Gauge Theory Gravity with Geometric Calculus

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

A new *gauge theory of gravity* on flat spacetime has recently been developed by Lasenby, Doran, and Gull. Einstein's principles of equivalence and general relativity are replaced by gauge principles asserting, respectively, local rotation and global displacement gauge invariance. A new unitary formulation of Einstein's tensor illuminates long-standing problems with energy-momentum conservation in general relativity. *Geometric calculus* provides many simplifications and fresh insights in theoretical formulation and physical applications of the theory.

I. Introduction

More than a decade before the advent of Einstein's general theory of relativity (GR), and after a lengthy and profound analysis of the relation between physics and geometry [1], Henri Poincaré concluded that:

> "One geometry cannot be more true than another; it can only be more convenient. *Now, Euclidean geometry is and will remain, the most convenient.* ... What we call a straight line in astronomy is simply the path of a ray of light. If, therefore, we were to discover negative parallaxes, or to prove that all parallaxes are higher than a certain limit, we should have a choice between two conclusions: we could give up Euclidean geometry, or modify the laws of optics, and suppose that light is not rigorously propagated in a straight line. *It is needless to add that every one would look upon this solution as the more advantageous.*" [italics added]

Applied to GR, this amounts to claiming that any curved space formulation of physics can be replaced by an equivalent and simpler flat space formulation. Ironically, the curved space formulation has been preferred by nearly everyone since the inception of GR, and many attempts at alternative flat space formulations have failed to exhibit the simplicity anticipated by Poincaré. One wonders if the trend might have been different if Poincaré were still alive to promote his view when GR made its spectacular appearance on the scene.

A dramatic new twist on the physics-geometry connection has been introduced by Cambridge physicists Lasenby, Doran, and Gull with their flat space

[1]electronic mail:hestenes@asu.edu

alternative to GR called *gauge theory gravity* (GTG) [2, 3]. With geometric calculus (GC) as an essential tool, they clarify the foundations of GR and provide many examples of computational simplifications in the flat space gauge theory. All this amounts to compelling evidence that Poincaré was right in the first place [4].

The chief innovation of GTG is a *displacement gauge principle* that asserts global homogeneity of spacetime. This is a new kind of gauge principle that clarifies and revitalizes Einstein's *general relativity principle* by cleanly separating arbitrary coordinate transformations from physically significant field transformations. GTG also posits a *rotation gauge principle* of standard type in nonAbelian gauge theory. The present account is unique in interpreting this principle as a realization of Einstein's Equivalence Principle asserting local isotropy of spacetime. With rotation gauge equivalence and displacement gauge equivalence combined, GTG synthesizes Einstein's principles of equivalence and general relativity into a new general *principle of gauge equivalence.*

A unique consequence of GTG, announced here for the first time, is the existence of a displacement gauge invariant energy-momentum conservation law.

Many previous attempts to formulate a flat space alternative to GR are awkward and unconvincing. The most noteworthy alternative is the self-consistent field theory of a massless spin-2 particle (the graviton) [5]. That is a gauge theory with local Lorentz rotations as gauge transformations. However, it lacks the simplicity brought by the Displacement Gauge Principle in GTG.

GTG is also unique in its use of *geometric calculus* (GC). It is, of course, possible to reformulate GTG in terms of more standard mathematics, but at the loss of much of the theory's elegance and simplicity. Indeed, GC is so uniquely suited to GTG it is doubtful that the theory would have developed without it. That should be evident in the details that follow.

This is the third in a series of articles promoting geometric algebra (GA) as a unified mathematical language for physics [6, 7]. Here, GA is extended to a *geometric calculus* (GC) that includes the tools of differential geometry needed for Einstein's theory of *general relativity* (GR) on flat spacetime. 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 teaching and research. This provides the last essential piece for a comprehensive *geometric algebra and calculus* expressly designed to serve the purposes of theoretical physics [8, 9].

The preceding article [7] (hereafter referred to as GA2) is a preferred prerequisite, but to make the present article reasonably self-contained, a summary of essential concepts, notations and results from GA2 is included. Of course, prior familiarity with standard treatments of GR will be helpful as well. However, for students who have mastered GA2, the present article may serve as a suitable entrée to GR. As a comprehensive treatment of GR is not possible here, it will be necessary for students to coordinate study of this article with one of the many fine textbooks on GR. In a course for undergraduates, the textbook by d'Inverno [10] would be a good choice for that purpose. In my experience, the challenge of reformulating GR in terms of GC is a great stimulus to student learning.

As emphasized in GA2, one great advantage of adopting GC as a unified language for physics is that it eliminates unnecessary language and conceptual barriers between classical, quantum and relativistic physics. I submit that the simplifications introduced by GC are essential to fitting an adequate treatment of GR into the undergraduate physics curriculum. It is about time for black holes and cosmology to be incorporated into the standard curriculum. Although space limitations preclude addressing such topics here, the mathematical tools introduced are sufficient to treat any topic with GC, and more details are given in a recent book [2].

Recognition that GR should be formulated as a gauge theory has been a long time coming, and, though it is often discussed in the research literature [11], 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 in 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. The GC approach is pedagogically efficient as well, as it develops GR by a straightforward generalization of Special Relativity using mathematical tools already well developed in GA2.

Essential mathematical tools introduced in GA2 are summarized in Section II, though mastery of GA2 may be necessary to understand the more subtle aspects of the theory. Section III extends the tool kit to unique mathematical tools for linear algebra and induced transformations on manifolds. These tools are indispensible for GTG and useful throughout the rest of physics.

Section IV introduces the gauge principles of GTG and shows how they generate an induced geometry on spacetime that is mathematically equivalent to the Riemannian geometry of GR. Some facility with GC is needed to appreciate how it streamlines the formulation, analysis and application of differential geometry, so the more subtle derivations have been relegated to appendices.

Section V discusses the formulation of field equations and equations of motion in GTG. Besides standard results of GR, it includes a straightforward extension of spinor methods in GA2 to treat gravitational precession and gravitational interactions in the Dirac equation.

Section VI discusses simplifications that GTG brings to the formulation and analysis of solutions to the gravitational field equations, in particular, motion in the field of a black hole.

Section VII introduces a new split of the Einstein tensor to produce a new definition for the energy-momentum tensor and a general energy-momentum conservation law.

Section VIII summarizes the basic principles of GTG and discusses their status in comparison with Newton's Laws as *universal principles for physics*.

II. Spacetime Algebra Background

This section reviews and extends concepts, notations and results from GA2 that will be applied and generalized in this paper. Our primary mathematical tool is the the *real Spacetime Algebra* (STA). One great advantage of STA is that it enables coordinate-free representation and computation of physical systems and processes. Another is that it incorporates the spinors of quantum mechanics along with the tensors (in coordinate-free form) of classical field theory. Some results from GA2 are summarized below without proof, and the reader is encouraged to consult GA2 for more details. Other results from GA2 will be recalled as needed. *The reader should note that most of GA2 carries over without change; the essential differences all come from a generalized gauge concept of differentiation.* We show in subsequent sections that the gauge covariant derivative accommodates curved-space geometry in flat space to achieve a flat-space gauge theory of gravity. However, the accomodation is subtle, so it was discovered only recently.

A. STA elements and operations

For physicists familiar with the Dirac matrix algebra, the quickest approach to STA is by reinterpreting the Dirac matrices as an *orthonormal basis* $\{\gamma_\mu; \mu = 0, 1, 2, 3\}$ for a 4D *real Minkowski vector space* $\mathcal{V}^4$ with signature specified by the rules:

$$\gamma_0^2 = 1 \qquad \text{and} \qquad \gamma_1^2 = \gamma_2^2 = \gamma_3^2 = -1\,. \tag{1}$$

Note that the scalar 1 in these equations would be replaced by the identity matrix if the γ_μ were Dirac matrices. Thus, (1) is no mere shorthand for matrix equations but a defining relation of vectors to scalars that encodes spacetime signature in algebraic form.

The frame $\{\gamma_\mu\}$ generates an associative geometric algebra that is isomorphic to the Dirac algebra. The product $\gamma_\mu\gamma_\nu$ of two vectors is called the *geometric product*. The usual *inner product* of vectors is defined by

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

where $\eta_\mu = \gamma_\mu^2$ is the *signature indicator*. The *outer product*

$$\gamma_\mu \wedge \gamma_\nu \equiv \tfrac{1}{2}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu) = -\gamma_\nu \wedge \gamma_\mu\,, \tag{3}$$

defines a new entity called a *bivector* (or 2-vector), which can be interpreted as a directed plane segment representing the plane containing the two vectors.

STA is the *geometric algebra* $\mathcal{G}_4 = \mathcal{G}(\mathcal{V}^4)$ generated by $\mathcal{V}^4$. A full basis for the algebra is given by the set:

1	$\{\gamma_\mu\}$	$\{\gamma_\mu \wedge \gamma_\nu\}$	$\{\gamma_\mu i\}$	i
1 scalar	4 vectors	6 bivectors	4 trivectors	1 pseudoscalar
grade 0	grade 1	grade 2	grade 3	grade 4

where the *unit pseudoscalar*

$$i \equiv \gamma_0\gamma_1\gamma_2\gamma_3 \tag{4}$$

squares to -1, anticommutes with all odd grade elements and commutes with even grade elements. Thus, $\mathcal{G}_4$ is a linear space of dimension $1+4+6+4+1 = 2^4 = 16$.

A generic element of $\mathcal{G}_4$ is called a *multivector*. Any multivector can be expressed as a linear combination of the basis elements. For example, a bivector F has the expansion

$$F = \tfrac{1}{2}F^{\mu\nu}\gamma_\mu \wedge \gamma_\nu\,, \tag{5}$$

with its "scalar components" $F^{\mu\nu}$ in the usual tensorial form. For components, we use the usual tensor algebra conventions for raising and lowering indices and summing over repeated upper and lower index pairs.

Any multivector M can be written in the *expanded forms*

$$M = \alpha + a + F + bi + \beta i = \sum_{k=0}^{4} \langle M\rangle_k, \tag{6}$$

where $\alpha = \langle M\rangle_0$ and β are scalars, $a = \langle M\rangle_1$ and b are vectors, and $F = \langle M\rangle_2$ is a bivector, while $bi = \langle M\rangle_3$ is a trivector (or pseudovector) and $\beta i = \langle M\rangle_4$ is a pseudoscalar. It is often convenient to drop the subscript on the scalar part so $\langle M\rangle = \langle M\rangle_0$. The scalar part behaves like the "trace" in matrix algebra; for example, $\langle MN\rangle = \langle NM\rangle$ for arbitrary multivectors M and N. A multivector is said to be *even* if the grades of its nonvanishing components are all even. The even multivectors compose the *even subalgebra* of $\mathcal{G}_4$, which is, of course, closed under the geometric product.

Coordinate-free computations are facilitated by various definitions. The operation of *reversion* reverses the order in a product of vectors, so for vectors a, b, c it is defined by

$$(abc)^{\sim} = cba\,. \tag{7}$$

It follows for any multivector M in the expanded form (6) that the *reverse* $\widetilde{M}$ is given by

$$\widetilde{M} = \alpha + a - F - bi + \beta i\,. \tag{8}$$

Computations are also facilitated by defining various products in terms of the fundamental geometric product.

The inner and outer products of vectors, (2) and (3) can be generalized to arbitrary multivectors as follows. For $A = \langle A \rangle_r$ and $B = \langle B \rangle_s$ of grades $r, s \geq 0$, inner and outer products can be defined by

$$A \cdot B \equiv \langle AB \rangle_{|r-s|}, \qquad A \wedge B \equiv \langle AB \rangle_{r+s} . \tag{9}$$

Coordinate-free manipulations are facilitated by a system of identities involving inner and outer products [8]. These identities generalize to arbitrary dimensions the well-known identities for dot and cross products in ordinary 3D vector algebra. Only a few of the most commonly used identities are listed here. For a vector a, the geometric product is related to inner and outer products by

$$aB = a \cdot B + a \wedge B . \tag{10}$$

For vectors a, b, c the most commonly used identity is

$$a \cdot (b \wedge c) = (a \cdot b)c - (a \cdot c)b = a \cdot b\, c - a \cdot c\, b . \tag{11}$$

This is a special case of

$$a \cdot (b \wedge B) = a \cdot b\, B - b \wedge (a \cdot B) \tag{12}$$

for grade$(B) \geq 2$, and we have a related identity

$$a \cdot (b \cdot B) = (a \wedge b) \cdot B . \tag{13}$$

We also need the *commutator product*

$$A \times B \equiv \tfrac{1}{2}(AB - BA). \tag{14}$$

This is especially useful when A is a bivector. Then we have the identity

$$AB = A \cdot B + A \times B + A \wedge B, \tag{15}$$

which should be compared with (10). For $s =$ grade(B), it is easy to prove that the three terms on the right side of (15) have grades $s-2, s, s+2$ respectively.

Note the dropping of parentheses on the right hand side of (11). To reduce the number of parentheses in an expression, we sometimes use a *precedence convention*, which allows that inner, outer and commutator products take precedence over the geometric product in ambiguous expressions.

Finally, we note that the concept of *duality*, which appears throughout mathematics, has a very simple realization in GA. The *dual* of any multivector is simply obtained by multiplying it by the pseudoscalar i (or, sometimes, by a scalar multiple thereof). Thus, in (6) the trivector bi is the dual of the vector b. Inner and outer products are related by the duality identities

$$a \cdot (Bi) = (a \wedge B)i , \qquad a \wedge (Bi) = (a \cdot B)i . \tag{16}$$

B. Lorentz rotations and rotors

A complete treatment of Lorentz transformations is given in GA2. We are concerned here only with transformations continuously connected to the identity called *Lorentz rotations*.

Every Lorentz rotation $\underline{R}$ has an explicit algebraic representation in the canonical form

$$\underline{R}(a) = Ra\widetilde{R}\,, \tag{17}$$

where R is an even multivector called a *rotor*, which is subject to the normalization condition

$$R\widetilde{R} = 1\,. \tag{18}$$

The rotors form a multiplicative group called the *rotor group*, which is a double-valued representation of the Lorentz rotation group, also called the *spin group* or SU(2). The underbar serves to indicate that $\underline{R}$ is a linear operator. An overbar is used to designate its adjoint $\overline{R}$. For a rotation, the adjoint is also the inverse; thus,

$$\overline{R}\,(a) = \widetilde{R}\,aR\,. \tag{19}$$

A rotor is a special kind of spinor — as useful in classical physics as in quantum mechanics.

For a particle history parametrized by proper time τ, the particle velocity $v = v(\tau)$ is a unit vector, so it can be obtained at any time τ from the fixed unit timelike vector γ_0 by a Lorentz rotation

$$v = R\gamma_0\widetilde{R}\,, \tag{20}$$

where $R = R(\tau)$ is a one-parameter family of rotors. It follows from (18) that R must satisfy the *rotor equation of motion*

$$\dot{R} = \tfrac{1}{2}\Omega R\,, \tag{21}$$

where the overdot indicates differentiation with respect to proper time and $\Omega = \Omega(\tau)$ is a bivector-valued function, which can be interpreted as a generalized *rotational velocity*. It follows by differentiating (20) that the equation of motion for the particle velocity must have the form

$$\dot{v} = \Omega \cdot v\,. \tag{22}$$

The dynamics of particle motion is therefore completely determined by $\Omega(\tau)$. For example, for a classical particle with mass m and charge q in an electromagnetic field F, the dynamics is specified by

$$\Omega = \frac{q}{m}\,F\,, \tag{23}$$

which gives the standard *Lorentz Force* when inserted into (22). One of our objectives in this paper is to ascertain what Ω should be for a particle subject to a gravitational force.

One major advantage of the rotor equation (20) for the velocity is that it generalizes immediately to an orthonormal frame comoving with the velocity:

$$e_\mu = R\gamma_\mu \widetilde{R} \,, \tag{24}$$

where, of course, $v = e_0$. Solution of the rotor equation of motion (21) therefore gives precession of the comoving frame along with the velocity vector. In Section VII we adapt this approach to gyroscope precession in a gravitational field.

C. Vector derivatives, differential forms and field equations

Geometric calculus is the extension of a geometric algebra like STA to include differentiation and integration. The fundamental differential operator in geometric calculus is the *vector derivative*. Although the vector derivative can be defined in a coordinate-free way [9, 8], the quickest approach is to use the reader's prior knowledge about partial differentiation.

For a vector variable $a = a^\mu \gamma_\mu$ defined on $\mathcal{V}^4$, the *vector derivative* can be given the operator definition

$$\partial_a = \gamma^\mu \frac{\partial}{\partial a^\mu} \,. \tag{25}$$

This can be used to evaluate the following specific derivatives where $K = \langle K \rangle_k$ is a constant k-vector:

$$\partial_a \, a \cdot K = kK \,, \tag{26}$$

$$\partial_a \, a \wedge K = (n-k)K \,, \tag{27}$$

$$\partial_a \, Ka = \gamma^\mu K \gamma_\mu = (-1)^k (n-2k)K \,, \tag{28}$$

where n is the dimension of the vector space ($n = 4$ for spacetime).

