^\nu \gamma_\nu$ is therefore given by

$$g_0 = e^{\Phi}\gamma^0 + h_k\gamma_k, \qquad g_k = e^{\lambda}\gamma_k\,. \tag{145}$$

For the purpose of comparison with the literature, we evaluate the metric tensor $g_{\mu\nu} = g_\mu \cdot g_\nu$, with the results

$$\begin{aligned}g_{00} &= e^{2\Phi} - \sum_k h_k^2 \\ g_{0k} &= -e^{\lambda}h_k \\ g_{ij} &= -e^{2\lambda}\delta_{ij}\,,\end{aligned} \tag{146}$$

We can solve equations (145) to get the reciprocal frame $g^\mu = \nabla x^\mu$ by the general method in the Appendix. However, this case is so simple that we can immediately write down the result

$$\begin{aligned}g^0 &= \nabla x^0 = e^{-\Phi}\gamma^0 \\ g_k &= \nabla x^k = e^{-\lambda}(\gamma^k - e^{-\Phi}h_k\gamma_0)\,.\end{aligned} \tag{147}$$

Now, taking the curl of (143) and using (147), we get

$$\begin{aligned}D\wedge\gamma_0 &= (\nabla\Phi)\wedge\gamma_0 \\ D\wedge\gamma_k &= \gamma^k\wedge\nabla\lambda + e^{-\Phi}\gamma_0\wedge(\nabla h_k - h_k\nabla\lambda))\,,\end{aligned} \tag{148}$$

from which we obtain the trivector

$$T \equiv \gamma^\mu \wedge D \wedge \gamma_\mu = e^{-\Phi}\gamma^k\wedge\gamma_0\wedge(\nabla h_k - h_k\nabla\lambda))\,. \tag{149}$$

We could substitute (148) and (149) immediately into (66) to get expressions for the $\omega(\gamma_\mu)$. However, we want our results in terms of relative variables, so let us first introduce the necessary notation for that end.

We represent the time coordinate by $t = x^0$ and introduce the notations

$$\partial_t = \gamma_0 \cdot \nabla, \qquad \partial_k = \gamma_k \cdot \nabla = \boldsymbol{\sigma}_k \cdot \nabla,$$

$$\nabla = \gamma_0 \wedge \nabla = \boldsymbol{\sigma}_k \partial_k, \qquad \mathbf{h} = h_k \boldsymbol{\sigma}_k \,. \tag{150}$$

Now (148) assumes the form

$$\begin{aligned} D \wedge \gamma_0 &= -\nabla\Phi, \\ D \wedge \gamma_k &= -\boldsymbol{\sigma}_k \wedge \nabla\lambda - \boldsymbol{\sigma}_k \partial_t \lambda + e^{-\Phi}(\nabla h_k - h_k \nabla\lambda) \,. \end{aligned} \tag{151}$$

And from (149) we obtain

$$\begin{aligned} T\cdot\gamma_0 &= -e^{-\Phi}[\nabla \wedge \mathbf{h} + \mathbf{h} \wedge \nabla\lambda], \\ T\cdot\gamma_k &= -e^{-\Phi}[\partial_k \mathbf{h} - \mathbf{h}\partial_k \lambda - \nabla h_k + h_k \nabla\lambda] \,. \end{aligned} \tag{152}$$

Inserting (151) and (152) into (66), we obtain

$$\begin{aligned} \omega(\gamma_0) &= \nabla\Phi - \tfrac{1}{2} e^{-\Phi}[\nabla \wedge \mathbf{h} + \mathbf{h} \wedge \nabla\lambda], \\ \omega(\gamma_k) &= \boldsymbol{\sigma}_k \wedge \nabla\lambda + \boldsymbol{\sigma}_k \partial_t \lambda - \tfrac{1}{2} e^{-\Phi}[\partial_k \mathbf{h} - \mathbf{h}\partial_k \lambda + \nabla h_k - h_k \nabla\lambda] \,. \end{aligned} \tag{153}$$

For a weak static gravitational field, the case of greatest experimental interest, these equations assume the approximate form

$$\omega(\gamma_0) = \nabla\Phi - \tfrac{1}{2}\nabla \wedge \mathbf{h}, \tag{154}$$

$$\omega(\gamma_k) = \boldsymbol{\sigma}_k \wedge \nabla\lambda - \tfrac{1}{2}\nabla h_k \,. \tag{155}$$

Inserting (155) into (142), we get

$$\omega(v) = \beta[(\nabla(\Phi - \tfrac{1}{2}\mathbf{h}\cdot\mathbf{v}) + i(\mathbf{v} \times \nabla\lambda - \tfrac{1}{2}\nabla \times \mathbf{h})] \,, \tag{156}$$

from which one can read off the desired values of **e** and **b**. Substituting these results into (134), we get the relative equation of motion for a particle in a gravitational field:

$$\frac{d\mathbf{v}}{dt} = \beta[(\nabla(\Phi - \tfrac{1}{2}\mathbf{h}\cdot\mathbf{v}) + \mathbf{v} \times (\mathbf{v} \times \nabla\lambda - \tfrac{1}{2}\nabla \times \mathbf{h})] \,. \tag{157}$$

Similarly, the same substitution into (141) yields the standard result for small velocities [16]:

$$\Omega = \mathbf{v} \times \nabla(\lambda - \tfrac{1}{2}\Phi) - \tfrac{1}{2}\nabla \times \mathbf{h} \,. \tag{158}$$

Physical applications of this result are discussed in many textbooks on GR.

