On the universality of Maxwell's equations

On the Universality
of Maxwell's Equations

D. H. Sattinger

Department of Mathematics

University of Arizona

Tucson, Arizona

June 5, 2017

Revised

Abstract

Einstein's theory of relativity is based on the Principle of Equivalence, Hilbert's on invariant theory and the calculus of variations. The two paradigms are not equivalent. Using the universality of Maxwell's equations, Hilbert's variational method is used to determine the energy-momentum tensor uniquely, and to show that general relativity can be formulated on the basis of Maxwellian, rather than specific physical force fields. A unified field theory is proved in which the Maxwellian force fields are all on an equal footing, distinct from the geometric field.

Keywords: Maxwell's Equations, Universality, Relativity.

dsattinger@math.arizona.edu
http://math.arizona.edu/∼dsattinger/

617 N. Santa Rita Ave.
P.O. Box 210089
Tucson, AZ 85721-0089 USA

1 Introduction

Newtonian physics played a fundamental role in the prediction and discovery of Neptune; it accurately described the motions of all the planets in the solar system save Mercury; but it could not account for the anomalous 38 seconds of arc per century observed in its orbit [11]. The astounding accuracy of Einstein's geometric theory in explaining those anomalies – together with his calculation of the correct bending of light rays from a distant star passing by the Sun – led to a rapid acceptance of his theory of gravitation.

Einstein's investigation of gravitation began in his 1907 "review" article [3]. He discussed many aspects of relativity there, and proposed that the axiom of special relativity – that the laws of physics are valid in all inertial coordinate systems – be extended to accelerated coordinate systems as well. He called this the Principle of Equivalence. It can be stated in a number of ways; one is that an observer cannot detect the difference between a force field and an accelerated reference frame. In a brilliant *gedanken Experiment* he posited two identical clocks on a line, one in uniform acceleration, the second fixed in a uniform gravitational field. Assuming the two clocks keep the same time, he showed by an entirely mathematical argument that the rate of time of the stationary clock depended solely on the gravitational potential.

Einstein's intuition and determination ultimately led him to a system of partial differential equations coupling two unknown quantities, the metric tensor of the space-time continuum, and the energy-momentum tensor representing the physical source term causing space-time to bend. The energy-momentum tensor is not determined by the Principle of Equivalence, but Einstein did not need it in the two problems he was treating; for by his *gedanken Experiment* he knew the time-component of the metric tensor was to first order a perturbation of the Minkowski metric by the addition of the gravitational potential ϕ. The source was clearly the mass of the Sun, which could be taken as a point source; Einstein argued that the first order approximation was sufficient, and thereby obtained the solution to his problem.

But how do we model intense gravitational fields generated by objects in motion, a binary star, for example, or the collision of two black holes? Or competing fields, such as the electromagnetic and gravitational fields in the interior of a star?

The energy-momentum tensor is fundamental to modeling such

problems. It was uniquely determined in Hilbert's formulation of general relativity as a problem in the calculus of variations on the space-time manifold (see §3). But there was a glitch: it turned out to be the classical Maxwell energy-stress tensor and Maxwell's equations were seen as specific to the electromagnetic field. This leads to an inexplicable "coupling" of the gravitational and electromagnetic fields. Hilbert [8] describes it as "the appearance of electrodynamics as the source of gravitation" (§82, *Die elektrodynamischen Erscheinungen als Wirkung der Gravitation*). Perhaps for that reason, Hilbert's contribution has never been fully appreciated or acknowledged. The variational method is often presented in the literature as an alternative derivation of the equations of general relativity, without an acknowledgement of Hilbert.

The Einstein field equations, with the classical Maxwell stress-tensor for the electromagnetic field as the energy-momentum tensor, are called the Einstein-Maxwell equations. Einstein himself discussed the Maxwell stress-tensor for electromagnetism as an example of a co-variant energy-momentum tensor; but since he identified the geometry of space-time with the gravitational field, he did not propose that tensor for his equations. His skepticism about using the Maxwell stress tensor as the energy-momentum theory in his theory of gravity was justified; it makes no sense to argue that the electromagnetic field generates the gravitational field.