In special relativity the location of an event is designated by a vector x in $\mathcal{V}^4$. For the derivative with respect to a spacetime point x we usually use the special symbol

$$\bigtriangledown \equiv \partial_x \,. \tag{29}$$

This agrees with the standard symbol $\bigtriangledown \varphi$ for the gradient of a scalar field $\varphi = \varphi(x)$. Moreover, the same symbol is used for the vector derivative of any multivector field. It is most helpful to consider a specific example.

Let $F = F(x)$ be an electromagnetic field. This is a bivector field, with standard tensor components given by (5). The vector derivative enables us to express *Maxwell's equation* in the compact coordinate-free form

$$\bigtriangledown F = J \,, \tag{30}$$

where the vector field $J = J(x)$ is the charge current density.

The vector derivative can be separated into two parts by using the the *differential identity*

$$\nabla F = \nabla \cdot F + \nabla \wedge F \,, \tag{31}$$

which is an easy consequence of the identity (10). The terms on the right-hand side of (29) are called, respectively, the *divergence* and *curl* of F. Since $\nabla \cdot F$ is a vector and $\nabla \wedge F$ is a trivector, we can separate vector and trivector parts of (29) to get two *Maxwell's equations*:

$$\nabla \cdot F = J \,, \tag{32}$$

$$\nabla \wedge F = 0 \,. \tag{33}$$

Although these separate equations have distinct applications, the single equation (30) is easier to solve for given source J and boundary conditions [9].

Something needs to be said here about differential forms, though we shall not need them in this paper, because we will not be doing surface integrals. Differential forms have become increasingly popular in theoretical physics and especially in general relativity since they were strongly advocated by Misner, Thorne and Wheeler [12]. The following brief discussion is intended to convince the reader that all the affordances of differential forms are present in geometric calculus and to enable translation from one language to the other.

A differential k-form ω is a scalar-valued linear function, which can be given the explicit form

$$\omega = d^k x \cdot K \,, \tag{34}$$

where $K = K(x)$ is a k-vector field and $d^k x$ is a k-vector-valued volume element. In this representation, Cartan's *exterior derivative* $d\omega$ is given by

$$d\omega = d^{k+1}x \cdot (\nabla \wedge K) \,. \tag{35}$$

Thus, the exterior derivative of a k-form is equivalent to the curl of a k-vector. Differential forms are often used to cast the *Fundamental Theorem of Integral Calculus* (also known as the generalized Stokes' Theorem) in the form

$$\int_{\mathcal{R}} d\omega = \oint_{\partial\mathcal{R}} \omega \,. \tag{36}$$

By virtue of (34) and (35), this theorem applies to a differentiable k-vector field $K = K(x)$ defined on a $(k+1)$-dimensional surface $\mathcal{R}$ with a k-dimensional boundary $\partial\mathcal{R}$.

Now it is easy to translate Maxwell's equations from STA into differential forms. The electromagnetic field F and current J become a 2-form and a 1-form:

$$\omega = d^2 x \cdot F \,, \qquad \alpha = J \cdot dx \,. \tag{37}$$

Dual forms can be defined by

$$*\omega \equiv d^2x \cdot (Fi)\,, \qquad *\alpha \equiv d^3x \cdot (Ji)\,. \tag{38}$$

The duality identity (16) implies the duality of divergence and curl:

$$(\nabla \cdot F)i = \nabla \wedge (Fi). \tag{39}$$

Consequently, the Maxwell equations (32) and (33) are equivalent to the equations

$$d * \omega = *\alpha\,, \tag{40}$$

$$d\omega = 0\,. \tag{41}$$

The standard formalism of differential forms does not allow us to combine these two equations into a single equation like $\nabla F = J$. This is related to a more serious limitation of the differential forms formalism: It is not suitable for dealing with spinors.

As we are interested to see how quantum mechanics generalizes to curved spacetime, we record without proof the essential results from GA2, which should be consulted for a detailed explanation. To be specific, we consider the electron as the quintessential fermion.

The electron wave function is a *real spinor field* $\psi = \psi(x)$, an even multivector in the real STA. For the electron with charge e and mass m, the field equation for ψ is the *real Dirac equation*

$$\nabla\psi\gamma_2\gamma_1\hbar - eA\psi = m\psi\gamma_0\,, \tag{42}$$

where $A = A_\mu\gamma^\mu$ is the electromagnetic vector potential. This coordinate-free version of the Dirac equation has a number of remarkable properties, beginning with the fact that it is formulated entirely in the real STA and so implies that complex numbers in the standard matrix version of the Dirac equation have a geometric interpretation. Indeed, the unit imaginary is explicitly identified in (38) with the spacelike bivector $\gamma_2\gamma_1$, which does square to minus one. The vector derivative $\nabla = \gamma^\mu\partial_\mu$ will be recognized as the famous differential operator introduced by Dirac, except that the γ^μ are vectors rather than matrices. Our reformulation of Dirac's operator as a vector derivative shows that it is the fundamental differential operator for all of spacetime physics, not just quantum physics. The present paper shows that adapting this operator is the main problem in adapting quantum mechanics to curved spacetime.

Physical interpretation of the Dirac equation depends on the specification of "observables," which are bilinear functions of the wave function. The Dirac wave function determines an orthonormal frame field of *local observables*

$$\psi\gamma_\mu\widetilde{\psi} = \rho e_\mu\,, \tag{43}$$

where $\rho = \rho(x)$ is a scalar probability density and

$$e_\mu = R\gamma_\mu \widetilde{R}\,, \tag{44}$$

where $R = R(x)$ is a rotor field. The vector field $\psi\gamma_0\widetilde{\psi} = \rho e_0$ is the *Dirac probability current*, which doubles as a charge current when multiplied by the charge e. The vector field $e_3 = R\gamma_3\widetilde{R}$ specifies the local direction of electron spin. The vector fields e_1 and e_2 specify the local phase of the electron, and $e_2 e_1 = R\gamma_2\gamma_1\widetilde{R}$ relates the unit imaginary in the Dirac equation to electron spin. A full understanding of this point requires more explanation than can be given here. Details are given in GA2.

Finally, note the strong similarity between the spinor frame field (40) and the comoving frame (24) for a classical particle. This provides a common ground for describing spin precession in both classical and quantum mechanics.

III. Mathematical Tools

This section introduces powerful mathematical tools and theorems of general utility throughout physics, though the treatment here is limited to results needed for gauge gravity theory. It is particularly noteworthy that these tools enable linear algebra and differential transformations without matrices or coordinates.

A. Linear Algebra

Within geometric calculus (GC), *linear algebra* is the theory of linear vector-valued functions of a vector variable. GC makes it possible to perform coordinate-free computations in linear algebra without resorting to matrices, as demonstrated in the basic concepts, notations and theorems reviewed below. Linear algebra is a large subject, so we restrict our attention to the essentials needed for gravitation theory. A more extensive treatment of linear algebra with GC is given elsewhere [13, 8, 14] as well as [2].

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

$$\underline{f}a = \underline{f}(a)\,. \tag{45}$$

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

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

$$\underline{f}(A \wedge B) = (\underline{f}A) \wedge (\underline{f}B) \tag{46}$$

for arbitrary multivectors A, B, and

$$\underline{f}\alpha = \alpha \tag{47}$$

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

$$\underline{f}(A) = (\underline{f}a_1) \wedge (\underline{f}a_2) \wedge \ldots \wedge (\underline{f}a_r)\,. \tag{48}$$

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

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

$$\underline{f}\langle\, M\,\rangle_k = \langle\, \underline{f}M\,\rangle_k \tag{49}$$

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

$$\underline{f}(i) = (\det \underline{f})i \qquad \text{or} \qquad \det \underline{f} = -i\,\underline{f}(i)\,, \tag{50}$$

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

The "product" of two linear transformations, expressed by

$$\underline{h} = \underline{g}\,\underline{f} \tag{51}$$

applies also to their outermorphisms. In other words, the outermorphism of a product equals the product of outermorphisms. It follows immediately from (51) that

$$\det(\underline{g}\,\underline{f}) = (\det \underline{g})(\det \underline{f})\,, \tag{52}$$

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

$$\det(\underline{f}^{-1}) = (\det \underline{f})^{-1} \tag{53}$$

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

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

$$\langle\, M\bar{f}N\,\rangle = \langle\, N\underline{f}M\,\rangle\,, \tag{54}$$

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

$$b \cdot \bar{f}(a) = a \cdot \underline{f}(b) . \tag{55}$$

Differentiating with respect to b we obtain [15],

$$\bar{f}(a) = \partial_b a \cdot \underline{f}(b) . \tag{56}$$

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

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

$$\bar{f}[\underline{f}(A) \cdot B] = A \cdot \bar{f}(B) \tag{57}$$

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

$$\bar{f}[A \cdot B] = \underline{f}^{-1}(A) \cdot \bar{f}(B) . \tag{58}$$

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

$$\underline{f}^{-1} A = [\bar{f}(Ai)](\bar{f} i)^{-1} = (\det \underline{f})^{-1} \bar{f}(Ai) i^{-1} . \tag{59}$$

This relation shows explicitly the double duality inherent in computing the inverse of a linear transformation, but not at all obvious in the matrix formulation.

B. Transformations and Covariants

This section describes the apparatus of geometric calculus [8] for handling transformations of spacetime and corresponding induced transformations of multivector fields on spacetime. We concentrate on transformations (or mappings) of 4-dimensional regions, including the whole of spacetime, but our apparatus applies with only minor adjustments to mapping vector manifolds of any dimension including submanifolds in spacetime. It also applies to curved as well as flat manifolds and allows the target of a mapping to be a different manifold. As a matter of course, we assume whatever differentiability is required for performing indicated operations. Accordingly, all transformations are presumed to be smooth and invertible unless otherwise indicated.

In the next section we model spacetime as a vector manifold $\mathcal{M}^4 = \{x\}$ with tangent space $\mathcal{V}^4(x)$ at each spacetime point x. Tangent vectors in $\mathcal{V}^4(x)$ are differentiably attached to $\mathcal{M}^4$ in two distinct ways: first, as tangent to a curve; second as gradient of a scalar. Thus, a smooth curve $x = x(\lambda)$ through a point x has tangent

$$\frac{dx}{d\lambda} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \{x(\lambda + \epsilon) - x(\lambda)\}. \tag{60}$$

A scalar field $\phi = \phi(x)$ has a gradient determined by the vector derivative

$$\nabla\phi = \nabla_x \phi(x) = \partial_x \phi(x). \tag{61}$$

The two kinds of derivative are related by the *chain rule*:

$$\frac{d\phi(x(\lambda))}{d\lambda} = \frac{dx}{d\lambda} \cdot \nabla_x \phi(x). \tag{62}$$

By the way, we usually write $\nabla = \nabla_x$ for the derivative with respect to a spacetime point and ∂_a for the derivative with repect to a vector variable a in $\mathcal{V}^4(x)$. The vector-valued gradient $\nabla\phi$ is commonly called a *covector* to distinguish its kind from the *vector* $dx/d\lambda$. The difference in kind is manifest in the different ways they transform, as we see next.

Let f be a smooth mapping (i.e. diffeomorphism) that transforms each point x in some region of spacetime into another point x', as expressed by

$$f: \; x \quad \to \quad x' = f(x)\,. \tag{63}$$

The mapping (63) induces a linear transformation of tangent vectors at x to tangent vectors at x', given by the differential

$$\underline{f}: \quad a \quad \to \quad a' = \underline{f}(a) = a \cdot \nabla f\,. \tag{64}$$

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

$$a' = a'(x') \equiv \underline{f}[a(x);\, x] = \underline{f}[\,a(f^{-1}(x'));\; f^{-1}(x')\,]\,. \tag{65}$$

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

$$\underline{f}(i) = J_f i, \qquad \text{where} \qquad J_f = \det \underline{f} = -i\,\underline{f}i \tag{66}$$

is known as the *Jacobian* of f.

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

$$\bar{f}: \quad b' \quad \to \quad b = \bar{f}(b') \equiv \grave{\nabla}\grave{f} \cdot b' = \nabla_x f(x) \cdot b'\,. \tag{67}$$

More explicitly, for covector fields

$$\bar{f}: \quad b'(x') \quad \to \quad b(x) = \bar{f}\,[\,b'(x');\, x\,] = \bar{f}\,[\,b'(f(x));\, x\,]\,. \tag{68}$$

As in (55), the differential is related to the adjoint by

$$b' \cdot \underline{f}(a) = a \cdot \bar{f}\,(b')\,. \tag{69}$$

Since the induced transformations for vectors and covectors are linear transformations, they generalize as outermorphisms to differential and adjoint transformations of the entire tangent algebra at each point of the manifold. According to (59), therefore, the outermorphism $\bar{f}$ determines the inverse transformation

$$\underline{f}^{-1}(a') = \bar{f}(a'i)(J_f i)^{-1} = a\,. \tag{70}$$

Also, however,

$$\underline{f}^{-1}(a') = a' \cdot \nabla_{x'} f^{-1}(x')\,. \tag{71}$$

Hence, the inverse of the differential equals the differential of the inverse.

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

$$\overline{f^{-1}}: \quad b(x) \quad \to \quad b'(x') = \overline{f^{-1}}\,[\,b(f^{-1}(x'))\,]\,. \tag{72}$$

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

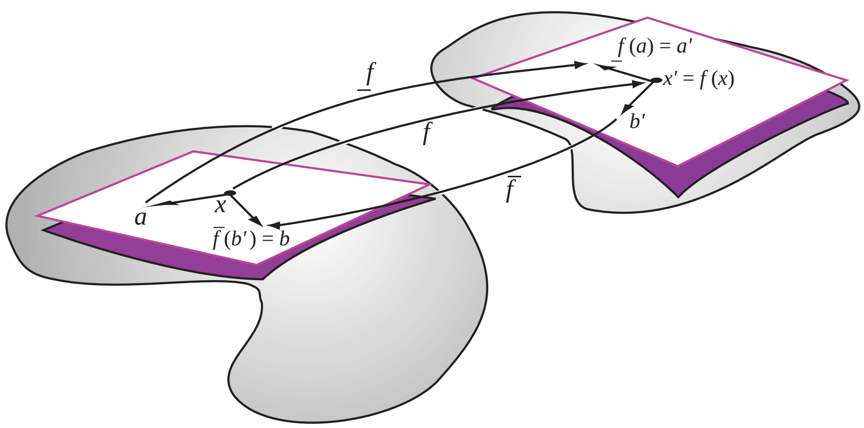

Fig. 1. A differentiable transformation f of points induces a differential $\underline{f}$ and its adjoint $\bar{f}$ transforming tangent vectors between the points.

To summarize, we have two kinds of *induced transformations* for multivector fields: the *differential* $\underline{f}$ and the *adjoint* $\bar{f}$, as shown in Fig. 1. Multivector fields transformed by $\underline{f}$ are commonly said to be *contravariant* or to "transform like a vector," while fields transformed by $\bar{f}$ are said to be *covariant* or to "transform like a *covector*." The term "vector" is thus associated with the differential while "covector" is associated with the adjoint. This linking of the vector concept to a transformation law is axiomatic in ordinary tensor calculus, because vectors are defined in terms of coordinates. In geometric calculus, however, the two concepts are separate, because the algebraic concept of vector is determined by the axioms of geometric algebra without reference to any coordinates or transformations.

As we have just seen, one way to assign a transformation law to a vector or a vector field is by relating it to differentiation. Another way is by the rule of *direct substitution:* A field $F = F(x)$ is transformed to $F(f^{-1}(x')) = F'(x')$ or, equivalently, to

$$F'(x') \equiv F'(f(x)) = F(x)\,. \tag{73}$$

Thus, the values of the field are unchanged — although they are associated with different points by changing the functional form of the field.

Directional derivatives of the two different functions in (73) are related by the *chain rule:*

$$a \cdot \nabla F = a \cdot \nabla_x F'(f(x)) = (a \cdot \nabla_x f(x)) \cdot \nabla_{x'} F'(x') = \underline{f}(a) \cdot \nabla' F' = a' \cdot \nabla' F'\,. \tag{74}$$

The chain rule is more simply expressed as an operator identity

$$a \cdot \nabla = a \cdot \bar{f}(\nabla') = \underline{f}(a) \cdot \nabla' = a' \cdot \nabla'\,. \tag{75}$$

Differentiation with ∂_a yields the general transformation law for the vector derivative:

$$\nabla = \bar{f}(\nabla') \qquad \text{or} \qquad \nabla' = \bar{f}^{\,-1}(\nabla)\,. \tag{76}$$

This is the *chain rule for the vector derivative*, the most basic form of the chain rule for differentiation on vector manifolds. All properties of induced transformations are essentially implications of this rule, including the transformation law for the differential, as (75) shows.