VIII. Dirac Equation with Gravitational Interaction

Recall from [2] that a real Dirac spinor field $\psi = \psi(x)$ determines an orthonormal frame of vector fields $e_\mu = e_\mu(x)$ defined by

$$\psi \gamma_\mu \widetilde{\psi} = \rho e_\mu \,, \tag{159}$$

where scalar $\rho = \rho(x)$ is interpreted as electron probability density, and $\psi\gamma_0\widetilde{\psi} = \rho e_0$ is the Dirac current. We can adopt this relation without change by interpreting $\{\gamma_\mu\}$ as a fiducial frame and writing

$$e_\mu = R\gamma_\mu\widetilde{R}\,, \tag{160}$$

where $R = R(x)$ is a rotor field. This equation has exactly the same form as the equation (46) for a change of fiducial frame. Therefore, the Dirac wave function determines a unique, physically significant fiducial frame $\{e_\mu\}$ on spacetime. Accordingly, its gauge invariant directional coderivative is given by

$$D_\nu e_\mu = \partial_\nu e_\mu + \omega_\nu \cdot e_\mu\,, \tag{161}$$

where ω_ν is the fiducial connexion for the frame $\{\gamma_\mu\}$. This is consistent with defining the *coderivative of the Dirac spinor* by

$$D_\nu\psi = (\partial_\nu + \tfrac{1}{2}\omega_\nu)\psi\,. \tag{162}$$

The spinor coderivative (162) is form invariant under the *spinor gauge transformation*

$$\psi \qquad \rightarrow \qquad \psi' = \Lambda\psi\,, \tag{163}$$

where $\Lambda = \Lambda(x)$ is a rotor field. This induces a transformation of (162) to

$$D_\nu\psi' = (\partial_\nu + \tfrac{1}{2}\omega'_\nu)\psi'\,, \tag{164}$$

where

$$\omega'_\nu = \Lambda\omega_\nu\widetilde{\Lambda} - 2(\partial_\nu\Lambda)\widetilde{\Lambda}\,. \tag{165}$$

We could have used this *spinor gauge transformation* to define the spinor coderivative. But note that it is not (explicitly, at least) related to the gauge equivalence of fiducial frames, so it raises new issues of physical interpretation. It is an active transformation that changes the fields on spacetime, rather than a passive transformation that changes the reference system but leaves fields unchanged. We shall return to this issue in the sequel to this paper.

The generalization of the real Dirac equation in [2] to include gravitational interaction is obtained simply by replacing the partial derivative ∂_μ by the coderivative D_μ. Thus, we obtain

$$g^\mu D_\mu\psi\gamma_2\gamma_1\hbar = g^\mu(\partial_\mu + \tfrac{1}{2}\omega_\mu)\psi\gamma_2\gamma_1\hbar = eA\psi + m\psi\gamma_0\,. \tag{166}$$

This is equivalent to the standard matrix form of the Dirac equation with gravitational interaction, but it is obviously much simpler in formulation and application. This is not the time and place for solving the *real gravitational Dirac equation* (166). However, comparison of the spinor coderivative (162) with the rotor coderivative (124) tells us immediately that *gravitational effects on electron motion, including spin precession, are exactly the same as on classical rigid body motion.*

With the spinor coderivative in hand, the rest of Dirac theory in [2] is easily adapted to gravitational interactions [3, 4].

IX. Vector Manifolds

The spacetime manifold $\mathcal{M}^4 = \{x\}$ was introduced as a *vector manifold* in Section II, and a *coordinate frame* $\{g_\mu = g_\mu(x)\}$ was generated from partial derivatives of a parametrized point in the manifold, as expressed by

$$g_\mu = \partial_\mu x\,. \tag{167}$$

At each spacetime point x the coordinate frame provides a basis for the tangent space $\mathcal{V}^4(x)$ and generates the tangent algebra $\mathcal{G}_4(x) = \mathcal{G}(\mathcal{V}^4(x))$.

The reader may have noticed that the role of $\mathcal{M}^4$ itself in subsequent developments is hardly more than a shadow. All the geometry and physics — the vector, multivector and spinor fields, the connexion and the curvature — occur in the tangent algebra. It could be argued that even the spacetime points $\{x\}$ are superfluous, as coordinates are sufficient to index points of the manifold. This argument is taken to the extreme in most recent mathematical works on differential geometry, where the x is eliminated and (167) is replaced by

$$g_\mu = \partial_\mu\,. \tag{168}$$

In other words, vectors in a coordinate frame are identified with coordinate partial derivatives; consequently, all vectors $\{a = a^\mu g_\mu\}$ in the tangent space are identified with the directional derivatives $\{a^\mu\partial_\mu\}$.

The purported problem with (167) is that it is deficient in mathematical rigor, because the partial derivative is defined as the limit of a difference quotient

$$\partial_\mu x = \frac{\triangle x}{\triangle x^\mu}\,, \tag{169}$$

and the difference vector $\triangle x$ requires subtraction of one point from another, which is not well defined unless they are vectors in a vector space of higher dimension. In other words, it is argued that the equation (169) presumes embedding of $\mathcal{M}^4$ in a vector space of higher dimension, whereas GR is concerned with intrinsic properties of manifolds irrespective of any embedding in a higher dimensional space. The definition (168) of tangent vectors as differential operators finesses this issue with a "don't ask, don't tell" approach that doesn't specify what is to be differentiated. Nevertheless, it has been argued that (167) has great heuristic value [17].