Nevertheless, the Einstein-Maxwell equations occur repeatedly throughout the literature. An early attempt to combine electromagnetism and gravitation in a single relativistic theory appeared in Nordström [15], predating both Hilbert and Einstein. Subsequent to the publication of Einstein's theory and Schwarzschild's construction of the associated line element, Reissner in 1916 [21] and Nordström in 1918 [16] independently obtained metric tensors for the Maxwell-Einstein equations for special choices of the electromagnetic field which purport to represent the gravitational field generated by spherical bodies with both mass and charge. The associated metric is known as the Reissner-Nordström metric [25] Wald, Chapter 6, Problem 3. Metric tensors for the Einstein-Maxwell equations are central the modern theory of black holes (see §12.3 Wald).

This discrepancy between theory and practice is rectified by Theorem 2.1, which proves that Maxwell's equations are universal, independent of the physical nature of the material source generating the field. Theorem 2.2 then follows using Hilbert's variational method combined

3

with the universality of Maxwell's equations to obtain a formulation of general relativity based on Maxwellian fields rather than specific force fields.

There is a second problem with Einstein's formulation of his theory. In [4], he stated

> "According to the general theory of relativity, gravitation occupies an exceptional position with regard to other forces, particularly the electromagnetic forces, since the ten functions representing the gravitational field at the same time define the metrical properties of the space measured."

This premise put the electromagnetic and gravitational fields on incompatible footings – the gravitational field as the Riemannian structure on space-time, the electromagnetic field as a linear field theory on flat space-time. Owing to that viewpoint, Kaluza [9] modified the work of Nordström and Reissner to obtain a unification of gravitation and electromagnetism consistent with Einstein's premise. Like Hilbert, Kaluza's theory was based on a variational method.

Electromagnetism and gravitation certainly differ at the microscopic level – the interaction of the field with matter, for example Snell's laws of reflection and refraction in electromagnetism; and in subatomic interactions, which are the province of quantum mechanics. But Einstein's premise is not supported by any physical experiment in the macroscopic case, and is contradicted by Theorem 2.1.

Relativistic theories of gravitation built on Maxwell's equations and Lorentz invariance go back to Heaviside [7], Lorentz [12] and Poincaré [20]. Theorem 2.1 is *not* to be interpreted as a proof that Maxwell's equations are a valid physical model for gravitation. Quite the contrary: current notions about the role of Maxwell's equations must be re-examined. They are only approximate models of field theory – in the electromagnetic as well as the gravitational case – valid for weak fields. Curvature effects become noticeable in the presence of strong fields,

When Theorem 2.1 is incorporated into Hilbert's variational method, the Maxwell stress-energy tensor associated with the field in question is found to be the unique energy-momentum tensor to be used in the equations of general relativity. Maxwell's equations are thus central to the determination of the energy-momentum tensor.

Universality puts all force fields generated by a material source on the same footing, leading to a simple unified field theory in which

the geometric structure of Einstein's theory is no longer specific any physical force field. Instead, it is the mechanism for the transmission and interaction of physical forces via the bending of the space-time continuum. The semi-Riemannian manifold of space-time thus plays the role of Maxwell's *luminiferous medium* in his original theory of the electromagnetic field.

These results have significant consequences for general relativity. Since the correct energy-momentum tensor has never been used, a good deal of the theory will have to be reexamined. For example, the extensive body of research on gravitational collapse deals with the gravitational field alone, and ignores the competing electromagnetic field (see §4); whereas the stability and ultimate gravitational collapse of a star are the consequence of competing fields, for example the electromagnetic and gravitational fields.

Historically, Maxwell derived his dynamical equations for the electromagnetic field using the empirical laws of electromagnetism, in particular Faraday's law of electromagnetic induction. While the original theory has been considerably streamlined and simplified, all existing derivations are based in one way or another on Faraday's Law.

It is essential to Hilbert's invariant approach to general relativity that Maxwell's equations be invariant under the entire diffeomorphism group of space-time. Minkowski's formulation of them is restricted to flat space time, where the invariance group is the Poincaré group. That formulation is easily extended to obtain the tensorial representation of Maxwell's equations, and was known to both Einstein and Hilbert. The invariant Lagrangian is given in equation (7); and the invariant (tensorial) form of the equations is given in (38), §9.

The vector form of the equations, the coupled system of partial differential equations for **E** and **B**, can be obtained directly from Minkowski's formulation. The vector form of the equations is the simplest form of Maxwell's equations, and is ubiquitous in the literature, e.g. [5], [10], [14]. It is specific to flat space-time and to the electromagnetic field, and its symmetry group is the Poincaré group, the Lorentz group plus translations.