Sometimes it is convenient to use a subscript notation for the differential:

$$f_a \equiv \underline{f}(a) = a \cdot \nabla f. \tag{77}$$

Then the *second differential* of the transformation (63) can be written

$$f_{ab} \equiv b \cdot \grave{\nabla} a \cdot \nabla \grave{f} = b \cdot \nabla[\underline{f}(a)] - [\underline{f}(b \cdot \nabla a)]\,, \tag{78}$$

and we note that

$$f_{ab} = f_{ba}. \tag{79}$$

This symmetry is equivalent to the fact that the adjoint function has vanishing curl. Thus, using $\nabla = \partial_a a \cdot \nabla$ we prove

$$\grave{\nabla} \wedge \grave{\bar{f}}(a') = \partial_b \wedge \bar{f}_b(a') = \partial_b \wedge \partial_c f_{cb} \cdot a' = \nabla \wedge \nabla f \cdot a' = 0\,. \tag{80}$$

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

$$\nabla \wedge a = \nabla \wedge \bar{f}(a') = \bar{f}(\nabla' \wedge a'). \tag{81}$$

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

$$\bar{f}_b(A' \wedge B') = \bar{f}_b(A') \wedge \bar{f}(B') + \bar{f}(A') \wedge \bar{f}_b(B'). \tag{82}$$

Therefore, it follows from (80) that the curl of the adjoint outermorphism vanishes, and (81) generalizes to

$$\nabla \wedge A = \bar{f}(\nabla' \wedge A') \qquad \text{or} \qquad \nabla' \wedge A' = \bar{f}^{-1}(\nabla \wedge A), \tag{83}$$

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

The transformation rule for the divergence is more complex, but it can be derived from that of the curl by exploiting the duality of inner and outer products

$$a \wedge (Ai) = (a \cdot A)i \tag{84}$$

and the transformation law (58) relating them. Thus,

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

Then, using (83) and (66) we obtain

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

where $A' = \underline{f}(A)$. For the divergence, therefore, we have the transformation rule

$$\nabla' \cdot A' = \nabla' \cdot \underline{f}(A) = J_f{}^{-1}\underline{f}[\nabla \cdot (J_f A)] = \underline{f}[\nabla \cdot A + (\nabla \ln J_f) \cdot A], \tag{85}$$

This formula can be separated into two parts:

$$\grave{\nabla}' \cdot \grave{\underline{f}}(A) = \underline{f}[(\nabla \ln J_f) \cdot A] = (\nabla' \ln J_f) \cdot \underline{f}(A), \tag{86}$$

$$\grave{\nabla}' \cdot \underline{f}(\grave{A}) = \underline{f}(\nabla \cdot A). \tag{87}$$

The whole can be recovered from the parts by using the following generalization of (75) (which can also be derived from (58)):

$$\underline{f}(A) \cdot \nabla' = \underline{f}(A \cdot \nabla). \tag{88}$$

IV. Induced Geometry on Flat Spacetime

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 relativity theories the spacetime of real physical entities is a 4-dimensional continuum modeled mathematically by a 4D differentiable manifold $\mathcal{M}^4$. The standard formulation of *general relativity* (GR) employs a *curved space model* of spacetime, which places points and vector fields in different spaces. In contrast, our *flat space model* of spacetime identifies the spacetime manifold $\mathcal{M}^4$ with the Minkowski vector space of special relativity. A *spacetime map* describes a spatiotemporal partial ordering of physical events by representing the events as points (vectors) in $\mathcal{M}^4 = \{x\}$. All other properties of physical entities are represented by fields on spacetime with values in the tangent algebra $\mathcal{V}^4(x)$ and the geometric algebra $\mathcal{G}(\mathcal{V}^4)$ that it generates. For this reason it is worthwhile to maintain a distinction between $\mathcal{V}^4$ and $\mathcal{M}^4$ even though they are algebraically identical vector spaces.

To incorporate gravity into flat space theory we represent metrical relations among events by fields in $\mathcal{V}^4$ rather than intrinsic properties of $\mathcal{M}^4$ as in curved space theory. The following subsections explain how to do that in a systematic way that is easily compared to Einstein's curved space theory.

A. Displacement Gauge Invariance

A given ordering of events can be represented by a map in many different ways, just as the surface of the earth can be represented by Mercator projection, stereographic projection or many other equivalent maps. As the physical world is independent of the way we construct our maps, we seek a physical theory which is equally independent. The Cambridge group has formulated this idea as a new kind of gauge principle, which can be expressed as follows:

> **Displacement Gauge Principle (DGP):** *The equations of physics must be invariant under arbitrary smooth remappings of events onto spacetime.*

To give this principle a precise mathematical formulation, from here on we interpret the transformation (63) as a smooth remapping of flat spacetime $\mathcal{M}^4$ onto itself. It will be convenient to cast the direct substitution transformation (73) of a field $F = F(x)$ in the alternative form

$$F'(x) \equiv F(x') = F(f(x)) . \tag{89}$$

This simple transformation law, assumed to hold for all physical fields and field equations, is the mathematical formulation of the DGP. The descriptive term "displacement" is justified by the fact that (89) describes displacement of the field from point x to point x' without changing its values. The DGP should be recognized as a vast generalization of "*translational invariance*" in special relativity, so it has a comparable physical interpretation. Accordingly, the DGP can be interpreted as asserting that *"spacetime is globally homogeneous."* In other words, with respect to the equations of physics all spacetime points are equivalent. Thus, the DGP has implications for measurement, as it establishes a means for comparing field configurations at different places and times.

The Cambridge group had the brilliant insight that enforcement of the transformation law (89) requires introduction of a new physical field that can be identified with the gravitational field. Properties of this field, called the *gauge field* or *gauge tensor*, derive from the transformation laws of the preceding section. The most essential step is defining a gauge invariant derivative. To that end, consider a gauge invariant scalar field $\phi = \phi(x)$ transformed by (89); according to (76) the transformation of its gradient is given by

$$\nabla\phi'(x) = \nabla\phi(f(x)) = \bar{f}\,[\nabla'\phi(x')]\,. \tag{90}$$

To make this equation gauge invariant, we introduce an invertible *tensor field* $\bar{h}$ defined on covectors so that

$$\bar{h}'[(\nabla\phi'(x)] = \bar{h}[(\nabla'\phi(x')]\,. \tag{91}$$

Comparison with (90) shows that $\bar{h}$ must be a linear operator on covectors with the transformation law

$$\bar{h}' = \bar{h}\bar{f}^{-1}\,. \tag{92}$$

More explicitly, when operating on a covector $b = \nabla\phi$ the rule is

$$\bar{h}'(b;x) = \bar{h}(\bar{f}^{-1}(b;x);x')\,. \tag{93}$$

We usually suppress the position dependence and write

$$\bar{h}'(b) = \bar{h}(\bar{f}^{-1}(b))\,. \tag{94}$$

Applying this rule to the vector derivative ∇ with its transformation rule (76), we can define a *position gauge invariant derivative* by

$$\bar{\nabla} \equiv \bar{h}(\nabla)\,. \tag{95}$$

The symbol $\bar{\nabla}$ is a convenient *abuse of notation* to remind us that the linear operator $\bar{h}$ is involved in its definition.

From the operator $\bar{\nabla}$ we obtain a *position gauge invariant directional derivative*

$$a\cdot\bar{\nabla} = a\cdot\bar{h}(\nabla) = (\underline{h}a)\cdot\nabla\,, \tag{96}$$

where $\underline{h}$ is the adjoint of $\bar{h}$ and a is a "free vector," which is to say that it can be regarded as constant or as a vector transforming by the direct substitution rule (89). This is clarified by a specific example.

Let $x = x(\tau)$ be a timelike curve representing a *particle history.* According to (64), the diffeomorphism (63) induces the transformation

$$\dot{x} = \frac{dx}{d\tau} \qquad \to \qquad \dot{x}' = \underline{f}(\dot{x})\,. \tag{97}$$

Thus the description of particle velocity by $\dot{x}$ is "contravariant" under spacetime diffeomorphisms. Comparing (62) with (96), it is evident that an "invariant" *velocity* vector $v = v(x(\tau))$ can be introduced by writing

$$\dot{x} = \underline{h}(v) \,, \tag{98}$$

where $\underline{h}$ obeys the adjoint of the "gauge transformation" rule (92), that is

$$\underline{f}^{-1}\,\underline{h}' = \underline{h} \qquad \text{or} \qquad \underline{h}' = \underline{f}\,\underline{h} \,. \tag{99}$$

Accordingly, (97) implies that

$$\dot{x}' = \underline{h}'(v) \,, \tag{100}$$

where v has the same value as in (98), but it is taken as a function of $x'(\tau)$ instead of $x(\tau)$. To distinguish the $\dot{x}$ from the invariant velocity

$$v = \underline{h}^{-1}(\dot{x}) = \underline{h}'^{-1}(\dot{x}') \,, \tag{101}$$

one could refer to $\dot{x}$ as the *map velocity*. Otherwise the term "*velocity*" designates v. The invariant normalization

$$v^2 = 1 \tag{102}$$

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

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

$$\frac{d}{d\tau} = \dot{x} \cdot \bigtriangledown = \underline{h}(v) \cdot \bigtriangledown = v \cdot \bar{h}(\bigtriangledown) = v \cdot \bar{\bigtriangledown} \,, \tag{103}$$

so that

$$\frac{dF}{d\tau} = v \cdot \bar{\bigtriangledown} F \,. \tag{104}$$

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

$$\frac{dx}{d\tau} = v \cdot \bar{\bigtriangledown} x = (\underline{h} v) \cdot \bigtriangledown x = \underline{h} v \,, \tag{105}$$

recovering (98).

Since $\underline{h}$ and $\bar{h}$ are mutually determined, we can refer to either as the "gauge tensor" or, simply, the "gauge" on spacetime. Actually, the term "gauge" is more appropriate here than elsewhere in physics, because $\underline{h}$ does indeed determine the "gauging" of a metric on spacetime. To see that, use (97) in (101) to derive the following expression for the *invariant line element* on a timelike particle history:

$$d\tau^2 = [\, \underline{h}^{-1}(dx)\,]^2 = dx \cdot \underline{g}(dx) \tag{106}$$

where

$$\underline{g} = \bar{h}^{-1}\underline{h}^{-1} \tag{107}$$

is a symmetric *metric tensor*. This formulation suggests that *gauge* is a more fundamental geometric entity than metric, and that view is confirmed by developments below. Einstein has taught us to interpret the metric tensor physically as a gravitational potential, making it easy and natural to transfer this interpretation to the gauge tensor. Some readers will recognize $\underline{h}$ as equivalent to a "tetrad field," which has been proposed before to represent gravitational fields on flat spacetime [16]. However, the spacetime calculus makes all the difference in turning the tetrad into a practical tool.

To verify that (106) is equivalent to the standard invariant line element in general relativity, coordinates must be introduced. Let $x = x(x^0, x^1, x^2, x^3)$ be a parametrization of the points, in some spacetime region, by an arbitrary set of coordinates $\{x^\mu\}$. Partial derivatives then give tangent vectors to the coordinate curves $\partial_\mu x$, which, in direct analogy to (98) and (101), determine a set of displacement gauge invariant vector fields $\{g_\mu\}$ according to the equation

$$e_\mu \equiv \partial_\mu x = \frac{\partial x}{\partial x^\mu} = \underline{h}(g_\mu)\,, \tag{108}$$

or, equivalently, by

$$g_\mu = \underline{h}^{-1}(e_\mu)\,. \tag{109}$$

The components for this coordinate system are then given by

$$g_{\mu\nu} = g_\mu \cdot g_\nu = e_\mu \cdot \underline{g}(e_\nu)\,. \tag{110}$$

Therefore, with $dx = dx^\mu e_\mu$, the line element (106) can be put in the form

$$d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu\,, \tag{111}$$

which is familiar from the tensor formulation of GR.

To complete the introduction of coordinates into the flat space gauge theory, we introduce coordinate functions $x^\mu = x^\mu(x)$ and the *coordinate frame* vectors

$$e^\mu = \nabla x^\mu \tag{112}$$

reciprocal to the frame $\{e_\mu\}$ defined by (108); whence

$$e_\mu \cdot e^\nu = e_\mu \cdot \nabla x^\nu = \partial_\mu x^\nu = \delta_\mu^\nu\,. \tag{113}$$

The displacement gauge invariant frame $\{g^\mu\}$ reciprocal to (109) is then given by

$$g^\mu = \bar{h}(e^\mu) = \bar{\nabla} x^\mu\,. \tag{114}$$

So the gauge invariant derivative is given by

$$\bar{\nabla} = \bar{h}(\nabla) = g^\mu \partial_\mu \,. \tag{115}$$

The flat space and curved space theories can now be compared through their use of coordinates. The vector fields g_μ encode all the information in the metric tensor $g_{\mu\nu} = g_\mu \cdot g_\nu$ of GR, as is evident in (110). However, equations (111) and (114) decouple coordinate dependence of the metric, expressed by $e_\mu = \partial_\mu x$, from the physical dependence, expressed by the gauge tensor $\bar{h}$ or its adjoint $\underline{h}$. 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 the Displacement Gauge Principle has physical consequences, whereas Einstein's general relativity principle does not.

The precise theoretical status of Einstein's *general relativity principle* (GRP), also known as the *general covariance principle*, has been a matter of perennial confusion and dispute that some say originated in confusion on Einstein's part [17]. One of Einstein's main motivations in formulating it was epistemological: to extend the *special relativity principle*, requiring physical equivalence of all inertial frames, to physical equivalence of arbitrary reference frames. He formulated the GRP as requiring covariance of the equations of physics under the group of arbitrary coordinate transformations (also known as *general covariance*). While acknowledging criticisms that the requirement of general covariance is devoid of physical content, Einstein continued to believe in the GRP as a cornerstone of his gravitation theory.

Gauge theory throws new light on the GRP, which may be regarded as vindicating Einstein's obstinate stance. The problem with Einstein's GRP is that it is not a true symmetry principle [17]. For a transformation group to be a physical symmetry group, there must be a well defined "geometric object" that the group leaves invariant. Thus the "*displacement group*" of the DGP is a symmetry group, because it leaves the flat spacetime background invariant. That is why it can have the physical consequences described above. There is no comparable symmetry group for curved spacetime, because each mapping produces a new spacetime, so there is no geometric object to be left invariant. Moreover, general covariance does not discriminate between passive coordinate transformations and active physical transformations. Those distinctions are generally made on an *ad hoc* basis in applications. The intuitive idea of physically equivalent situations was surely at the back of Einstein's mind when he insisted on heuristic significance of the GPR over the objections of Kretschmann and others on its lack of physical content. I submit that the DGP provides, at last, a precise mathematical formulation for the kind of GRP that Einstein was looking for. Once again I am impressed by Einstein's profound physical insight, which served him so well in assessing the significance of mathematical equations in physics. Of course, his conclusions depended critically on the mathematics at his disposal, and displacement gauge theory was not an option available to him.

B. Rotation Gauge Covariance

Gauge theory gravity requires one other gauge principle, which we formulate as follows:

> **Rotation Gauge Principle (RGP):** *The equations of physics must be covariant under local Lorentz rotations.*

Generalizing the treatment in Section IIIB, we characterize a local Lorentz rotation by a position dependent rotor field $L = L(x)$ with $L\widetilde{L} = 1$. This enables us to define two kinds of covariant fields, a multivector field $M = M(x)$ and a spinor field $\psi = \psi(x)$, with respective transformation laws:

$$\underline{L} : M \quad \to \quad M' = \underline{L}\, M \equiv LM\widetilde{L}\,, \tag{116}$$

$$L : \psi \quad \to \quad \psi' = L\psi\,. \tag{117}$$

Spinors and multivectors are related by the fact that the spinor field determines a frame of multivector observables $\{\psi\gamma_\mu\widetilde{\psi}\}$, where $\{\gamma_\mu\}$ is the standard orthonormal frame of constant vectors introduced in equation (1). Frame independence of the gauge tensor is ensured by (116), which implies that the gauge tensor satisfies the operator transformation law

$$\underline{L} : \bar{h} \quad \to \quad \bar{h}' = \underline{L}\, \bar{h}\,. \tag{118}$$

From (114) it follows that this transformation leaves the metric tensor $g^{\mu\nu} = g^\mu \cdot g^\nu$ invariant.

To construct field equations that satisfy the RGP, we define a "gauge covariant derivative" or *coderivative* operator D as follows. With respect to a coordinate frame $\{g_\mu = \underline{h}^{-1}(e_\mu)\}$, components $D_\mu = g_\mu \cdot D$ of the coderivative are defined for spinors and multivectors respectively by

$$D_\mu\psi = (\partial_\mu + \tfrac{1}{2}\omega_\mu)\psi\,, \tag{119}$$

$$D_\mu M = \partial_\mu M + \omega_\mu \times M\,, \tag{120}$$

where the *connexion* for the derivative

$$\omega_\mu = \omega(g_\mu) \tag{121}$$

is a bivector-valued tensor evaluated on the frame $\{g_\mu\}$. From Section IA, we know that the bivector property of ω_μ implies that $\omega_\mu \times M$ is "grade preserving." Hence D_μ is a "scalar differential operator" in the sense that grade($D_\mu M$) =grade(M). Note that D_μ is not grade-preserving on the spinor in (119).