It has been almost universally overlooked in the mathematics and physics literature that the identification (168) of tangent vectors with differential operators precludes assigning them the algebraic properties of vectors in geometric algebra as done in this paper. Such conflation of vectors with differential operators has enormous drawbacks. It is sufficient to note that if tangent vectors are not allowed to generate a geometric algebra in the first place, then the algebra must be artifically imposed on the manifold later on, because it is absolutely essential for spinors and quantum mechanics. Indeed, standard practice [18, 7, 19] is to attach the Dirac algebra to the spacetime manifold as an afterthought, and recently the elaborate formalism of fibre bundles has been employed for that [20].

To avoid all that unnecessary gymnastics, it is necessary to return mathematical respectability to equation (167) — that requires reconsidering the concept of a differentiable manifold.

The standard definition of a differentiable manifold employs coordinates to impose differentiable structures on a set [20]. Alternatively, the definition of a vector manifold has been expressly designed to incorporate differentiability directly into the structure of the set [12, 21]. This entails regarding the vectorial difference quotient (169) as a well-defined quantity with the well-defined limit (167). Contrary to common belief, it does not *require* any assumptions about embedding the spacetime manifold in a (flat) vector space of higher dimension. Indeed, no mention of an embedding space appears in this paper. However, if one insists that an embedding vector space must be assumed to make vectorial operations like (169) meaningful, there is still no loss of generality in representing the spacetime manifold as a vector manifold, because it has been proved that every Riemannian manifold can be embedded in a flat manifold of sufficiently high dimension [22]. Indeed, the theory of vector manifolds may be the ideal venue for investigating embedding theorems, because it offers a powerful new method for differential geometry that efficiently coordinates characterization and analysis of the intrinsic and extrinsic properties of a manifold without presumptions about embedding [12]. As that method is based on the same concept of vector manifold employed here, it is an attractive alternative to the method in this paper, and the two methods can be regarded as complementary. Sobczyk has taken the first steps in the use of GC for an embedding approach to spacetime manifolds [23]. The main interest of physicists in studying extrinsic geometry of spacetime manifolds is the possibility of relating it to fundamental interactions that have not yet been given a satifactory geometric interpretation. There has been little research in that direction [22], but for those who are interested, the theory of vector manifolds with GC can be recommended as providing ideal mathematical tools [12].

With the above brief background on vector manifolds, we are better able to assess the significance of equation (167). We can read that equation as extracting the algebraic structure of a Minkowski tangent space from the manifold $\mathcal{M}^4$. However, in the intrinsic approach to manifold geometry taken in this paper, the differentiable structure connecting neighboring tangent spaces is not extracted from the manifold, it is imposed on the manifold by defining a connexion and curvature. Consequently, differential computations throughout the present paper involve only the STA at a single point, and all the tangent algebras are isomorphic. This leads one to wonder if we can simplify the theory and get by with a single copy of the STA. The answer is yes, and the result is the flat space theory of spacetime geometry in [2, 4, 3]. Finally, one should note that all the geometry can be extracted from $\mathcal{M}^4$ itself only if it is an embedded manifold.

X. Historical notes

The present approach to GR was initiated in 1966 by my book *Space-Time Algebra* [1]. The crucial innovation there was to reduce the standard representation of spacetime geometry by the metric tensor $g_{\mu\nu}$ to representation by a coordinate frame of vectors g_μ that generate a real geometric algebra at each spacetime point, as described in Section II. I also introduced the local Lorentz transformations of equation (46) and translated Utiyama's gauge formulation of GR [24] into the real STA. However, the gauge concept was not given the central role it has here. My purpose then was just to incorporate the real Dirac equation into GR. I could not have anticipated the rise of gauge theory to the supreme status that it enjoys in theoretical physics today [25].

In 1966 I was blissfully unaware of similar work decades before. Perhaps that was all to the good, as it might have been intimidating or discouraging. At any rate it would have been an unnecessary distraction, because, as I believed then and know now, all my predecessors had missed a key point, namely, the geometric significance of the Dirac algebra. Thus, Schroedinger [26] and others made the Dirac matrices spacetime dependent and related their products to the metric tensor, as in equation (3), in order to incorporate the Dirac equation into GR. To the same end, Fock and Iwanenko [27, 28] and, independently, Weyl[29] in his seminal paper on gauge theory, were evidently the first to introduce the "spin connexion" (165) expressed in terms of Dirac matrices. To do that they were forced to introduce orthonormal frames called *vierbeins* or *tetrads*, which are equivalent to fiducial frames represented by matrix elements. However, they all failed to recognize the Dirac matrices as representations of vectors, so they interpreted their constructions as essentially quantum mechanical rather than fundamentally geometrical. At the same time, their treatment of tetrads as mere auxiliary quantities, shows that they failed to recognize the primary physical significance that we have attributed to fiducial frames.

The main limitation of my 1966 book was the lack of mathematical methods to solve field equations that take advantage of simplifications introduced by GA. To remedy that deficiency I embarked on the development of a Geometric Calculus that culminated in publication of a monograph [12] that, among other things, first formulated the theory of vector manifolds.

The STA formulation of GR in the present paper was developed by 1976, but not published until 1986 [11, 30], because I had originally intended to include it in the GC monograph. The claim in those papers that my method is more efficient than Cartan's exterior calculus in geometric computations was soon supported by direct comparison of computer calculations [31]. However, the most important consequence of that work was stimulating creation of the flat-space gauge theory of gravity by Lasenby, Doran and Gull [3, 4]. That, in turn, stimulated emphasis on gauge theory and the Equivalence Principle in the present paper. Finally, the present approach will be coordinated with the flat-space theory in a sequel to this paper.

Appendix A. Flows, Frames and Integrability

A differentiable one-parameter family of mappings of spacetime into itself is called a *flow*. Although flows appear in a variety of physical contexts, we are 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. Our main objective will be to define a suitable "flow derivative" to describe how quantities change along a flow.