To ensure the covariant transformations for the derivatives

$$L : D_\mu\psi \quad \to \quad L(D_\mu\psi) = D'_\mu\psi' = (\partial_\mu + \tfrac{1}{2}\omega'_\mu)\psi'\,, \tag{122}$$

$$\overline{L}\,:D_\mu M \quad\rightarrow\quad \overline{L}\,(D_\mu M) = D'_\mu M' = \partial_\mu M' + \omega'_\mu \times M' \,, \tag{123}$$

the connexion must obey the familiar gauge theory transformation law

$$\omega_\mu = \quad\rightarrow\quad \omega'_\mu = L\omega_\mu \widetilde{L} - 2(\partial_\mu L)\widetilde{L} \,, \tag{124}$$

where it should be noted that the operator $\partial_\mu = g_\mu \cdot \overline{\overline{\nabla}}$ is rotation gauge invariant. The full *vector coderivative* D can now be defined by the operator equation

$$D = g^\mu D_\mu \,, \tag{125}$$

for which gauge covariance follows from (118) and the definition of D_μ. Since D_μ preserves the grade of M in (120), the coderivative is a vector differential operator, so we can decompose it in the same way that we decomposed the vector derivative ∇ into divergence and curl in equation (31). Thus, for an arbitrary covariant multivector field $M = M(x)$, we can write

$$DM = D \cdot M + D \wedge M \,, \tag{126}$$

where $D \cdot M$ and $D \wedge M$ are respectively the *codivergence* and the *cocurl*.

Combining (118) with (92), we see that the most general transformation of the gauge tensor is given by

$$\bar{h}' = \overline{L}\,\bar{h}\bar{f}^{-1}, \quad \underline{h}' = \underline{f}\,\underline{h}\underline{L} \,. \tag{127}$$

In other words, every gauge transformation on spacetime is a combination of displacement and local rotation. This operator equation enables us to turn any multivector field on spacetime into a gauge covariant field, as shown explicitly in the next section for the electromagnetic field. Finally, it should be noted that all the above rotation covariant quantities, including the connexion ω_μ, are displacement gauge invariant as well.

In special relativity, Lorentz transformations are passive rotations expressing equivalence of physics with respect to different inertial reference frames. Here, however, covariance under active rotations expresses local physical equivalence of different directions in spacetime. In other words, *the rotation gauge principle asserts that spacetime is locally isotropic.* Thus, "passive equivalence" is an equivalence of observers, while "active equivalence" is an equivalence of states. This distinction generalizes to the physical interpretation of any symmetry group principle: *Active transformations relate equivalent physical states; passive transformations relate equivalent observers.*

To establish the connection of GTG to GR, we need to relate the coderivative to the usual covariant derivative in GR. Like the directional derivative $\partial_\mu = g_\mu \cdot \overline{\overline{\nabla}}$, the *directional coderivative* $D_\mu = g_\mu \cdot D$ is a "scalar differential operator" that maps vectors into vectors. Accordingly, we can write

$$D_\mu g_\nu = L^\alpha_{\mu\nu} g_\alpha \,, \tag{128}$$

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 $g^\alpha \cdot g_\nu = \delta^\alpha_\nu$, we find the complementary equation

$$D_\mu g^\alpha = -L^{\alpha}_{\mu\nu} g^\nu \,. \tag{129}$$

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

$$D_\mu a = (\partial_\mu a^\alpha + a^\nu L^{\alpha}_{\mu\nu}) g_\alpha \,. \tag{130}$$

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

Further properties of the connection are obtained by contracting (129) with g^μ to obtain

$$D \wedge g^\alpha = g^\nu \wedge g^\mu L^{\alpha}_{\mu\nu} \,. \tag{131}$$

This bivector-valued quantity is called *torsion*. In the Riemannian geometry of GR torsion vanishes, so we restrict our attention to that case. Considering the antisymmetry of the outer product on the right side of (131), we see that the torsion vanishes if and only if

$$L^{\mu}_{\alpha\beta} = L^{\mu}_{\beta\alpha} \,. \tag{132}$$

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

$$\partial_\mu g_{\alpha\beta} = g_{\alpha\nu} L^{\nu}_{\mu\beta} + g_{\beta\nu} L^{\nu}_{\mu\alpha} \,. \tag{133}$$

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

$$L^{\mu}_{\alpha\beta} = \tfrac{1}{2} g^{\mu\nu} (\partial_\alpha g_{\beta\nu} + \partial_\beta g_{\alpha\nu} - \partial_\nu g_{\alpha\beta}) \,. \tag{134}$$

This is the classical *Christoffel formula* for a *Riemannian connexion*, so the desired relation of the coderivative to the covariant derivative in GR is established.

C. Curvature and Coderivative

In standard tensor analysis the curvature tensor is obtained from the commutator of covariant derivatives. Likewise, we obtain it here from the commutator of coderivatives. Differentiating (129) we get

$$[D_\mu, D_\nu] g^\alpha = R^{\alpha}_{\mu\nu\beta} g^\beta \,, \tag{135}$$

where the operator commutator has the usual definition

$$[D_\mu, D_\nu] \equiv D_\mu D_\nu - D_\nu D_\mu \,, \tag{136}$$

and

$$R^\alpha_{\mu\nu\beta} = \partial_\mu L^\alpha_{\nu\beta} - \partial_\nu L^\alpha_{\mu\beta} + L^\alpha_{\nu\sigma} L^\sigma_{\mu\beta} - L^\alpha_{\mu\sigma} L^\sigma_{\nu\beta} \,, \tag{137}$$

is the usual tensor expression for the *Riemannian curvature* of the manifold. This suffices to establish mathematical equivalence of our flat space gauge theory to the standard tensor formulation of Riemannian geometry and hence to GR. Now we can confidently turn to the full gauge theory treatment of curvature and gravitation to see what advantages it has over standard GR.

The commutator of coderivatives defined by (121) gives us

$$[D_\mu, D_\nu] M = R(g_\mu \wedge g_\nu) \times M \,, \tag{138}$$

where

$$\begin{aligned} R(g_\mu \wedge g_\nu) &= \partial_\mu \omega_\nu - \partial_\nu \omega_\mu + \omega_\mu \times \omega_\nu \\ &= D_\mu \omega_\nu - D_\nu \omega_\mu - \omega_\mu \times \omega_\nu \equiv \omega_{\mu\nu} \end{aligned} \tag{139}$$

is the *curvature tensor* expressed as a bivector-valued function of a bivector variable. This is the fundamental form for curvature in GTG. It follows from (123) that the curvature bivector has the covariance property

$$\bar{L} \; : R(g_\mu \wedge g_\nu) \quad \to \quad R'(g'_\mu \wedge g'_\nu) = \bar{L}\, R(\underline{L}\,(g_\mu \wedge g_\nu)), \tag{140}$$

where $\bar{L} = \underline{L}^{-1}$.

For the gravity field variables g^μ, equations (138) and (135) give us

$$[D_\mu, D_\nu] g^\alpha = \omega_{\mu\nu} \cdot g^\alpha = R^\alpha_{\mu\nu\beta}\, g^\beta \,; \tag{141}$$

whence

$$R^\alpha_{\mu\nu\beta} = \omega_{\mu\nu} \cdot (g^\alpha \wedge g_\beta). \tag{142}$$

Coordinate frame fields $\{g_\mu\}$ and $\{g^\mu\}$ have been used to introduce the coderivative and curvature in order to make connection with the standard tensor formalism as direct as possible. However, the *vector coderivative* and the *curvature bivector* can be defined and all their properties can be derived in a completely coordinate-free way. Derivations are given in Appendix B, and results are summarized in Table 1.

Curvature

$R(a \wedge b) \equiv a \cdot \overline{\overline{\nabla}}\omega(b) - b \cdot \overline{\overline{\nabla}}\omega(a) + \omega(a) \times \omega(b) - \omega([a,b])$

$[a \cdot \dot{D}, b \cdot \dot{D}]\dot{M} = R(a \wedge b) \times M = [a \cdot D, b \cdot D]M - \omega([a,b])$

$[a,b] \equiv a \cdot \overline{\overline{\nabla}} b - b \cdot \overline{\overline{\nabla}} a = a \cdot Db - b \cdot Da$

$R(g_\mu \wedge g_\nu) = \partial_\mu \omega_\nu - \partial_\nu \omega_\mu + \omega_\mu \times \omega_\nu$

$[D_\mu, D_\nu] g^\alpha = R(g_\mu \wedge g_\nu) \times g^\alpha = R^\alpha_{\mu\nu\beta}\, g^\beta$

$R(a \wedge b) = a^\mu b^\nu R(g_\mu \wedge g_\nu)$ $\qquad A \cdot R(B) = B \cdot R(A)$

Curvature Contractions and Coderivative Identities

$R(a) \equiv \partial_b \cdot R(b \wedge a) = \partial_b R(b \wedge a)$ $\qquad \partial_a \wedge R(a \wedge b) = 0$

$R \equiv \partial_a \cdot R(a) = \partial_a R(a)$ $\qquad \partial_a \wedge R(a) = 0$

$D \wedge D\, a = (D \wedge D) \cdot a = R(a)$ $\qquad (D \wedge D) \cdot M = D \cdot (D \cdot M) = 0$

$D \wedge D \wedge M = 0$ $\qquad D \wedge D\, M = (D \wedge D) \times M$

Bianchi Identity and its Contractions

$\dot{D} \wedge \dot{R}(a \wedge b) = 0$ $\qquad a \cdot \dot{D}\dot{R}(b \wedge c) + b \cdot \dot{D}\dot{R}(c \wedge a) + c \cdot \dot{D}\dot{R}(a \wedge b) = 0$

$\dot{R}(\dot{D} \wedge b) = \dot{D} \wedge \dot{R}(b)$ $\qquad \dot{R}(\dot{D}) = \frac{1}{2} DR$

$G(a) \equiv R(a) - \frac{1}{2} aR$ $\qquad \dot{G}(\dot{D}) = 0$

Table 1. Coderivative and Curvature. All derivatives are rotation gauge covariant if the fields are. However, coderivatives are well-defined even for quantities that are not covariant. The fields a and b are vector-valued, while A and B are bivectors, and the multivector M must have $\text{grade}(M) \geq 2$ in the double codivergence identity. Note that accents are employed in some equations to indicate which quantities are differentiated.

V. Field Equations and Particle Motion

With the differential geometry of gauge fields well in hand, we are ready to apply it to gravity theory.

A. Einstein's Equation

Einstein's gravitational field equation can be created from contractions of the curvature tensor in the usual way. Contraction of the curvature bivector (139) gives the usual *Ricci tensor* expressed as a vector-valued function of a vector variable.

$$R(g_\nu) = g^\mu \cdot R(g_\mu \wedge g_\nu)\,, \qquad (143)$$

With help from the identity (13), a second contraction gives the usual *scalar curvature*

$$R \equiv g^{\nu} \cdot R(g_{\nu}) = (g^{\nu} \wedge g^{\mu}) \cdot R(g_{\mu} \wedge g_{\nu}) \,. \tag{144}$$

The gravitational field equation can thus be written in the form

$$G(g_{\nu}) \equiv R(g_{\nu}) - \tfrac{1}{2} g_{\nu} \, R = \kappa T(g_{\nu}) \,, \tag{145}$$

where $\kappa = 8\pi$ in "natural units" with the gravitational constant set equal to unity. The vector-valued function of a vector variable $G(a)$ is *Einstein's tensor*, and its source $T(a)$ is the total *energy-momentum tensor* for material particles and non-gravitational fields. Of course, decomposition into tensor components $G_{\mu\nu} = g_{\mu} \cdot G(g_{\nu})$ and $T_{\mu\nu} = g_{\mu} \cdot T(g_{\nu})$ yields Einstein's equation in its standard tensor component form.

The spacetime algebra reviewed in Section I enables us to express Einstein's tensor $G^{\beta} = G(g^{\beta})$ in the new *unitary form*:

$$G^{\beta} = \tfrac{1}{2} (g^{\beta} \wedge g^{\mu} \wedge g^{\nu}) \cdot \omega_{\mu\nu} = \tfrac{1}{2} (g^{\beta} \wedge \omega_{\mu\nu}) \cdot (g^{\mu} \wedge g^{\nu}) \,. \tag{146}$$

The last equality is a consequence of the well-known symmetry of the curvature tensor,

$$(g_{\alpha} \wedge g_{\beta}) \cdot R(g_{\mu} \wedge g_{\nu}) = (g_{\mu} \wedge g_{\nu}) \cdot R(g_{\alpha} \wedge g_{\beta}) \,. \tag{147}$$

Identities (12) and (13) can be used to expand the inner product in (146) to get

$$G^{\beta} = g^{\mu} \cdot \omega_{\mu\nu} g^{\nu\beta} - \tfrac{1}{2} g^{\beta} (g^{\nu} \wedge g^{\mu}) \cdot \omega_{\mu\nu} \,, \tag{148}$$

which will be recognized as equivalent to the standard form (145) for the Einstein tensor. Thus, expansion of the inner product has split the "unitary" Einstein tensor into two parts. Let us refer to this as the *Ricci split* of the Einstein tensor.

As is well known, Einstein originally chose the particular combination of Ricci tensor and scalar curvature in (145) because, as shown in Appendix B and Table 1, its codivergence $\grave{G}(\grave{D})$ vanishes in consequence of the Bianchi identity. Using $D_{\mu}(g g^{\mu}) = 0$ from appendix A, we can write this version of the Bianchi identity in the form

$$D_{\mu}[g G(g^{\mu})] = g \grave{G}(\grave{D}) = 0 \,. \tag{149}$$

It follows, then, from Einstein's field equation (145) that

$$D_{\mu}[g T(g^{\mu})] = g \grave{T}(\grave{D}) = 0 \,. \tag{150}$$

Initially, Einstein wanted to interpret this as a general energy-momentum conservation law. However, that interpretation is not so straightforward, because a true conservation law requires vanishing of an ordinary divergence rather than a codivergence. Instead, we shall see in Section VC that (150) determines the equation of motion for matter. Then in Section VII we see that an alternative to the Ricci split is more suitable for deriving the general energy-momentum conservation law that Einstein wanted.

B. Electrodynamics with Gravity

GTG differs from GR in providing explicit equations for the influence of gravity on physical fields. This is best illustrated by electrodynamics. Maxwell's equation $\nabla F = J$ with $F = \nabla \wedge A$ is not gravitation gauge covariant. To make the vector field A gauge covariant and incorporate the effect of gravity, we simply write [15]

$$\overline{A} = \bar{h}(A)\,, \tag{151}$$

and gauge covariance is assured by (127). In fact, the influence of gravity on the electromagnetic field is fully characterized by the gauge tensor $\bar{h}$.

Now applying (240) from Appendix A to (151), we immediately obtain a gauge covariant expression for the electromagnetic field strength:

$$\overline{F} = \bar{h}(F) = D \wedge \overline{A} = \bar{h}(\nabla \wedge A)\,. \tag{152}$$

Similarly, applying the cocurl again and using Table 1, we obtain

$$D \wedge \overline{F} = D \wedge D \wedge \overline{A} = \bar{h}(\nabla \wedge \nabla \wedge A) = 0\,. \tag{153}$$

Thus, Maxwell's equation $\nabla \wedge F = 0$ applies irrespective of the gravitational field. Now it is obvious that the gauge covariant form of Maxwell's equation must be

$$D\,\overline{F} = D \cdot \overline{F} + D \wedge \overline{F} = \overline{J}\,. \tag{154}$$

From the vector part of this equation, we have, using the double codivergence identity in Table 1,

$$D \cdot (D \cdot \overline{F}) = 0 = D \cdot \overline{J}\,, \tag{155}$$

A little more analysis is needed to relate the right side of this equation to the standard charge current conservation law.

Explicit dependence of the codivergence $D \cdot \overline{F}$ on the gauge tensor can be derived from (152) by exploiting duality. The derivation is essentially the same as that for (85) and yields the result

$$D \cdot \overline{F} = h\,\underline{h}^{-1}[\nabla \cdot (h^{-1}\,\underline{h}(\overline{F}\,))]\,, \tag{156}$$

where $h \equiv det(\underline{h})$. For the vector field $\overline{J}$ the result is a scalar, and we obtain the simpler equation

$$D \cdot \overline{J} = h \nabla \cdot (h^{-1}\,\underline{h}(\overline{J}\,))\,, \tag{157}$$

so comparison with (155) identifies $J = h^{-1}\underline{h}(\overline{J}\,)$ with the conserved charge current. Combining this with (156), we can write $D \cdot \overline{F} = \overline{J}$ in the equivalent form

$$\nabla \cdot (h^{-1}\underline{h}\bar{h}(F)) = J\,. \tag{158}$$

This enables us to interpret the operator $h^{-1}\underline{h}\bar{h} = h^{-1}\,\underline{g}^{-1}$ as a kind of dielectric tensor or *generalized index of refraction* for the *"gravitational medium"* [18]. Of course, equation (158) can be used to compute refraction of light by the sun, with a result in agreement with alternative GR calculations.