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

$$\frac{dx(\tau)}{d\tau} = v(x(\tau))\,. \tag{170}$$

According to a fundamental theorem in the theory of differential equations, for nonvanishing v equation (170) has the unique solution

$$x(\tau) = f(x,\tau) \tag{171}$$

for a given initial value $x(0) = f(x,0)$. 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 (171) with an initial value for the flow. That choice will be indicated with the subscript notation

$$f_\tau(x) = f(x,\tau) \qquad \text{so that} \qquad f_\tau(f(x,t)) = f(x,t+\tau)\,. \tag{172}$$

The function $f_\tau = f_\tau(x)$ is called the *relative flow* to distinguish it from the *flow* (171), though the difference is usually obvious in context. The relative flows have the following properties of a "transformation group":

(a) *Composition:* $\quad f_\tau \circ f_t = f_{t+\tau}$. (173)

(b) *Associativity:* $\quad (f_\tau \circ f_t)\circ f_s = f_\tau \circ (f_t \circ f_s)$, (174)

(c) *Identity:* $\quad f_0(x) = x$, (175)

(d) *Inverse:* $\quad f_\tau^{-1} = f_{-\tau}$. (176)

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

$$\underline{f}_\tau : \quad u(x) \quad \to \quad \underline{f}_\tau u \equiv u\cdot\nabla f_\tau(x) = u'(f_\tau(x))\,. \tag{177}$$

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

$$u(f_\tau(x)) = \underline{f}_\tau u(x) = u'(f_\tau(x))\,. \tag{178}$$

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

$$\begin{aligned}\mathcal{L}_v u &\equiv \lim_{\tau\to 0}\frac{1}{\tau}\left[\,u(f_\tau(x)) - \underline{f}_\tau u(x)\,\right] = \lim_{\tau\to 0}\frac{1}{\tau}\left[\,\underline{f}_\tau^{-1} u(f_\tau(x)) - u(x)\,\right]\\ &= \frac{d}{d\tau}\,\underline{f}_\tau^{-1} u\Big|_{\tau=0} = \underline{f}\,\frac{d}{d\tau}\left[\,\underline{f}^{-1}u\,\right].\end{aligned} \tag{179}$$

The last form has the advantage of applying to the arbitrary parametrization of points by (171) and so holds for any value of τ. The second form simplifies evaluation of the derivative.

Recalling the definition (27) of the Lie bracket and its relation (29) to the coderivative when torsion vanishes, we find

$$\mathcal{L}_v u = v\cdot\nabla u - u\cdot\nabla v \equiv [\,v, u\,] = v\cdot Du - u\cdot Dv\,. \tag{180}$$

The Lie derivative can also be expressed in terms of the codivergence using the differential identity

$$[\,v, u\,] = D\cdot(v\wedge u) - uD\cdot v + vD\cdot u\,. \tag{181}$$

The Lie derivative definition (179) is easily generalized to any multivector field $A = A(x)$ by extending the differential $\underline{f}_\tau u$, which is a linear transformation of tangent vectors, to a differential $\underline{f}_\tau A$ on the whole tangent algebra. Such an extension is called an *outermorphism* of $\underline{f}_\tau u$. We define it first for a a *simple* or *decomposable* k-vector field $K = K(x)$, for which there exist vector fields $v_i = v_i(x)$ such that

$$K = v_1\wedge v_2\wedge\ldots\wedge v_k\,. \tag{182}$$

This decomposition is only *local*, however, for it may not be possible to find a set of such vector fields covering the whole region on which K is defined. 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. Despite this proviso, we can define the differential of K by

$$\underline{f}_\tau K = \underline{f}_\tau(v_1\wedge v_2\wedge\ldots\wedge v_k) = (\,\underline{f}_\tau v_1)\wedge(\,\underline{f}_\tau v_2)\wedge\ldots\wedge(\,\underline{f}_\tau v_k)\,, \tag{183}$$

and the result is independent of the particular vectors into which K is factored. As it applies to all k-vector components of any multivector, it determines a unique differential for any multivector.

Now we can define the Lie derivative of an arbitrary differentiable multivector field $A = A(x)$ by

$$\mathcal{L}_v A = \frac{d}{d\tau}\, \underline{f}_\tau{}^{-1} A \Big|_{\tau=0} = \underline{f} \frac{d}{d\tau} \left[\, \underline{f}^{-1} A \,\right], \tag{184}$$

which evaluates to

$$\mathcal{L}_v A = v \cdot DA - \grave{v} \wedge (\grave{D} \cdot A) \equiv [\, v, A \,], \tag{185}$$

defining a "generalized Lie-bracket" $[v, A]$. A further generalization of the bracket to arbitrary fields is treated in [12]. Equation (181) likewise generalizes to

$$[\, v, A \,] = D \cdot (v \wedge A) - AD \cdot v + v \wedge (D \cdot A)\,. \tag{186}$$

For a bivector field $A = a \wedge b$, we find

$$[\, v, a \wedge b \,] = [\, v, a \,] \wedge b + a \wedge [\, v, b \,]\,. \tag{187}$$

This generalizes to the "derivation property"

$$\mathcal{L}_v (A \wedge B) = (\mathcal{L}_v A) \wedge B + A \wedge (\mathcal{L}_v B)\,. \tag{188}$$

Thus, the Lie derivative is a derivation with respect to the outer product.

Although the formula (185) does not apply to a scalar field, the definition (184) does hold, since outermorphisms do not alter scalars. Thus, for a scalar field $\varphi = \varphi(x)$ we find

$$\mathcal{L}_v \varphi = v \cdot D\varphi = v \cdot \nabla \varphi\,. \tag{189}$$

With the generalized Lie derivative in hand, we are equipped to survey 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.