The standard Maxwell energy-momentum tensor (as given in GA2) is easily generalized to the gauge covariant form:

$$T_{EM}(a) = -\tfrac{1}{2}\overline{F}\,a\overline{F} \tag{159}$$

Using (153) to compute its coderivative , we obtain

$$\grave{T}_{EM}(\grave{D}) = -\tfrac{1}{2}(\grave{\overline{F}}\,\grave{D}\overline{F} + \overline{F}\,D\overline{F}\,) = \tfrac{1}{2}(\overline{J}\,\overline{F} - \overline{F}\,\overline{J}\,)\,. \tag{160}$$

In other words,

$$D_\mu[gT_{EM}(g^\mu)] = g\grave{T}_{EM}(\grave{D}) = g\overline{F} \cdot \overline{J}\,. \tag{161}$$

One way to assure mutual consistency of Einstein's field equation with Maxwell's equation is to derive them from a common Lagrangian. This has been done by the Cambridge group in a novel way with GC, and the method has been extended to spinor fields with provocative new results. The reader is referred to the literature for details [2].

C. Equations of Motion from Einstein's Equation

In classical electrodynamics Maxwell's field equation and Lorentz's equation of motion for a charged particle must be postulated separately. It came as quite a surprise, therefore, to discover that in GR the equation of motion for a material particle can be derived from Einstein's field equation, provided it is assumed to hold everywhere, including at any singularities in the gravitational field. This discovery was made independently by several people, though most of the credit is unfairly given to Einstein [19]. Here we sketch essential elements in one especially simple approach to deriving equations of motion to illustrate advantages of using GA. The form of the equation of motion is determined by the form of the energy-momentum tensor. The idea is to describe a localized material system as a point particle by a multipole expansion of the mass distribution. Then the problem is to reformulate the result as a differential equation for the particle.

We consider only the simplest case of a structureless point particle. For a fluid of non-interacting material particles, the energy-momentum tensor can be put in the form

$$T_M(a) = m\rho(a \cdot v)v, \tag{162}$$

where ρ is the *proper particle density*, v is the *proper particle velocity* and m is the *mass per particle*. Its codivergence is

$$\grave{T}_M(\grave{D}) = \grave{T}_M(\grave{\nabla}) + \omega_\mu \cdot T(g^\mu) = mvD \cdot (\rho v) + m\rho v \cdot Dv. \tag{163}$$

As shown in (157) for a charge current, the codivergence $D \cdot (\rho v)$ can be expressed as a true divergence, so its vanishing expresses particle conservation, which we assume holds here. For a single particle, ρ is a δ-function on the particle history, and it can be eliminated by integration, but we don't need that step here.

For a charged particle the total energy-momentum tensor is

$$T(a) = T_M(a) + T_{EM}(a) = m\rho a \cdot vv - \tfrac{1}{2}\overline{F}\, a\overline{F}\,. \tag{164}$$

According to (161) and (163), its codivergence is

$$D_\mu[g(T_M(g^\mu) + T_{EM}(g^\mu)] = g\rho mv \cdot Dv + g\overline{J}\ \cdot \overline{F}\,. \tag{165}$$

For a particle with charge q the current density is $\overline{J}\ = q\rho v$, so the vanishing of (165) gives us

$$mv \cdot Dv - q\overline{F}\ \cdot v = 0. \tag{166}$$

This has the desired form for the equation of motion of a point charge. For $\overline{F}\ = 0$ it reduces to the geodesic equation, as described in the next Section.

The electromagnetic field in (166) can be expressed as a sum $\overline{F}\ = \overline{F}_{self} + \overline{F}_{ext}$, where $\overline{F}_{self}$ is the particle's own field and $\overline{F}_{ext}$ is the field of other particles. Evaluated on the particle history, $\overline{F}_{self}$ describes reaction of the particle to radiation it emits. This process has been much studied in the literature [19], but the analysis is too complicated to discuss here. If we neglect radiation, we can write $\overline{F}\ = \overline{F}_{ext}$ in (166), and $q\overline{F}\ \cdot v$ is the usual Lorentz force on a "test charge."

D. Particle Motion and Parallel Transfer

The equation for a particle history $x = x(\tau)$ generated by a timelike velocity $v = v(x(\tau))$ in the presence of an "ambient" gauge tensor $\underline{h}$ is $\dot{x} = \underline{h}(v)$. For any covariant multivector field $M = M(x(\tau))$, eqn. (120) gives us the directional *coderivative*

$$v \cdot DM = \frac{d}{d\tau}M + \omega(v) \times M\,, \tag{167}$$

where the gauge invariant proper time derivative is given by (104). Obviously, all this applies to arbitrary differentiable curves in spacetime if the requirement that v be timelike is dropped.

The equation for a *geodesic* can now be written

$$v \cdot Dv = \dot{v} + \omega(v) \cdot v = 0 \tag{168}$$

With $\omega(v)$ specified by the "ambient geometry," this equation can be integrated for $v = \underline{h}^{-1}(\dot{x})$ and then a second integration gives the geodesic curve $x(\tau)$. The solution is facilitated by noting that (168) implies

$$v \cdot \dot{v} = \tfrac{1}{2} \frac{dv^2}{d\tau} = v \cdot \omega(v) \cdot v = \omega(v) \cdot (v \wedge v) = 0\,. \tag{169}$$

Therefore v^2 is constant on the curve and, as noted in GA2, the value of v at any point on the curve can be obtained from a timelike reference value v_0 by a Lorentz rotation. Thus,

$$v = \overline{R}\, v_0 = R v_0 \widetilde{R}\,, \tag{170}$$

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

$$v \cdot DR = \frac{dR}{d\tau} + \tfrac{1}{2}\omega(v)R = 0\,. \tag{171}$$

This is equivalent to the spinor derivative (119), so the rotor R can be regarded as a special kind of spinor.

More generally, the *parallel transport* of any fixed multivector M_0 to a field $M = M(x(\tau))$ defined on the whole curve is given by

$$M = \overline{R}\, M_0 = R M_0 \widetilde{R}\,. \tag{172}$$

Of course, it satisfies the differential equation

$$v \cdot DM = \dot{M} + \omega(v) \times M = 0\,. \tag{173}$$

The formal similarity of this formulation for parallel transport to a directional gauge transformation is obvious, but the two should not be confused.

With the spinor coderivative defined by (171) (and not necessarily vanishing), we can rewrite (167) as an operator equation

$$v \cdot D = \overline{R}\, \frac{d}{d\tau}\, \underline{R}, \tag{174}$$

or equivalently, as

$$\underline{R} v \cdot D = v \cdot \overline{\overline{\nabla}} \underline{R} = \frac{d}{d\tau}\, \underline{R}\,. \tag{175}$$

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

According to (166), for a particle with unit mass and charge q in an electromagnetic field $\overline{F} = \bar{h}(F) = D \wedge \overline{A} = \bar{h}(\nabla \wedge A)$, the geodesic equation (168) generalizes to

$$\dot{v} = v \cdot \omega(v) + q\overline{F} \cdot v\,, \tag{176}$$

The first term on the right can be interpreted as a "gravitational force," though the standard GR interpretation regards it as part of "force-free" geodesic motion. From (246) and (242) in Appendix A, we see that $v \cdot (v \cdot H) = (v \wedge v) \cdot H = 0$, so

$$v \cdot \omega(v) = v \cdot H(v) = -v \cdot \bar{h}\,[\grave{\nabla} \wedge \grave{\bar{h}}^{-1}(v)]\,, \tag{177}$$

where the accent indicates which quantity is differentiated. Thus, the "gravitational force" is determined by the bivector-valued function $H(v)$, and we can write

$$\begin{aligned} v \cdot H(v) + q\bar{F} \cdot v &= \bar{h}[(\grave{\nabla} \wedge \grave{\bar{h}}^{-1}(v) + \nabla \wedge A)] \cdot \bar{h}^{-1}(\dot{x}) \\ &= \bar{h}[(\grave{\nabla} \wedge \grave{\bar{h}}^{-1}(v) + \nabla \wedge A) \cdot \dot{x}]. \end{aligned} \tag{178}$$

Finally, on expanding the last term and introducing the metric tensor $\underline{g}$ defined by (107), we can put the equation of motion (176) in the form

$$\frac{d}{d\tau}\,[\,\underline{g}(\dot{x}) + qA] = \grave{\nabla}[\,\tfrac{1}{2}\dot{x} \cdot \grave{\underline{g}}(\dot{x}) + q\grave{A} \cdot \dot{x}]\,. \tag{179}$$

Except for the presence of the metric tensor $\underline{g}$, this is identical to the equation of motion for a point charge in Special Relativity, so we can use familiar techniques from that domain to solve it (details in Section VII). Ignoring the vector potential and solving for $\ddot{x}$, it can be put in the form

$$\ddot{x} \,=\, \underline{g}^{-1}\,[\,\tfrac{1}{2}\grave{\nabla}\dot{x} \cdot \grave{\underline{g}}(\dot{x}) - \dot{\underline{g}}(\dot{x})]\,. \tag{180}$$

This is equivalent to the standard formulation for the geodesic equation in terms of the "Christoffel connection." Note how the metric tensor appears when the equation of motion is formulated in terms of a spacetime map rather than in covariant form.

E. Gravitational Precession and Dirac Equation

The solution of equation (171) gives more than the velocity. In perfect analogy with the treatment of electromagnetic spin precession in GA2, it 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 gravitation theory. To the particle worldline we attach 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 fixed orthonormal frame $\{\gamma_\mu\}$ is described by

$$e_\mu = R\,\gamma_\mu \widetilde{R}\,, \tag{181}$$

where $R = R(x(\tau))$ is a unimodular rotor. This has the same form as equation (24) for a mobile in flat spacetime because the background spacetime is flat.

Applying (167) to (181) we obtain the coderivative of the mobile:

$$v \cdot De_\mu = \dot{e}_\mu + \omega(v) \cdot e_\mu \,. \tag{182}$$

The coderivative here includes a gauge invariant description of gravitational forces on the mobile. In accordance with (21) and (23), effects of any nongravitational forces can be incorporated by writing

$$v \cdot De_\mu = \dot{e}_\mu + \omega(v) \cdot e_\mu = \Omega \cdot e_\mu \,, \tag{183}$$

where $\Omega = \Omega(x)$ is a bivector such as an external electromagnetic field acting on the mobile. The four equations (183) include the equation of motion

$$\frac{dv}{d\tau} = (\Omega - \omega(v)) \cdot v \tag{184}$$

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

$$\frac{dR}{d\tau} = \tfrac{1}{2}(\Omega - \omega(v))R \,. \tag{185}$$

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. This equation has been applied to gravitational precession in [20], so there is no need to elaborate here.

Our rotor treatment of classical gravitational precession generalizes directly to gravitational interactions in quantum mechanics. We saw in (40) that a real Dirac spinor field $\psi = \psi(x)$ determines an orthonormal frame of vector fields $e_\mu = e_\mu(x) = R\gamma_\mu \widetilde{R}$. Generalization of the real Dirac equation (38) to include gravitational interaction is obtained simply by replacing the partial derivative ∂_μ by the coderivative D_μ defined by equation (119). Thus, we obtain

$$g^\mu D_\mu \psi \gamma_2\gamma_1 \hbar = g^\mu(\partial_\mu + \tfrac{1}{2}\omega_\mu)\psi\gamma_2\gamma_1\hbar = e\overline{A}\,\psi + m\psi\gamma_0 \,, \tag{186}$$

where $\overline{A} = \bar{h}(A)$ is the gauge covariant vector potential. 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. Comparison of the spinor coderivative (119) with the rotor coderivative (185) tells us immediately that *gravitational effects on electron motion, including spin precession, are exactly the same as on classical rigid body motion.* The Cambridge group has studied solutions of the *real Dirac equation* (186) in the field of a black hole extensively [21].

VI. Solutions of Einstein's Equation

Though GC facilitates curvature calculations, the nonlinearity of Einstein's equation still makes it difficult to solve. Gauge theory gravity introduces simplifications both in calculation and physical interpretation by cleanly separating the functional form of the gauge tensor from arbitrary choices of coordinate system. Using results in the Appendices and Table 1, gravitational curvature can be calculated from a given gauge tensor by the straightforward sequence of steps:

$$\bar{h}(a) \quad \rightarrow \quad \grave{\nabla} \wedge \grave{\bar{h}}^{-1}(a) \quad \rightarrow \quad \omega(a) \quad \rightarrow \quad R(a \wedge b)\,. \tag{187}$$

To solve Einstein's equation for a given physical situation, symmetry considerations are used to restrict the functional form of the gauge tensor, from which a functional form for the curvature tensor is derived by (187). This, in turn, is used to simplify the form of Einstein's tensor and reduce Einstein's equation to solvable form. Finally, gauge freedom is used to analyze alternative forms for the solution and ascertain the simplest for a given application. Details, framed somewhat differently, are given in [2] and other publications by the Cambridge group, so they need not be repeated here. We confine our attention to a review of results that showcase advantages of the GC formulation and analysis.

A. Static Black Hole

The most fundamental result of GR is the gravitational field of a point particle at rest called a *black hole*. To construct a spacetime map of this field and the motion of particles within it, we represent the velocity of the hole by the constant unit vector γ_0 and locate the origin on its history. As explained in GA2, this defines a preferred reference frame and a spacetime split of each spacetime point x into designation by a time $t = x \cdot \gamma_0$ and a relative position vector $\mathbf{x} = x \wedge \gamma_0$, and a split of the vector derivative $\partial_t = \gamma_0 \cdot \nabla$ and $\boldsymbol{\nabla} = \gamma_0 \wedge \nabla$, as expressed by

$$x\gamma_0 = t + \mathbf{x}, \qquad \gamma_0 \nabla = \partial_t + \boldsymbol{\nabla}. \tag{188}$$

It will also be convenient to write $r = |x \wedge \gamma_0| = |\mathbf{x}|$ so that $\mathbf{x} = r\hat{\mathbf{x}}$, and to introduce the unit radius vector

$$e_r = \partial_r x = -\nabla r = \hat{\mathbf{x}}\gamma_0. \tag{189}$$

The Cambridge group has shown that the curvature tensor for a black hole can be written in the wonderfully compact form [2, 3]

$$R(B) = -\frac{M}{2r^3}(B + 3\hat{\mathbf{x}}B\hat{\mathbf{x}}), \tag{190}$$

where M is the mass of the black hole. Considering the local Lorentz rotation of the curvature (140), it is evident at once that the curvature tensor is invariant under a *boost* in the $\hat{\mathbf{x}} = e_r\gamma_0$ plane, as expressed by

$$\underline{L}\,\hat{\mathbf{x}} = L\hat{\mathbf{x}}\widetilde{L} = \hat{\mathbf{x}}. \tag{191}$$

This simple symmetry of the curvature tensor has long remained unrecognized in standard tensor formulations of GR.

Using the vector derivative identity (28), it is not difficult to verify by direct calculation that

$$R(b) = \grave{\partial}_a R(\grave{a} \wedge b) = 0, \tag{192}$$

so Einstein's tensor $G(a)$ vanishes when $r \neq 0$. As is evident in (190), the curvature is singular at $r = 0$; this can be shown to arise from a delta-function in the energy-momentum tensor:

$$T(a) = 4\pi\delta^3(\mathbf{x})(a \cdot \gamma_0)\gamma_0 \,. \tag{193}$$

This simple expression for the singularity contrasts with standard GR where the black hole singularity is not so well defined, and it is often claimed that such entities as "wormholes" are allowed [12]. Unfortunately for science fiction buffs, GTG rules out wormholes by requiring global solutions of Einstein's equation, as explained below.

The standard Schwarzschild solution of Einstein's field equation for a black hole is equivalent to the symmetric gauge tensor

$$\bar{h}_S(a) = (\alpha^{-1} a \cdot \gamma_0 + \alpha a \wedge \gamma_0)\gamma_0, \qquad \alpha = \left(1 - \frac{2M}{r}\right)^{\frac{1}{2}} . \tag{194}$$

Alternative solutions can be generated by gauge transformations that leave the curvature (190) invariant. Allowed transformations are determined by three parameters. One parameter is fixed by boundary conditions at infinity, a second generates displacements, and the third generates boosts defined by (191). The Cambridge group [3] has found boost-displacements that transform the "Schwarzschild gauge" (194) into a "*Kerr-Schild gauge*"

$$\bar{h}_{KS}(a) = a + \frac{M}{r} a \cdot (\gamma_0 - e_r)(\gamma_0 - e_r), \tag{195}$$

and a "*Newtonian gauge*"

$$\bar{h}(a) = a - \varphi(a \cdot \gamma_0)e_r, \qquad \varphi = \left(\frac{2M}{r}\right)^{\frac{1}{2}} . \tag{196}$$

The former corresponds to such well-known solutions as the Eddington-Finkelstein metric, while the latter has hardly been noticed in the GR literature, despite its considerable virtues, which we now expose. It should be noted that the Newtonian gauge (196) is well-defined for every $r \neq 0$, while the Schwarzschild gauge (194) has a well-known singularity at $r = 2M$. Therefore, the former is a more general solution to Einstein's equation than the latter.

To interpret results of physical measurements in the gravitational field, we follow the usual practice of setting up reference systems of imaginary observers. First, consider a *stationary observer* at a fixed distance r from the black hole, so his world line is defined by the map velocity

$$\dot{x} = \dot{t}\gamma_0, \tag{197}$$

so $\dot{\mathbf{x}} = 0$. As only gauge covariant quantities are physically significant, we must consider the observer's "gauge velocity" determined by (196):

$$v = \underline{h}^{-1}\dot{x} = \dot{t}\,(\gamma_0 + \varphi e_r). \tag{198}$$

Whence the constraint $v^2 = 1$ implies

$$\dot{t} = \left(1 - \frac{2M}{r}\right)^{-\frac{1}{2}}. \tag{199}$$

This tells us that the stationary observer's clock (reading proper time τ) runs slow compared to the proper time t of the black hole. But how can we "read" a clock on the black hole? That is answered by considering a *freely-falling observer* with a constant gauge velocity $v = \gamma_0$ equal to that of the black hole, so his map velocity is

$$\dot{x} = \underline{h}^{-1}\gamma_0 = \gamma_0 + \varphi e_r, \tag{200}$$

which we solve immediately for

$$\dot{t} = \dot{x}\cdot\gamma_0, \qquad \dot{r} = -\dot{x}\cdot e_r = -\varphi = -\left(\frac{2M}{r}\right)^{\frac{1}{2}}. \tag{201}$$

This tells us that the clock of a freely falling observer has the same rate as the proper time of the black hole, and his velocity $\dot{r}$ is identical to the infall velocity in Newtonian theory. This justifies the name "Newtonian gauge" introduced above and suggests referring to time t as "*Newtonian time.*"

As our discussion of observers suggests, the Newtonian gauge simplifies physical interpretation and analysis of motion in the gravitational field of a point particle. It has the great advantage of showing precisely how GR differs from Newtonian theory in this case. To compare with standard GR, we need the "Newtonian" gauge tensor, which from (196) is

$$\underline{g}\,a = \bar{h}^{-1}\,\underline{h}^{-1}a = a + \varphi(a\cdot\gamma_0\,e_r + a\cdot e_r\gamma_0) - \varphi^2 a\cdot\gamma_0\gamma_0. \tag{202}$$

This determines a line element

$$ds^2 = dx\cdot\,\underline{g}\,dx = dt^2 - (dr + \varphi dt)^2 - r^2 d\Omega^2. \tag{203}$$

This line element does not appear in any GR textbook. Only Ron Gautreau has expounded its pedagogical benefits in a series of AJP articles [22]. All his analysis can be neatly reformulated in Newtonian gauge theory.

Now, to analyze the general motion of a point charge in the Newtonian gauge, we multiply equation (179) by γ_0 to do a space-time split. Considering first the scalar part of the split, for static fields $\partial_t\,\underline{g} = 0$ and $\partial_t A = 0$ we find immediately the constant of motion

$$E = (\,\underline{g}\dot{x} + qA)\cdot\gamma_0 = v\cdot u + V\,, \tag{204}$$

where $V = qA \cdot \gamma_0$ is the usual electric potential and

$$v \cdot u = \gamma_0 \cdot \underline{g}\dot{x} = \underline{h}^{-1}(\gamma_0) \cdot \underline{h}^{-1}(\dot{x}) = \dot{t}\,(1 - \varphi^2) - \varphi\dot{r}\,. \tag{205}$$

We recognize this as a gauge invariant generalization of energy conservation in special relativity. The time derivative can be eliminated from this expression with the constraint

$$1 = v^2 = (\dot{x} + \varphi\dot{t}e_r)^2 = (1 - \varphi^2)\dot{t}^2 - 2\dot{t}\dot{r} - \dot{\mathbf{x}}^2\,. \tag{206}$$

As our analysis to this point is sufficient to show how electromagnetic interactions can be included, let us omit them from the rest of our analysis to concentrate on the distinctive features of the gravitational interaction. Introducing the angular momentum (per unit mass)

$$\mathbf{L} = \mathbf{x} \wedge \dot{\mathbf{x}} = r^2\hat{\mathbf{x}}\dot{\hat{\mathbf{x}}} \tag{207}$$

so we can write

$$\dot{\mathbf{x}}^2 = \dot{r}^2 + \frac{L^2}{r^2} \tag{208}$$

with $L^2 = -\mathbf{L}^2$, we eliminate $\dot{t}$ between (204) and (206) to get an equation governing radial motion of the particle:

$$E^2 = (1 - \varphi^2)\left(1 + \frac{L^2}{r^2}\right) + \dot{r}^2. \tag{209}$$

A pedagogically excellent analysis of the physical implications this equation and related issues in elementary black hole physics is given by Taylor and Wheeler [23]. Their treatment could be further simplified by adopting the Newtonian gauge.

To complete our analysis of motion in the Newtonian gauge, we return to the space-time split of equation (179) to consider the bivector part, which gives us the spatial equation of motion

$$\dot{\mathbf{g}} = -\dot{\nabla}\,\tfrac{1}{2}\dot{x} \cdot \underline{\grave{g}}(\dot{x})\,, \tag{210}$$

and, from (202), we get the more explicit expressions

$$\mathbf{g} \equiv (\,\underline{g}\dot{x}) \wedge \gamma_0 = \dot{\mathbf{x}} + \varphi\dot{t}\hat{\mathbf{x}} \tag{211}$$

$$-\dot{\nabla}\tfrac{1}{2}\dot{x} \cdot \underline{\grave{g}}(\dot{x}) = \dot{t}\left[\dot{r}\nabla\varphi + \frac{\varphi}{r}(\dot{\mathbf{x}} - \dot{r}\hat{\mathbf{x}}) + \tfrac{1}{2}\dot{t}\nabla\varphi^2\right]. \tag{212}$$

Inserting these expressions into (210), we get the equation of motion in the form

$$\ddot{\mathbf{x}} = \lambda\hat{\mathbf{x}}\,, \tag{213}$$

where λ is a complicated expression that is difficult to simplify. However, all we need is to recognize that the right side of (213) is a central force, so angular momentum is conserved. Then, with the help of (208), we obtain

$$\ddot{\mathbf{x}} = \left(\ddot{r} - \frac{L^2}{r^3}\right)\hat{\mathbf{x}}\,, \tag{214}$$

and, differentiating the radial equation (209) to eliminate $\ddot{r}$, we get the equation of motion in the simple explicit form

$$\ddot{\mathbf{x}} = -\frac{M}{r^2}\hat{\mathbf{x}} - \frac{3ML^2}{r^3}\hat{\mathbf{x}}\,. \tag{215}$$

Thus, the deviation of GR from the Newtonian gravitational force has been reduced to the single term on the right side of this equation.

To solve the equation of motion (215), we first neglect the GR correction, so the equation has the familiar Newtonian form, although the parameter is proper time rather than Newtonian time. This equation is most elegantly and easily solved by using (207) to put it in the form

$$\mathbf{L}\ddot{\mathbf{x}} = -\frac{M}{r^2}\mathbf{L}\hat{\mathbf{x}} = M\dot{\hat{\mathbf{x}}}\,, \tag{216}$$

which reveals at once the constant of motion

$$M\boldsymbol{\varepsilon} = \mathbf{L}\dot{\mathbf{x}} - M\hat{\mathbf{x}}\,. \tag{217}$$

The solution is now completely determined by the two constants of motion $\boldsymbol{\varepsilon}$ and $\mathbf{L}$, and GA facilitates analysis of all its properties, as demonstrated exhaustively elsewhere [24]. The solution is a conic section with eccentricity $\varepsilon = |\boldsymbol{\varepsilon}|$, and $\hat{\boldsymbol{\varepsilon}}$ is the direction of its major axis. For $0 \le \varepsilon < 1$ the solution is an ellipse. Changing the independent variable from proper time to Newtonian time has the effect of making the ellipse precess. That is a purely kinematic effect of special relativity.

To ascertain the GR effect on orbital precession, we use (217) to write (215) in the form

$$\frac{d\boldsymbol{\varepsilon}}{d\tau} = 3L^3\frac{\hat{\mathbf{x}}\cdot\hat{\mathbf{L}}}{r^3}\,. \tag{218}$$

Averaging over an orbital period, we get [2]

$$\Delta\boldsymbol{\varepsilon} = \frac{6\pi M}{a(1-\epsilon^2)}\boldsymbol{\varepsilon}\cdot\hat{\mathbf{L}}\,, \tag{219}$$

where a is the semimajor axis of the ellipse. For Mercury, this gives the famous GR perihelion advance of 43 arcseconds per century.

B. Other Solutions

The Cambridge group has applied GTG to most of the domains where GR is applied, with many new simplifications, insights and results. For example, anyone who has studied the Kerr solution for a rotating black hole should be pleased, if not astounded, to learn that the Kerr curvature tensor can be cast in the compact form

$$R(B) = -\frac{M}{2(r - iL\cos\theta)^3}(B + 3\hat{\mathbf{x}}B\hat{\mathbf{x}}), \tag{220}$$

where L is a constant and θ is the angle with respect to the axis of rotation. It has been known for a long time that the Kerr curvature can be obtained from the Schwarzschild curvature by complexifying the coordinates, but (220) shows that it can be reduced to a substitution for the radius in (190), and the unit imaginary is properly identified as the unit pseudoscalar, so the substitution can be interpreted as a duality rotation-scaling of the Schwarzschild curvature. GTG provides other insights into Kerr geometry, but so far they have not been sufficient to solve the important problem of matching the Kerr solution to the boundary of a rotating star.

The Cambridge group, including a number of graduate students and post-docs, is actively applying GTG to a number of outstanding problems in GR, particularly in cosmology and black hole quantum theory. As this work is continually progressing, interested readers are referred to the Cambridge website [21] for the latest news.

VII. Energy-momentum Conservation

Shortly after publishing his general theory of relativity, Einstein endeavored to define an energy-momentum tensor for the gravitational field so he could interpret his gravitational field equation as a general equation for energy-momentum conservation [25]. His proposal was immediately confronted with severe difficulties that have remained unresolved to this day despite considerable efforts of many physicists. The general consensus today is that a unique energy-momentum tensor cannot be defined, so only global energy-momentum conservation is allowed in GR [26]. However, there are dissenters who maintain that local energy-momentum conservation is a fundamental experimental fact, so it must be and can be incorporated in GR [27]. In this section we show how GTG provides a new approach to the problem of local energy-momentum conservation in GR.

An alternative to the Ricci split (148) is obtained by substituting the curvature bivector (139) into the new unitary form for Einstein's tensor (146) to get Einstein's equation in the form

$$gG(g^\beta) = g(g^\beta \wedge g^\mu \wedge g^\nu)\cdot(D_\mu\omega_\nu - \tfrac{1}{2}\omega_\mu \times \omega_\nu) = \kappa g T(g^\beta) = \kappa T^\beta\,, \tag{221}$$

where $g = |\det(g_\mu \cdot g_\nu)|^{\frac{1}{2}} = |\det(g_{\mu\nu})|^{\frac{1}{2}}$ is the "metric density," so $T^\beta \equiv gT(g^\beta)$ is a "tensor density." Next, we define an *"energy-momentum superpotential"* by

$$\kappa U^{\beta\mu} \equiv g(g^\beta \wedge g^\mu \wedge g^\nu)\cdot\omega_\nu\,, \tag{222}$$

and, using the following identity derived from an identity at the end of Appendix A

$$D_\mu[g(g^\beta \wedge g^\mu \wedge g^\nu)] = 0\,, \tag{223}$$

we write

$$g(g^\beta \wedge g^\mu \wedge g^\nu)\cdot D_\mu\omega_\nu = D_\mu U^{\beta\mu} = \partial_\mu\, U^{\beta\mu} + \omega_\mu \times U^{\beta\mu}\,. \tag{224}$$

Inserting this into (221) we obtain a split of Einstein's equation in the following form:

$$\kappa^{-1} g G(g^\beta) = \mathcal{T}^\beta - t^\beta = T^\beta\,, \tag{225}$$

where

$$t^\beta \equiv U^{\beta\mu} \times \omega_\mu + \tfrac{1}{2}\kappa^{-1} g(g^\beta \wedge g^\mu \wedge g^\nu)\cdot(\omega_\mu \times \omega_\nu) \tag{226}$$

is identified as the *energy-momentum tensor (density) for the graviational field*, so

$$\mathcal{T}^\beta \equiv \partial_\mu U^{\beta\mu}\,, \tag{227}$$

can be identified as the *total energy-momentum tensor (density)*. It follows immediately that we have the true *energy-momentum conservation law*:

$$\partial_\beta \mathcal{T}^\beta = \partial_\beta(t^\beta + T^\beta) = 0\,, \tag{228}$$

because $\partial_\beta\partial_\mu U^{\beta\mu} = -\partial_\beta\partial_\mu U^{\beta\mu} = 0$. Because of its evident importance in giving us a general conservation law, let us refer to (225) with the associated definitions (226) and (227) as the *canonical energy-momentum split*.

The most striking and perhaps the most profound feature of the canonical split is the simple linear relation (222) between the the superpotential, which determines energy-momentum density, and the connexion, which determines gravitational force. It is a kind of duality between a vector-valued function of a bivector variable $\kappa U^{\beta\mu}$ and a bivector-valued function of a vector variable ω_σ. The duality relation can be inverted by expanding the right side of (222) to get

$$\kappa U^{\mu\nu} = g(g^\mu \wedge g^\nu \wedge g^\sigma)\cdot\omega_\sigma = \tfrac{1}{2}\kappa(g^\mu \wedge g^\nu)\cdot I + g(g^\mu \wedge g^\nu)\cdot\omega_\sigma g^\sigma\,, \tag{229}$$

where

$$I \equiv \tfrac{1}{2}(g_\nu \wedge g_\mu)\cdot U^{\mu\nu} = g_\nu(g_\mu\cdot U^{\mu\nu}) = 2\kappa^{-1} g g^\sigma\cdot\omega_\sigma \tag{230}$$

is a "frame-independent" vector field. Solving (229) for the connexion bivector we obtain

$$\kappa^{-1} g\omega_\sigma = \tfrac{1}{2}(g_\sigma \wedge g_\nu \wedge g_\mu)\cdot U^{\mu\nu} + \tfrac{1}{2} I \wedge g_\sigma\,. \tag{231}$$

This surprising dual equivalence of the connexion ω_σ to the superpotential $\kappa U^{\beta\mu}$ can be regarded as a hitherto unsuspected relation of gravitational force to energy-momentum tensor. Accordingly, it can be expected to provide new physical insights, but there is insufficient space to explore the possibilities in this paper. A more complete analysis of the canonical split, including its relation to alternative splits by Einstein and others, will be published elsewhere [28].

Physically, the most important question about the canonical split is the uniqueness of the energy-momentum tensor defined by (227). In other words, to what degree is energy-momentum localizable? We have already noted that the connexion $\omega_\mu = \omega(g_\mu)$ is invariant under the position gauge transformations. Hence, the superpotential and all other elements of the canonical split are gauge invariant. This completely solves Einstein's problem of defining a coordinate independent energy-momentum tensor.

In contrast to the Ricci split (148), however, the canonical split is not rotation gauge covariant. According to (124), the connexion transforms as a "*pseudotensor*" under local rotations. Since the superpotential $U^{\mu\nu}$ is directly proportional to ω_μ and the energy-momentum tensor is a gradient thereof, these quantities depend on the choice of the local rotation gauge, which is not uniquely defined by the theory. However, local energy-momentum conservation is not necessarily destroyed by this gauge variability, because the local gravitational force retains its duality to local energy-momentum density. To make the point explicit, consider the geodesic equation for a particle with invariant velocity $v = \underline{h}^{-1}(\dot{x})$:

$$v \cdot Dv = \dot{v} + \omega(v) \cdot v = 0. \tag{232}$$

For a given gauge, the term $-\omega(v) \cdot v$ can be regarded as the gravitational force on the particle. This force changes with a change in gauge, but the superpotential and energy-momentum tensor change in the same way. Therefore, the local relation of force to energy-momentum is preserved. In other words, with respect to local energy-momentum exchange, all gauges are equivalent. In this sense, *energy-momentum is localizable* whatever gauge is chosen. However, much more analysis will be needed to confirm this line of reasoning.

VIII. Spacetime Modeling Games

Every scientific theory is characterized by general principles, often ill-defined or unstated, that define the domain and structure of the theory. Accordingly, an essential step in evaluating a theory is articulating its defining principles. Einstein taught us that models of spacetime underlie every physical theory, so that is where we begin. To facilitate comparison with alternative theories, especially Einstein's general relativity (GR), I have outlined the basic principles of gauge theory gravity (GTG) as I see it in Table 2. Some commentary is needed to understand the design and intent of the table.