Our next task is to study the composition of flows. 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:

$$\underline{f}_\tau\, \underline{g}_s = \underline{g}_s\, \underline{f}_\tau \qquad \text{iff} \qquad [\, v, u \,] = 0\,. \tag{190}$$

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, (190) implies that the two parameter function

$$x(s,\tau) \equiv f_\tau \circ g_s(x_0) = g_s \circ f_\tau(x_0) \tag{191}$$

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

$$\partial_s x(s,\tau) = u(x(s,\tau)), \qquad \partial_\tau x(s,\tau) = v(x(s,\tau))\,. \tag{192}$$

Using this in (180) we find that

$$[\,v,u\,] = (\partial_\tau x)\cdot\nabla(\partial_s x) - (\partial_s x)\cdot\nabla(\partial_\tau x) = (\partial_\tau\partial_s - \partial_s\partial_\tau)x(s,\tau) = 0\,. \tag{193}$$

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 $K = K(x(s,\tau)) \equiv (\partial_\tau x)\wedge(\partial_s x) = u\wedge v$ is everywhere tangent to the surface. It determines a directed area element

$$d^2x = K d\tau ds = (d\tau\,\partial_\tau x)\wedge(ds\,\partial_s x) = (v d\tau)\wedge(u ds)\,. \tag{194}$$

This is the directed area element for a directed integral over the surface as an iterated integral with respect to the scalar parameters. The surface parametrized by (191) is said to be an *integral surface* 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 must be decomposable like $K = K(x)$ in (182). Unlike a vector field, which always has integral curves, a k-vector field might 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 present it in several forms, each of which offers its own insight. To facilitate comparison with standard treatments [20], 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

$$M = D\wedge A + D\cdot B\,. \tag{195}$$

This generalizes the well known *Helmholtz Theorem* of vector analysis [12].

A scalar field $\lambda = \lambda(x)$ is said to be an *integrating factor* for a field $W = W(x)$ if

$$D \wedge (\lambda W) = 0\,. \tag{196}$$

A field which has an integrating factor is also said to be *integrable* or *holonomic*; 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 (182) is *directly integrable*, while W in (196) is *normally integrable.* The reason for saying "normally" appears below.

Now we are ready to state and discuss the

Frobenius theorem: *For $k \geq 0$, 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 $v \wedge K = 0$,*

$$\mathcal{L}_v K = [\,v, K\,] = 0\,. \tag{197}$$

(2) *If $u = u(x)$ is also a vector field satisfying $u \wedge K = 0$, then*

$$[\,v, u\,] \wedge K = 0\,. \tag{198}$$

(3) *The dual of K has an integrating factor so that*

$$D \wedge (\lambda K i) = 0\,. \tag{199}$$

(4) *If $w = w(x)$ is a vector field satisfying $w \cdot K = 0$*, then

$$K \cdot (D \wedge w) = 0\,. \tag{200}$$

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 (197) and (198) are essentially the same. The equation $v \wedge K = 0$ can be interpreted as "*v is contained in K*." Then (198) 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 (182) form a complete set of vectors "contained in K," and therefore the Lie bracket closure condition (198) implies that

$$[\,v_i, v_j\,] = \alpha_{ij}^{m}\, v_m\,, \tag{201}$$

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 (201) is a "classical form" for the integrability condition. Its equivalence to (197) is easily proved by inserting (182) into (197) and using properties of the Lie derivative.

To prove that (201) or (198) 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.

Turn now to the dual versions of the Frobenius theorem (199,200). The condition $w \cdot K = 0$ on the vector w in (200) means that w is *normal* to the integral surfaces of K. The $(n-k)$-vector

$$\hat{K} \equiv Ki\,, \tag{202}$$

where i is the unit pseudoscalar, 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.)

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 (200) can then be written $(iw) \cdot (D \wedge w) = 0$ or, equivalently,

$$w \wedge D \wedge w = 0\,. \tag{203}$$

We can write $w = \hat{K}$, so that (199) becomes

$$D \wedge (\lambda w) = (D\lambda) \wedge w + \lambda D \wedge w = 0\,, \tag{204}$$

which is equivalent to (203). It is now clear that (196) and (199) are called "normal integrability conditions" because they are conditions on the normals of integral surfaces. The normal integrability condition (203) 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 $w_i = w_i(x)$:

$$\hat{K} = w_1 \wedge w_2 \wedge \ldots \wedge w_r\,. \tag{205}$$

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 (200) can be written

$$w_1 \wedge w_2 \wedge \ldots \wedge w_r \wedge D \wedge w_i = 0 \tag{206}$$

for $i = 1, 2, \ldots, r$. This is an obvious generalization of (203).

By virtue of the Poincaré Lemma, the integrability condition (199) implies that $\hat{K}$ is locally exact, so that

$$\lambda \hat{K} = D \wedge A\,, \tag{207}$$

where A is an $(r-1)$-vector field. If $\hat{K} = w$ is a vector, this becomes

$$\lambda w = D\varphi = \nabla \varphi\,, \tag{208}$$

where $\varphi = \varphi(x)$ is a scalar field. The equation

$$\varphi(x) = \mu \tag{209}$$

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

$$K \cdot D\varphi = K \cdot \nabla \varphi = 0\,, \tag{210}$$

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

$$(D\varphi^1) \wedge (D\varphi^2) \wedge \ldots \wedge (D\varphi^r) = \lambda \hat{K} \tag{211}$$

satisfies the integrability condition (199), and a specific A for (207) 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