In the early stages of GTG development the Cambridge group was keen to show how it differs from GR, particularly in the treatment of black holes [18]. I cannot speak for their current opinions, but my own view is that GTG is not fundamentally different from GR. I see it, rather, as simplifying and clarifying the structure of GR — in a sense, as perfecting the foundations and completing the development of GR. As a bonus, GTG brings the full power of geometric calculus to the analysis and solution of problems in the domain of GR, and it goes beyond GR when torsion is included [29].

As Newtonian mechanics is the original prototype for all physical theories, I have followed Newton's example in formulating a set of *Universal Laws* that define GTG. The Laws are "universal" in the sense that they apply to all models and explanations consistent with the theory, and they are intended to embrace all of physics. This is not to say that the Laws cannot be modified, improved or extended.

I have argued elsewhere that it takes six Universal Laws to define Newton's theory completely [30, 31], so I have constructed analogs of those six Laws for GTG in Table 2. This stimulates an enlightening comparison of Newtonian theory with GR and GTG, giving us a new slant on Newton's original formulation and how physics has evolved since then.

I have partitioned the six Laws into two major categories: *I. Laws of spacetime structure and measurement. II. Laws of motion and interaction.* The first category, embracing the Zeroth and First Laws, is a central concern of this paper, so it deserves the most thorough commentary.

The second category, embracing the Second through the Fifth Laws, consists of standard equations and concepts of GR expressed in the language of STA, augmented with the new definition of energy-momentum tensor in Section VII. I confess that I am not fully satified with the formulation of Laws in this category, but I offer it as a means to focus attention on fundamental theoretical issues. The categorization according to Newtonian Laws may not be not optimal, but it provides a convenient platform for comparison with Newtonian theory.

Table 2: UNIVERSAL LAWS FOR GAUGE GRAVITY THEORY

I. Laws of Spacetime Structure and Measurement

Zeroth Law: (Spacetime Structure)

(0.1) *Irreducible physical events are represented by points in 4D Minkowski spacetime.*

(0.2) *Physical entities are modeled as particles or fields on spacetime (with values in the spacetime algebra).*

(0.3) *Particle histories are positive timelike (or lightlike) curves.*

First Law: (Gauge Equivalence)

(1.1) *The equations of physics are gauge invariant under global displacements and covariant under local Lorentz rotations.*

(1.2) *The effects of gravity are represented by a gauge tensor* $\bar{h}$ *and its derivatives.*

(1.3) *Only gauge invariant quantities are physical observables.*

II. Laws of Motion and Interaction.

Second Law: (Particle Motion)

Every particle has a gauge covariant velocity $v = \underline{h}^{-1}(\dot{x})$ *and an equation of motion of the form*

$$\dot{v} = -\omega(v) \cdot v + \Omega \cdot v\,, \tag{2.1}$$

where the gravitational force is specified by the "connexion bivector" $\omega(v)$ *and nongravitational forces are specified by the bivector* Ω*. For a particle with rigid internal structure defined by a comoving frame* $\{e_\mu\}$ *with* $v = e_0$*, the equation of motion generalizes to*

$$\dot{e}_\mu = [\Omega - \omega(v)] \cdot e_\mu\,. \tag{2.2}$$

Third Law: (Gravity Field and Energy-momentum Conservation)

The gravitational field $\omega(a)$ *is determined (up to a rotation gauge transformation) by Einstein's equation*

$$\tfrac{1}{2}(a \wedge \partial_b \wedge \partial_c) \cdot R(b \wedge c) = \kappa T(a)\,, \tag{3.1}$$

where $R(b \wedge c)$ *is the curvature bivector (defined in Table 1),* $T(a)$ *is the total energy-momentum tensor for all entities present, and* κ *is a constant. With vanishing torsion, the total energy-momentum density is given by*

$$\mathcal{T}(a) = \grave{U}(a\wedge\grave{\bigtriangledown}) \quad \text{with} \quad U(a\wedge b) \equiv g\bar{h}(a\wedge b\wedge\partial_c) \cdot \omega[\underline{h}^{-1}(c)]\,, \tag{3.2}$$

so it satisfies the conservation law

$$\grave{\mathcal{T}}(\grave{\bigtriangledown}) = 0\,. \tag{3.3}$$

Fourth Law: (Energy-momentum Tensor)

Every physical system has an energy-momentum tensor $T(a)$ *with the following properties:*

Additivity: $T(a)$ is the sum of energy-momentum tensors for any subsystems. Thus, for a two component system,

$$T(a) = T_1(a) + T_2(a)\,. \tag{4.1}$$

Energy positivity: For every nonspacelike vector v,

$$v \cdot T(v) \geq 0\,, \qquad \text{and} \qquad [T(v)]^2 \geq 0\,. \tag{4.2}$$

Fifth Law: (Sources and Fields)

Every material system is the source of fields that propagate according to field equations and interact with particles according to the Second Law.

My formulation of Universal Laws for GTG is in accord with Einstein's view of GR as an extension of *classical field theory.* Even so, the formulation is general enough to include relativistic quantum mechanics, though I need not discuss the matter here, as the Cambridge group has treated the subject thoroughly with an elegant Lagrangian approach [3, 2].

Lagrangian field theory has many well-known advantages, including (a) derivation from a single Lagrangian guarantees mutual consistency of the various field equations, (b) derivation of conservation laws from symmetry principles (Noether's theorem). It sometimes seems that you get more out of the Lagrangian than you put in, but, of course, you can't really. The advantage of an explicit listing of Universal Laws like Table 2 is that it helps one identify critical components of the theory. That can serve as a guide in constructing a serviceable Lagrangian. Thus, the Cambridge Lagrangian (tacitly) presumes the Zeroth Law to begin with and builds in the First Law by introducing the gauge tensor and the connexion bivector as independent variables; additivity of interactions is introduced in the usual way through additive terms in the Lagrangian.

I. Laws of spacetime structure and measurement

Although the standard curved space version of GR is mathematically equivalent to GTG in all essential respects, there are striking differences between the two versions in regard to geometric intuition and physical interpretation. In GR physical entities and interactions are woven into the fabric of spacetime; whereas in GTG spacetime is little more than a bookkeeping system and physical entities are represented as geometric objects that live *on* rather than *in* spacetime — so to speak, they live in the tangent algebra. The Zeroth Law in Table 2 lays out the essential ideas.

The Zeroth Law is about making spacetime maps. Law (0.1) asserts that physical events are mapped into points in a 4D Minkowski vector space in exactly the same way as in special relativity [7]. Spacetime maps express contiguity and differentiability relations among events as well as a *partial ordering of events* into timelike, lightlike and spacelike relations as determined by the Minkowski lightcone. These relations among events are "*premetrical*," as they do not require a concept of spacetime distance between events. The metrical notion of distance is introduced by the First Law.

The ordering of events in a particle history is specified by Law (0.3). To give it a mathematical formulation, let $x = x(\lambda)$ and $y = y(\lambda)$ be two distinct histories that cross at some point and denote their derivatives by overdots. Then Law (0.3) asserts that at the crossing point

$$\dot{x}^2, \dot{y}^2 \geq 0 \qquad \text{and} \qquad \dot{x} \cdot \dot{y} > 0 \,. \tag{233}$$

The first inequality asserts that each history is timelike or lightlike, while the second asserts that the two histories share the same forward lightcone, which we designate as *positive.* No metrical concept is involved here, though the concept of spacetime signature is crucial. It is worth mentioning that Law (0.3) is easily generalized to allow piecewise differentiable curves that can be used to describe creation and annihilation of particles.

In GTG the concept of observer is replaced by, or if you will, augmented by the concept of *spacetime map.* In GTG we can regard each inertial system as identified with a single fixed frame $\{\gamma_\mu\}$ defining the flat space background of Minkowski space. So to speak, all inertial frames are reduced to a single flat spacetime where the observations of any observer can be recorded in a spacetime map using any convenient coordinate system. Different spacetime maps can then be compared with procedures specified by the First Law.

Besides designating locations of objects, maps designate the kind of object at specified locations. Law (0.2) specifies the class of geometric objects that GTG allows for modeling physical entities, namely, particles or fields with values in the real STA. We have seen that this class includes Dirac spinors and coordinate-free representation of the tensors that are needed for GR. The class also includes representations for particles of any spin within the real STA, but we cannot go into that here. All this is in full accord with the conventional view that fundamental physical entities should by modeled by representations of the Lorentz group, though association with the Lorentz group comes from the First Law in GTG.

The heart of the First Law is the *Displacement Gauge Principle* (DGP) and the *Rotation Gauge Principle* (RGP) that were introduced and thoroughly discussed in Section III. My purpose here is to place these principles in a broader theoretical context and reconcile them with the ideas of Einstein and Newton.

We have seen how the rich mathematical structure of differential geometry is generated by the DGP and RGP on the flat spacetime background. The DGP requires existence of the gauge tensor $\bar{h}(a)$, while the RGP entails existence of the coderivative D and the connexion bivector $\omega(a)$. By differentiation the connexion is related to the gauge tensor and the curvature is derived from the connexion. Law (1.2) identifies this geometry with effects of gravity. Indeed, Law (1.2) can be regarded as a formulation of Einstein's "Strong Principle of Equivalence," as it asserts *universality* of gravitational effects on all fields and particles.

The First Law provides the foundation for quantitative measurements on spacetime maps admitted by the Zeroth Law. Every measurement is a comparison of some sort. We have seen how the gauge tensor enables a gauge invariant definition of proper time on a particle history, thereby defining equivalent time intervals on the history, which is the foundation for time measurement. More generally, the gauge tensor determines a metric tensor that enables gauge invariant comparisons of lengths and angles at any spacetime point. However, we need more than that for a general theory of measurement. To compare physical measurements at different places and times, we need to generalize the Euclidean notion of *congruence*. That generalization is supplied by the DGP, which allows us to map any spacetime region into any other. In other words, the background *spacetime is globally homogeneous;* this can be regarded as an alternative formulation of the DGP. To complete our formulation of generalized congruence, we need a means to compare directions at neighboring points in any spacetime region. That is supplied by the RGP, which can be regarded as asserting that *spacetime is locally isotropic.*

The fundamental roles of the DGP and the RGP as "Laws of Measurement" can be neatly encapsulated by referring to the the First Law as the *Principle of Gauge Equivalence*. It is about equivalence of measurements or physical configurations rather than equivalence of observers. Displacement and rotation gauge transformations are *active* transformations comparing lengths and directions of physical configurations at different locations, rather than *passive* coordinate transformations comparing labels that different observers assign to the same configuration of events. In GTG the assignment of coordinates to events (spacetime maps) is completely separated from the assignment of measurable lengths and relative directions.

The RGP can be regarded as a clarification and refinement of Einstein's *weak principle of equivalence* (WPE), though the precise meaning and import of the WPE continues to be a matter of dispute [17]. For our purposes, it suffices to take the WPE as asserting that, for a given reference system at any spacetime point, a gravitational force is indistinguishable from an acceleration of the reference system; in other words, the gravitational force can be "cancelled" by accelerating the reference system. In accord with equation (124), the RGP allows us to interpret that local acceleration as a local Lorentz rotation that transforms the connection ω_μ (determining the gravitational force) to a physically equivalent vanishing connection $\omega'_\mu = L\omega_\mu \widetilde{L} - (\partial_\mu L)\widetilde{L} = 0$, with the cancelling acceleration expressed by the derivative $(\partial_\mu L)\widetilde{L}$. Since the WPE is presumed to hold for arbitrary gravitational forces, it requires the existence of arbitrary gauge rotations. Thus, the WPE can be regarded as an alternative formulation of the RGP, or better, perhaps, as an operational formulation in terms of physical measurements. Moreover, formulation of the WPE as a gauge principle makes it clear that the WPE implies existence of a gravitational force expressed as a connexion.

In Section IVA, we saw how Einstein's *general principle of relativity* can be purged of its adventitious reliance on coordinates by recasting it as a displacement gauge principle. With respect to physical measurements, this too can be regarded as and equivalence principle. Thus, it may be better to speak of "general (gauge) equivalence" rather than "general relativity," or, at least, regard gauge equivalence principles and relativity principles as different names for the same thing.

It may be useful to distinguish "special and general gauge theories," just as we distinguish "special and general relativity theories." We can define the *special (or restricted) gauge theory* by the requirement that the gauge tensor is constant, which we know implies a vanishing gravitational field. This can be neatly expressed by reformulating the First Law as follows.

Restricted First Law: *The history of a free particle is a straight line.*

Every straight line through a point x_0 has a parametric equation of the form $x(\tau) = x_0 + v\tau$, where v is a constant vector. The most general displacement that maps straight lines into straight lines is a *Poincarè transformation*:

$$f : x \quad \rightarrow \quad x' = f(x) = Rx\widetilde{R} + c\,, \tag{234}$$

where c is a constant vector and R is a constant rotor. This is not a (passive) relabeling of points, but an active transformation between congruent free particle histories. The differential of this transformation is $\underline{f}(a) = \underline{R}a = Ra\widetilde{R}$, so if we begin with $\underline{h} = 1$, this induces a gauge transformation to the constant gauge field $\underline{h}a = \underline{R}a$. Thus, the (constant) *Lorentz Group* defines a class of equivalent gauge tensors.

The restricted version of the First Law is a fairly straightforward generalization of Newton's First Law [30], which assigns uniform motion in a straight line to free particles. This establishes the First Law in GTG as a natural generalization of Newton's so-called Law of Inertia. The Newtonian version of the Zeroth Law models physical space with a Euclidean metric. In the present version all metrical assumptions have been transferred to the First Law. In retrospect we can see that Newton's First Law tacitly involved metrical assumptions for defining free particles and inertial frames — though it was not until the latter part of the nineteenth century that physicists recognized its necessity for defining inertial systems. I, for one, am amazed at the richness of the gauge theory generalization of a subtle Law that has often been dismissed as a trivial case of Newton's Second Law. It is time in the physics curriculum to recognize Newton's great First Law for what it is: a fundamental principle of gauge equivalence.

II. Laws of motion and interaction.

Law (2.1) in Table 2 is a fairly straightforward generalization of Newton's Second Law. As we noted in Section VD where it was studied at some length, Law (2.1) reduces to the gravitational geodesic equation when $\Omega = 0$. The gravitational force is gauge dependent, though it is well-defined when a physically significant gauge is specified, as we saw in our study of the Newtonian gauge for a black hole field.

I have suppressed "inertial mass" in Law (2.1) so it gives the geodesic equation for lightlike (massless) particles, and I have included its generalization (2.2) to include intrinsic spin, because these equations are so basic and useful. We have seen in Section VE how (2.2) can be reduced to a single spinor equation. Indeed, that equation it has been derived as a classical limit of the Dirac equation with electromagnetic interaction [7], and its generalization to include gravitational interactions is given here.

In Section VC we saw how the equation for particle motion can be derived from Einstein's field equation, at least for a particular case. One could argue that this fact renders the Second Law in Table 2 superfluous. However, I propose to regard equations (2.1) and (2.2) as general forms for any admissible equations of motion, with which the derivation of specific forms from specific energy-momentum tensors must be consistent. It is in the context of such specific derivations that the fundamental problems of defining mass and mass renormalization should be addressed.

I have used the coordinate-free method from Appendix B to formulate the equations in the Third Law, though a semi-coordinate method was used for the more detailed formulation and analysis in Section VII. It may be surprising to see Einstein's Law (3.1) proposed as a generalization of Newton's Third Law. However, according to (228) the "generalized conservation law" (3.3) implies the local balance of energy-momentum fluxes

$$\partial_\beta t^\beta = -\partial_\beta T^\beta, \tag{235}$$

which is indeed a generalization of Newton's Third Law. The balance among nongravitational fields and particles is covered by (4.1).

It should be mentioned, that the Universal Laws in Table 2 permit a gravitational connexion with *torsion*, though that requires redefining the energy-momentum tensor (3.2). Nonvanishing torsion is explicitly included in the Cambridge treatment of GTG [29, 3], so it need not be considered here. Suffice it to say that association of torsion with intrinsic spin has implications for angular momentum conservation and symmetry of the energy-momentum tensor.

In my expanded formulation of Newton's Laws [31], I followed Arnold Sommerfeld in formulating the "superposition principle" as a Fourth Law, independent of Newton's Second Law. The additivity of independent bivectors $\omega(v)$ and Ω in the Second Law of Table 2 is already a limited *superposition principle.* However, field theory (which was not available to Newton) enables us to trace superposition to properties of the field equations and sources. Thus, superposition of electromagnetic forces follows from the linearity of Maxwell's equation. In contrast, the nonlinearity of Einstein's equation precludes superposition of gravitational forces from different sources, though I propose the additivity Law (4.1) as a weaker kind of superposition principle. As that additivity is presumed to apply generally, I have folded it into a Fourth Law postulating general properties for any (non-gravitational) energy-momentum tensor. The importance of the energy positivity Law (4.2) has been demonstrated by Hawking and Ellis [32].