$$e \equiv e_0 \wedge e_1 \wedge e_2 \wedge e_3, \tag{212}$$

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

$$e^\mu \cdot e_\nu = \delta^\mu_\nu\,, \tag{213}$$

where $\mu, \nu = 0, 1, 2, 3$. These equations can be explicitly solved for the reciprocal vectors e^μ, with the result

$$e^\mu = (-1)^\mu (e_0 \wedge \ldots \wedge \check{e}_\mu \wedge \ldots \wedge e_3) e^{-1}\,, \tag{214}$$

where $\check{e}_\mu$ indicates that e_μ is omitted from the product. Moreover,

$$e^{-1} = \frac{e}{e^2} = e^0 \wedge e^1 \wedge e^2 \wedge e^3\,. \tag{215}$$

To interrelate derivatives of the frames $\{e_\mu\}$ and $\{e^\mu\}$, we consider

$$(e_\mu \wedge e_\nu)\cdot(D\wedge e^\alpha) = e_\mu\cdot(e_\mu\cdot De^\alpha) - e_\nu\cdot(e_\mu\cdot De^\alpha) = [\,e_\mu, e_\nu\,]\cdot e^\alpha\,, \quad (216)$$

where the last step involves differentiating (213). Solving for the Lie bracket, we obtain

$$[\,e_\mu, e_\nu\,] = e_\alpha(e_\mu\wedge e_\nu)\cdot(D\wedge e^\alpha)\,. \quad (217)$$

This is the integrability condition (201), and it gives an explicit expression for the scalar coefficients on the right side of (201). Alternatively, (216) can be solved for

$$D\wedge e^\alpha = \tfrac{1}{2}e^\mu\wedge e^\nu[\,e_\mu, e_\nu\,]\cdot e^\alpha\,. \quad (218)$$

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 identity $(a\cdot K)i = a\wedge(Ki)$ gives

$$|\,e\,|D\cdot(|\,e\,|^{-1}e_\mu) = [\,D\wedge(e_\mu e^{-1})\,]e\,. \quad (219)$$

On the other hand, using (215) 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 (219), we obtain the desired "duality relations"

$$|\,e\,|\,D\cdot(|\,e\,|^{-1}e_\mu) = (D\wedge e^\nu)\cdot(e_\nu\wedge e_\mu)\,. \quad (220)$$

Using (216) to express the right side of this expression in terms of Lie brackets, we find that

$$e^\nu\cdot[\,e_\nu, e_\mu\,] = D\cdot e_\mu - e^\nu\cdot(e_\mu\cdot De_\nu)\,, \quad (221)$$

whence

$$e_\mu\cdot D\,\ln|\,e\,| = e^\nu\cdot(e_\mu\cdot De_\nu) = -e_\nu\cdot(e_\mu\cdot De^\nu)\,. \quad (222)$$

This completes our collection of "differential identities" for arbitrary frames.

A frame $\{e_\mu\}$ is said to be *holonomic* or *integrable* if

$$[\,e_\nu, e_\mu\,] = 0 \quad (223)$$

for all its vectors. By virtue of (217), this is equivalent to the condition

$$D\wedge e^\mu = 0\,. \quad (224)$$

As mentioned already in Section III coordinate frames are holonomic.

The above theory of flows has extensive applications to motions and deformations of material bodies and fields. We must be content here with a single simple example to illustrate the method, derivation of the equation for geodesic deviation.

The location of a *material filament* or *string* at a given instant is described by a spacelike curve $x(\lambda) = g_\lambda(x_0)$. As the filament flows through spacetime with *proper time* τ, it generates a timelike surface

$$x = x(\tau, \lambda) = f_\tau \circ g_\lambda(x_0)\,. \tag{225}$$

The particles of the filament are parametrized by λ with *separation vector* $n = n(x) = \partial_\lambda x$ tangent to the filament and velocity $v = v(x) = \partial_\tau x$ normalized to $v^2 = 1$.

As explained above, the vector fields $v(x)$ and $n(x)$ generate the flow (225) if and only if

$$\mathcal{L}_v(n) = [\,v, n\,] = v \cdot Dn - n \cdot Dv = 0\,. \tag{226}$$

Consequently, the directional coderivative of n is given by

$$\frac{\delta n}{\delta \tau} \equiv v \cdot Dn = n \cdot Dv\,. \tag{227}$$

This describes how n changes along the flow.

To get an equation of motion for n, consider the second coderivative

$$\frac{\delta^2 n}{\delta \tau^2} = (v \cdot D)^2 n = v \cdot D(n \cdot Dv) = n \cdot D(v \cdot Dv) + [\,v \cdot D, n \cdot D\,]v\,. \tag{228}$$

Recall that $[\,v \cdot D, n \cdot D\,]v = R(v \wedge n) \cdot v$, where $R(v \wedge n)$ is the curvature tensor. Also, for a material particle on the filament subject to a net non-gravitational proper force F, the equation of motion is

$$\frac{\delta v}{\delta \tau} = v \cdot Dv = F\,, \tag{229}$$

and the differential of the force along the filament is

$$\underline{F}(n) \equiv n \cdot DF = n \cdot D(v \cdot Dv)\,. \tag{230}$$

Consequently, (228) gives us the equation of motion

$$\frac{\delta^2 n}{\delta \tau^2} = (v \cdot D)^2 n = R(v \wedge n) \cdot v + \underline{F}(n)\,. \tag{231}$$

This equation measures the relative acceleration of neighboring points in a filament. For a geodesic flow F vanishes, and (231) 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 "deformation force."

Appendix B. Comparison of flat and curved space formulations

This appendix is for readers who wish compare the gauge theory formulation of GR for curved space given in this paper with the flat space formulation given in [5].