To complete the formulation of Newton's Laws, I proposed a Fifth Law that specifies a general functional form for laws of force in the Newtonian system [31]. This sets the stage for Newton's general research program: to systematically search for fundamental force laws in nature. The ground rules for that search have been changed by classical field theory, which regards all forces as propagated by fields with sources in material particles. Accordingly, I propose a Fifth Law that sets the stage for investigating fields and field equations in nature. Of course, physics has gone a long way along that path. So far only electromagnetic field theory has been fully incorporated into GTG. Weak and strong interactions remain to be included, and some modification of GTG may be needed to do that in a geometrically fundamental way.

Of course, GTG is not the only spacetime modeling game in town, though it may be the most conservative. String theory proposes to generalize the spacetime manifold to 10 dimensions. It remains to be seen if extra spacetime dimensions are really needed.

Appendix A. Cocurl and Connexion

To solve the gravitational field equation, we need to relate the coderivative (or its connexion) to the gauge tensor. As explained in Section IVB, Riemannian geometry requires

$$D \wedge g^{\mu} = D \wedge \bar{\bar{\nabla}} x^{\mu} = 0\,. \tag{236}$$

This is equivalent to the condition that, for any scalar field $\phi = \phi(x)$,

$$D \wedge D\phi = D \wedge \bar{\bar{\nabla}}\phi = D \wedge \bar{h}(\nabla\phi) = 0\,. \tag{237}$$

To understand the geometric significance of vanishing torsion, observe that (237) is actually an *integrability condition* for scalar fields, as seen by applying (236) to get

$$D \wedge D\phi = D \wedge g^{\mu}\partial_{\mu}\phi = g^{\nu} \wedge g^{\mu}\partial_{\nu}\partial_{\mu}\phi = 0\,, \tag{238}$$

whence

$$\partial_{\nu}\partial_{\mu}\phi = \partial_{\mu}\partial_{\nu}\phi\,. \tag{239}$$

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

For $\phi = a \cdot x$ and constant vector a, (237) becomes

$$D \wedge \bar{h}(a) = 0\,. \tag{240}$$

By (123) this can be expanded to

$$\bar{\bar{\nabla}} \wedge \bar{h}(a) + \partial_b \wedge [\omega(b) \cdot \bar{h}(a)] = 0\,, \tag{241}$$

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

$$H(a) \equiv -\bar{h}(\nabla \wedge \bar{h}^{-1}(a)) \,, \tag{242}$$

and note that

$$H(a) = -\bar{h}(\grave{\nabla}) \wedge \bar{h}[\,\grave{\bar{h}}^{-1}(a)\,] = \grave{\bar{\bar{\nabla}}} \wedge \grave{\bar{h}}[\,\bar{h}^{-1}(a)\,] \,, \tag{243}$$

so (241) becomes

$$H(a) = \partial_b \wedge [\, a \cdot \omega(b)\,] \,. \tag{244}$$

Protraction of this equation gives the position gauge invariant quantity

$$H \equiv \partial_b \wedge \omega(b) = -\tfrac{1}{2}\partial_b \wedge H(b) \,. \tag{245}$$

It follows that

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

Adding this to (244), we obtain the desired result

$$\omega(a) = H(a) + a \cdot H \,. \tag{246}$$

This formula enables us to calculate $\omega(a)$ from a given $\bar{h}$ by first calculating $H(a)$ from (242) and then H from (245). The simple functional form (242) makes $H(a)$ seem almost as important as $\omega(a)$ itself.

As a check on our analysis, note that in the absence of gravity we can choose the gauge $\underline{h} = \underline{1}$, so $\omega(a)$ vanishes everywhere, and the coderivative reduces to the derivative. This is most easily proved from (242). Every other $\underline{h}$ can be generated from $\underline{1}$ by a transformation $f(x)$, so its adjoint $\bar{h}$ becomes a gradient and

$$\bar{h}^{-1}(a) = \bar{f}(a) = \nabla_x f(x) \cdot a \,. \tag{247}$$

Inserted into (242), this makes $H(a)$ vanish because of the operator identity $\nabla \wedge \nabla = 0$, so $\omega(a)$ vanishes by (246). Similarly, displacement gauge covariance of $H(a)$, and hence of $\omega(a)$, can be verified directly by substituting $\bar{h}' = \bar{h}\bar{f}$ into (242).

Many other differential identities are consequences of vanishing torsion. The following is needed for Section VII. For $g = |\det(g_{\mu\nu})|^{\frac{1}{2}}$, every textbook on GR derives the result

$$L^{\nu}_{\nu\mu} = \partial_\mu \ln g \,. \tag{248}$$

Whence we derive the identity

$$D_\mu(g g^\mu) = g D_\mu g^\mu + Dg = 0 \,, \tag{249}$$

and from this, the identity

$$D_\mu[g(g^\mu \wedge g^\nu)] = g D \wedge g^\nu = 0 \,. \tag{250}$$

Appendix B. Properties of the Coderivative and Curvature

This appendix is devoted to summarizing and analyzing properties of the coderivative and the curvature tensor using the coordinate-free techniques of GC to demonstrate its advantages.

Using (236) we can recast the curvature equation (135) in terms of the coderivative:

$$D \wedge Dg^\alpha = \tfrac{1}{2} R^\alpha_{\mu\nu\beta}(g^\mu \wedge g^\nu) g^\beta \,. \tag{251}$$

This can be analyzed further in the following way:

$$D^2 g^\alpha = (D \cdot D + D \wedge D) g^\alpha = D(D \cdot g^\alpha + D \wedge g^\alpha) \,. \tag{252}$$

Hence, using (236) again, we obtain

$$(D \wedge D) g^\alpha = D(D \cdot g^\alpha) - (D \cdot D) g^\alpha \,. \tag{253}$$

The right-hand side of this equation has only a vector part; hence the trivector part of (251) vanishes to give us

$$D \wedge D \wedge g^\alpha = \tfrac{1}{2} R^\alpha_{\mu\nu\beta}(g^\mu \wedge g^\nu \wedge g^\beta) = 0 \,. \tag{254}$$

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

$$R^\alpha_{\mu\nu\beta} + R^\alpha_{\beta\mu\nu} + R^\alpha_{\nu\beta\mu} = 0 \,. \tag{255}$$

However, its deep significance is that it implies

$$D \wedge D \wedge M = 0 \,. \tag{256}$$

for any k-vector field $M = M(x)$. This answers the question raised in Section VB about the existence of a vector potential for the electromagnetic field. It is a consequence of the condition (236) for vanishing torsion.

By the way, equation (251) reduces to

$$D \wedge Dg^\alpha = (D \wedge D) \cdot g^\alpha = R^\alpha_\beta g^\beta \,, \tag{257}$$

where

$$R^\alpha_\beta = R^\alpha_{\beta\mu\nu} g^{\mu\nu} \tag{258}$$

is the standard *Ricci tensor*. Comparing (257) with (253), we get the following provocative form for the Ricci tensor:

$$R(g^\alpha) \equiv R^\alpha_\beta g^\beta = D(D \cdot g^\alpha) - (D \cdot D) g^\alpha \,. \tag{259}$$

Though this suggests alternative forms for Einstein' equation, we will not investigate that in this paper.

For vector fields $a = a^\mu g_\mu$ and $b = b^\nu g_\nu$ the fundamental equation (138) can be put in the form

$$[a\cdot D, b\cdot D]M = R(a\wedge b)\times M\,, \tag{260}$$

where it has been assumed that $a\cdot Db - b\cdot Da = 0$ to avoid inessential complications.

Equation (260) shows that *curvature is a linear bivector-valued function of a bivector variable*. Thus, for an arbitrary bivector field $B = B(x)$ we can write

$$R(B) \equiv \tfrac{1}{2}B\cdot(\partial_b\wedge\partial_a)\,R(a\wedge b) = \tfrac{1}{2}B^{\nu\mu}R(g_\mu\wedge g_\nu)\,, \tag{261}$$

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 [8]. To that end, it is helpful to introduce the terminology *traction, contraction* and *protraction*, respectively, for the tensorial operations

$$\partial_a\, R(a\wedge b) = g^\mu R(g_\mu\wedge b) = \gamma^\mu R(\gamma_\mu\wedge b)\,, \tag{262}$$

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

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

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 (261). Of course, the replacement of vector derivatives by basis vectors and sums over indices in (262) is necessary to relate the following coordinate-free relations to the component forms of standard tensor analysis.

To reformulate (260) 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 (260) becomes

$$[a\cdot D, b\cdot D]c = R(a\wedge b)\cdot c\,. \tag{263}$$

To reformulate this as a condition on the vector coderivative, we simply eliminate the variables a and b by traction. Protraction of (263) 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

$$D\wedge D = \tfrac{1}{2}(\partial_b\wedge\partial_a)[\,a\cdot D, b\cdot D\,] \tag{264}$$

gives

$$D \wedge D \wedge c = [\,\partial_b \wedge \partial_a \wedge R(a \wedge b)\,] \cdot c + \partial_a \wedge R(a \wedge c)\,. \tag{265}$$

According to (256) 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

$$\partial_a \wedge R(a \wedge b) = 0\,. \tag{266}$$

This constraint on the Riemann curvature tensor is called the *Ricci identity*, or, in recent literature, the *Bianchi identity of the first kind.*

The requirement (266) that the curvature tensor is *protractionless* has an especially important consequence. The identity

$$\partial_b \wedge [B \cdot (\partial_a \wedge R(a \wedge b))] = \partial_b \wedge \partial_a B \cdot R(a \wedge b) - B \cdot (\partial_b \wedge \partial_a) R(a \wedge b) \tag{267}$$

vanishes on the left side because of (266), and the right side then implies that

$$A \cdot R(B) = R(A) \cdot B\,. \tag{268}$$

Thus, the curvature is a *symmetric* bivector function. This symmetry can be used to recast (266) in the equivalent form

$$R\big((a \wedge b \wedge c) \cdot \partial_e\big) \cdot e = 0\,. \tag{269}$$

On expanding the inner product in its argument, it becomes

$$R(a \wedge b) \cdot c + R(c \wedge a) \cdot b + R(b \wedge c) \cdot a = 0\,, \tag{270}$$

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

As noted in (258), contraction of the curvature tensor defines the *Ricci tensor*

$$R(a) \equiv \partial_b \cdot R(b \wedge a)\,. \tag{271}$$

The Ricci identity (266) implies that we can write

$$\partial_b \cdot R(b \wedge a) = \partial_b R(b \wedge a)\,, \tag{272}$$

and also that the Ricci tensor is protractionless:

$$\partial_a \wedge R(a) = 0\,. \tag{273}$$

This implies the symmetry

$$a \cdot R(b) = R(a) \cdot b\,. \tag{274}$$

An alternative expression for the Ricci tensor is obtained by operating on (263) with (264) and establishing the identity

$$\tfrac{1}{2}(\partial_a \wedge \partial_b) \cdot [\,R(a \wedge b) \cdot c\,] = R(c)\,. \tag{275}$$

The result is, in agreement with (257),

$$D \wedge D\, a = (D \wedge D) \cdot a = R(a)\,. \tag{276}$$

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 (276) 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 (260) 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

$$D \wedge D(a \wedge b) = R(a) \wedge b + a \wedge R(b) - 2R(a \wedge b)\,, \tag{277}$$

or equivalently

$$2R(a \wedge b) = (D \wedge Da) \wedge b + a \wedge (D \wedge Db) - D \wedge D(a \wedge b)\,. \tag{278}$$

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*

$$R \equiv \partial_a R(a) = \partial_a \cdot R(a)\,. \tag{279}$$

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

$$[a \cdot D, [\, b \cdot D, c \cdot D\,]] + [\, b \cdot D, [\, c \cdot D, a \cdot D\,]] + [\, c \cdot D, [\, a \cdot D, b \cdot D\,]] = 0\,. \tag{280}$$

By operating with this on an arbitrary nonscalar multivector M and using (260), we can translate it into a condition on the curvature tensor that is known as the *Bianchi identity:*

$$a \cdot DR(b \wedge c) + b \cdot DR(c \wedge a) + c \cdot DR(a \wedge b) = 0\,. \tag{281}$$

Like the Ricci identity (269), this can be expressed more compactly as

$$\grave{R}[\,(a \wedge b \wedge c) \cdot \grave{D}\,] = 0\,, \tag{282}$$

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

$$\grave{D} \wedge \grave{R}(a \wedge b) = 0\,. \tag{283}$$

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

Contraction of (283) with ∂_a gives

$$\grave{R}(\grave{D} \wedge b) - D \wedge R(b) = 0\,. \tag{284}$$

A second contraction yields the differential identity

$$\grave{G}(\grave{D}) = \grave{R}(\grave{D}) - \tfrac{1}{2} DR = 0\,, \tag{285}$$

where

$$G(a) \equiv R(a) - \tfrac{1}{2} aR \tag{286}$$

is the *Einstein tensor.*

References

[1] H. Poincaré, *Science and Hypothesis*, Original French edition 1902; English translation 1905 (Dover, New York, 1951).

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

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

[4] A. Grünbaum, *Philosophical Problems of Space and Time* (Knopf, New York, 1963). Poincaré's *conventionalism* (which holds that the distinction between flat and curved geometries for space and time rests on conventions in the physical definition of congruence in measurement) was much debated in the half century after the advent of GR. Grünbaum gives a thorough critical analysis of Poincaré's view (especially in Sect. B of Chap. IV) and its critics (including Einstein). The subject is ripe for reconsideration now that GTG has supplied a complete mathematical realization of Poincaré's position.

[5] R. Feynman, *Feynman Lectures on Gravitation* (Addison-Wesley, Reading, 1995).

[6] D. Hestenes, "Oersted Medal Lecture 2002: Reforming the mathematical language of physics," Am. J. Phys. **71** 104–121 (2003).

[7] D. Hestenes, "Spacetime physics with geometric algebra," Am. J. Phys. **71**: 691–704 (2003). Referred to as GA2 in the text.

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

[9] 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.

[10] R. d'Inverno, *Introducing Einstein's Relativity* (Clarendon Press, Oxford, 1992).

[11] 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).

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

[13] D. Hestenes, "The Design of Linear Algebra and Geometry," Acta Applicanda Mathematicae **23**: 65–93 (1991).

[14] C. Doran, D. Hestenes, F. Sommen & N. Van Acker, "Lie Groups as Spin Groups," J. Math. Phys. **34**: 3642–3669 (1993).

[15] Of course, this use of the "overbar notation" is an abuse of our notation for linear operators, but as the operator $\bar{h}$ is involved, it helps distinguish gauge covariant fields from fields that are not gauge covariant. Caligraphic font is used for the same purpose in [2].

[16] W. R. Davis, "The Role of Tetrad and Flat-Metric Fields in the Framework of the General Theory of Relativity," Nuovo Cimento **43B**: 2816–20 (1966).

[17] J. Earman, "Covariance, Invariance, and the Equivalence of Frames," Found. Phys. **4**: 267–289 (1974).

[18] S. Gull, A. Lasenby & C. Doran "Geometric algebra, spacetime physics and gravitation," In O. Lahav, E. Terlevich & R. Terlevich (Eds.) *Gravitational Dynamics* (Cambridge University Press, Cambridge, 1996). p. 171-180.

[19] P. Havas "The Early History of the 'Problem of Motion' in General Relativity," In D. Howard & J. Stachel (Eds.) *Einstein and the History of General Relativity* (Birkhäuser, Boston, 1989). p. 234-276.

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

[21] <www.mrao.cam.ac.uk>.

[22] R. Gautreau, "Light cones inside the Schwarzschild radius," Am. J. Phys. **63**: 431–439 (1995);"Cosmological Schwarzschild radii and Newtonian gravitational theory," Am. J. Phys. **64** 1457–1467 (1996). R. Gautreau & J. Cohen, "Gravitational collapse in a single coordinate system," Am. J. Phys. **63**: 991–999 (1995); "Gravitational expansion and the destruction of a black hole," Am. J. Phys. **65**: 198–201 (1997).

[23] E. Taylor & J. Wheeler, *Exploring Black Holes: Introduction to General Relativity* (Addison Wesley, Reading, 2004).

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

[25] A. Einstein, "Hamilton's Principle and the General Theory of Relativity," Sitzungsber.D. Preuss. Akad. D. Wiss., 1916. English translation in *The Principle of Relativity* (Dover, 1923).

[26] Page 46 in reference [12].

[27] S. Babak & L. Grishchuk, "Energy-momentum tensor for the gravitational field," Phys. Rev. D **61**, 1–18 (1999).

[28] D. Hestenes, "Energy-momentum Complex in General Relativity and Gauge Theory Gravity." To be published. Available at the website: <http://modelingnts.la.asu.edu>.

[29] C. Doran, A. Lasenby, A. Challinor & S. Gull, "Effects of Spin-Torsion in Gauge Theory Gravity," J. Math. Phys. **39**: 3303-3321 (1998).

[30] Chapter 9: "Foundations of Mechanics" in [24]. This chapter was replaced by a chapter on Relativity in the Second Edition. However, it is available at the website: <http://modelingnts.la.asu.edu>

[31] D. Hestenes, "Modeling Games in the Newtonian World," Am. J. Phys. **60**: 732-748 (1992).

[32] S. Hawking & G. Ellis, *The Large Scale Structure of Spacetime,* (Cambridge, 1973).