Table 1. Coordinate frames for flat and curved spacetime

Flat spacetime	Curved spacetime
$x = x(x^0, \ldots, x^3)$	$x = x(x^0, \ldots, x^3)$
$e_\mu = \partial_\mu x$	$g_\mu = \partial_\mu x$
$g_\mu \equiv \underline{h}^{-1}(e_\mu)$	$g_\mu = \underline{h}(\gamma_\mu)$
$x^\mu = x^\mu(x)$	$x^\mu = x^\mu(x)$
$g^\mu = \bar{h}(e^\mu)$	$g^\mu = \nabla(x^\mu)$
$e^\mu = \nabla x^\mu$	$g^\mu = \bar{h}^{-1}(\gamma^\mu)$
$g^\mu \cdot g_\nu = \delta^\mu_\nu$	$g^\mu \cdot g_\nu = \delta^\mu \cdot \delta_\nu = \delta^\mu_\nu$
$g_\mu \cdot g_\nu = e_\mu \cdot (\bar{h}^{-1}\, \underline{h}^{-1} e_\nu) = g_{\mu\nu}$	$g_\mu \cdot g_\nu = \gamma_\mu \cdot (\bar{h}\, \underline{h} \gamma_\nu)$
$\partial_\mu = e_\mu \cdot \nabla = (\underline{h} g_\mu) \cdot \nabla = g_\mu \cdot \bar{\nabla}$	$\partial_\mu = g_\mu \cdot \nabla$
$\nabla = e^\mu \partial_\mu$	$\nabla = g^\mu \partial_\mu$
$\bar{\nabla} = \bar{h}(\nabla) = g^\mu \partial_\mu$	$\nabla = g^\mu \partial_\mu$

The flat space and curved space theories differ primarily in their use of coordinates. Corresponding quantities are listed in Table 1. I have deliberately used the same symbol $\underline{h}$ for the *fiducial tensor* in curved space and for the *gauge tensor* in flat space to facilitate comparison. In surveying Table 1 it will be noticed that the fiducial tensor corresponds to the inverse of the gauge tensor. That trivial difference has been introduced for notational reasons, but it emphasizes that the two tensors map most naturally in opposite directions. The really significant difference is that *the fiducial tensor is coordinate dependent whereas the gauge tensor is not.* This comes about because $\{\gamma_\mu = \underline{h}^{-1}(\partial_\mu x)\}$ is necessarily an orthonormal frame in the fiducial case, whereas in the gauge case, $\{e_\mu = \partial_\mu x\}$ is an arbitrary coordinate frame that is completely decoupled from the gauge tensor. In other words, the remapping of events in spacetime is completely decoupled from changes in coordinates in the gauge theory, whereas the curved space theory has no means to separate passive coordinate changes from shifts in physical configurations. This crucial fact is the reason why in Gauge Theory Gravity the Displacement Gauge Principle has clear physical consequences, whereas in the curved space theory Einstein's General Relativity Principle does not.

Table 2. Comparison of Coderivatives and Connexions

Flat spacetime	Curved spacetime
$g'_\mu = \underline{R}\,g_\mu = \underline{h}'^{-1}(e_\mu)$	$\gamma'_\mu = \underline{R}\,\gamma_\mu = R\gamma_\mu\widetilde{R}$
$\omega_\mu = \omega(g_\mu)$	$\omega_\mu = \omega(g_\mu) = g_\mu\cdot\gamma^\nu\omega(\gamma_\nu)$
	$D_\mu\gamma_\nu \equiv \omega_\mu\cdot\gamma_\nu, \quad \partial_\mu\gamma_\nu \equiv 0$
$D_\mu M = \partial_\mu M + \omega_\mu \times M$	$D_\mu M = \partial_\mu M + \omega_\mu \times M$
$\omega'_\mu = R\omega_\mu\widetilde{R} - 2(\partial_\mu R)\widetilde{R}$	$\omega'_\mu = R\omega_\mu\widetilde{R} + 2(\partial_\mu R)\widetilde{R}$
$D \wedge g^\mu = D \wedge \not{\nabla} x^\mu = 0$	$D \wedge g^\mu = D \wedge \nabla x^\mu = 0$
$H \equiv g^\mu \wedge \omega_\mu = -\frac{1}{2} g^\mu \wedge H(g_\mu)$	$H \equiv g^\mu \wedge \omega_\mu = \frac{1}{2}\gamma^\mu \wedge (D \wedge \gamma_\mu)$
$\omega_\mu = H(g_\mu) + g_\mu \cdot H$	$\omega(\gamma_\mu) = -D \wedge \gamma_\mu + H \cdot \gamma_\mu$
$D_\mu g_\nu = L^\alpha_{\mu\nu} g_\alpha$	$D_\mu g_\nu = L^\alpha_{\mu\nu} g_\alpha$

Mathematical features of the coderivative for flat and curved spacetime are compared in Table 2. Note, in particular, that expressions for $D_\mu M$ have the same form in each case. However, they behave differently under rotation gauge transformations. Whereas the "curved version" simply changes its functional form, the "flat version" transforms according to

$$\overline{L}\,:D_\mu M \quad\to\quad \overline{L}\,(D_\mu M) = D'_\mu M' = \partial_\mu M' + \omega'_\mu \times M'\,, \tag{232}$$

induced by the *active* rotation gauge transformation

$$\overline{L}\,:M \quad\to\quad M' = \overline{L}\,M \equiv LM\widetilde{L}\,. \tag{233}$$

In other words, rotation gauge transformations are represented as passive in the curved version but active in the flat version. This difference translates to a difference in physical interpretation. In this paper we have interpreted passive rotations as expressing equivalence of physics with respect to different inertial reference frames. In the flat theory, however, covariance under active rotations expresses physical equivalence of different directions in spacetime. Thus, "passive equivalence" is an equivalence of observers, while "active equivalence" is an equivalence of states. This distinction generalizes to the interpretation of any relativity (symmetry group) principle: *Active transformations relate equivalent physical states; passive transformations relate equivalent observers.*

As Table 2 shows, the use of common tools of Geometric Calculus for both curved and flat space versions of GR has enabled us to define a coderivative with the same form on both versions, despite differences in the way that fields are attached to the base manifold. It follows that all computations with coderivatives have the same mathematical form in both versions; this includes the curvature tensor and all its properties as well as the whole panoply of GR. Accordingly, all such results in this paper and in [5] are identical, so further discussion is unnecessary. By the way, this fact can be regarded as a proof of equivalence of Einstein's curved space GR with flat space Gauge Theory Gravity.

References

[1] D. Hestenes, *Space-Time Algebra,* (Gordon & Breach, New York, 1966).

[2] D. Hestenes, "Spacetime physics with geometric algebra," Am. J. Phys. **71**: 691–704 (2003).

[3] A. Lasenby, C. Doran, & S. Gull, "Gravity, gauge theories and geometric algebra," Phil. Trans. R. Lond. A **356**: 487–582 (1998).

[4] C. Doran & A. Lasenby, *Geometric Algebra for Physicists* (Cambridge U Press, Cambridge, 2003)

[5] D. Hestenes, "Gauge Theory Gravity with Geometric Calculus," Am. J. Phys. **??**: ??? (2005).

[6] Much more literature on extensions and applications of Geometric Algebra and Calculus can be accessed from the websites
<http://modelingnts.la.asu.edu> and
<http://www.mrao.cam.ac.uk/˜clifford/>.

[7] S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).

[8] Misner, Thorne and Wheeler [18], pp. 226–228 provide a fine pedagogical discussion of the tangent vector concept.

[9] B. Iliev, "Is the principle of equivalence a principle?," Journal of Geometry and Physics **24**: 209–222 (1998).

[10] F. Heyl, P. von der Heyde & D. Kerlick, "General relativity with spin and torsion: Foundations and Prospects," Reviews of Modern Physics **48**: 393–416 (1976).

[11] D. Hestenes, "Curvature Calculations with Spacetime Algebra," Int. J. Theo. Phys., **25**: 581–588 (1986).

[12] D. Hestenes & G. Sobczyk, *CLIFFORD ALGEBRA to GEOMETRIC CALCULUS, a Unified Language for Mathematics and Physics* (Kluwer Academic, Dordrecht/Boston, 1986).

[13] Misner, Thorne and Wheeler [18], p. 356 demonstrate curvature calculation with differential forms.

[14] D. Hestenes, *New Foundations for Classical Mechanics,* (Kluwer, Dordrecht/Boston, 1986). Second Edition (1999).

[15] Actually, the last steps in reducing equation (134) in [2] to equation (141) here were inadvertently omitted from the text. However, they can be found in the original reference: D. Hestenes, "Proper Particle Mechanics," J. Math. Phys. **15**, 1768–1777 (1974).

[16] Equation (158) here agrees with equation (40.33) on p.1118 of [18], where $\lambda = \gamma U$ and $\Phi = -U$.

[17] Misner, Thorne and Wheeler [18], pp. 226–241, follow Cartan in arguing at length for the heuristic value of equation (167) and the value of Cartan's Calculus of Differential Forms.

[18] C. Misner, K. Thorne & J. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

[19] R. Wald, *General Relativity* (U. Chicago Press, Chicago, 1984).

[20] M. Göckler & T. Schücker, *Differential Geometry, Gauge Theories and Gravity* (Cambridge U Press, Cambridge, 1987).

[21] D. Hestenes, "Differential Forms in Geometric Calculus." In F. Brackx, R. Delanghe, and H. Serras (eds), *Clifford Algebras and their Applications in Mathematical Physics* (Kluwer Academic, Dordrecht/Boston, 1993). pp. 269–285.

[22] H. Goenner, "Local Isometric Embedding of Riemannian Manifolds and Einstein's Theory of Gravitation." In Held (Ed.), *General Relativity and Gravity, Vol. I,* (Plenum, New York, 1980).

[23] 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, Boston, 1986), pp. 117–244.

[24] R. Utiyama, "Invariant Theoretical Interpretation of Interaction," Phys. Rev. **101**: 1597–1607 (1956).

[25] L. OŔaifeartaigh & N. Straumann, "Gauge theory: Historical origins and some modern developments", Reviews of Modern Physics **72**: 1–23 (2000).

[26] E. Schroedinger, *Sitzungb. Preuss. Akad. Wiss. Phys.-Math. Kl.* 105 (1932).

[27] V. Fock, "Geometrisierung der Diracschen Theorie des Elektrons," Zeitshrift für Physik, **57**:261–277 (1929).

[28] V. Fock & D. Iwanenko , Géométrie quantique linéaire et déplacement parallèle, Comptes Rendus des séances de L'Académie des Sciences, **188**:1470–1472 (1929).

[29] H. Weyl, "Elektron and Gravitation I." Zeitshrift für Physik, **56**: 330 (1929).

[30] D. Hestenes, "A Spinor Approach to Gravitational Motion and Precession," Int. J. Theo. Phys., **25**, 589–598 (1986).

[31] A. Moussiaux & P. Tombal, "Geometric Calculus: A New Computational Tool for Riemannian Geometry," Int. J. Theo. Phys., **27**: 613–621 (1988).