Hodge theory is the natural language of potential theory, and Maxwell's equations have the structure of a potential problem on Minkowski 4 dimensional space-time. Misner, Thorne and Wheeler [14] obtained Maxwell's equations in the language of Hodge theory directly from their vector form. Since those equations are specific to the electromagnetic field, that derivation does not establish univer-

sality. To prove universality, one must derive the equations directly, with the use of Faraday's Law. This is done in §6 by replacing the laws of Faraday and Maxwell-Ampère by Einstein's two assumptions of special relativity. Both laws are found to be valid for all Maxwellian fields; they are not presumed in advance, but follow as mathematical corollaries of special relativity. The derivation is based on 2-forms which correspond to the laws of Maxwell-Ampère and Faraday; and Hodge duality plays a substantive role in the proof.

In order to extend universality to the tensorial and vector forms of the equations, it suffices to show that Minkowski's formulation can be obtained directly from the Hodge form of the equations; this is done in §7.

The drama and excitement of Einstein's discovery is captured in *The Hunt for Vulcan*, by Thomas Levenson [11]. Just as the discovery of Neptune validated Newtonian physics in explaining the perturbations of Uranus' orbit, astronomers postulated the existence of another planet, Vulcan, to explain the perturbations of Mercury. The search for Vulcan proved fruitless, and Einstein finally showed that the problem lay not in the heavens, but in the mathematical model itself – a point to keep in mind in the present search for dark matter, or dark energy.

The Schwarzschild metric is a singular solution of the Einstein field equations – the Euler-Lagrange equations for the "Riemannische Krümmungsinvariant" (the Ricci scalar curvature). Is it a solution to Hilbert's variational problem? It is not even clear that the integral exists, due to the Schwarzschild singularity. The mathematical analysis of such singularities plays a central role in modern cosmology, in particular, the theory of black holes (Wald, Chapters 9 and 12).

Hilbert's paradigm, as realized in Theorem 2.2, opens the door to an alternative mathematical approach to cosmology based on the calculus of variations. Some examples are briefly discussed in §3; but a good deal of mathematical machinery remains to be developed.

2 Statement of Results

A force field $\mathbf{F}$ is said to be generated by a material source consisting of a density ρ and current $\mathbf{J}$, both compactly supported, such that

$$\rho_t + \mathrm{div} \cdot \mathbf{J} = 0; \tag{1}$$

$$\nabla \times \mathbf{F} = \mathbf{J}, \qquad \mathrm{div} \cdot \mathbf{F} = \rho. \tag{2}$$

We assume in addition that the dynamical equations satisfy Einstein's two axioms of special relativity [2] – that the equations are valid in all inertial frames, and that there is a universal velocity c which is the same in all such frames. Force fields satisfying these postulates will be called *Maxwellian fields*.

Theorem 2.1 *In an isotropic, homogeneous medium the electromagnetic field equations are not specific to the electromagnetic field, but rather comprise a mathematical theory of Maxwellian fields, regardless of the physical nature of their source.*

The dynamical field equations are Maxwell's equations, and the laws of Maxwell-Ampère and Faraday are mathematical corollaries of the theory, valid for all force fields, regardless of their physical nature.

Theorem 2.1 is of mathematical interest in that its proof constitutes the first mathematical derivation of Maxwell's equations based on first principles, rather than on empirical laws such as Faraday's. Its physical significance is that it is central to the proof of Theorem 2.2 below. That theorem shows, among other things, that Maxwell's equations for gravitation are necessary for the proper determination of the energy-momentum tensor of Einstein's theory. This will be explained in §3.

Theorem 2.2 *Maxwell's equations couple to Einstein's field equations of general relativity using Hilbert's variational method.*

The energy-momentum tensor is uniquely determined, and the theory of general relativity can be formulated in terms of Maxwellian fields, rather than specific force fields.

The Maxwell potentials corresponding to distinct force fields can be superposed, leading to a unified field theory.

Theorem 2.2 will be proved in §3. A unified field theory is essential in modeling massive structures with internal electromagnetic

7

fields such as a star §4. The theorem shows not only that Maxwell's equations are relevant to gravitation, but also that general relativity is relevant in cases where the electromagnetic field is intense, in the interior of a star, or the early universe, for example.

The proof of Theorem 2.1 makes substantive use of Hodge theory. A summary of the basic ideas are presented in §9. The static equations require the conservation of material, and are obtained in §5 as an application of the Hodge decomposition theorem on $\mathbb{E}^2, \mathbb{E}^3$. Two parameters ϵ, μ arise in the static theory, and in the course of the derivation of the dynamic equations we obtain

$$\mu\epsilon = \frac{1}{c^2}. \tag{3}$$

Equation (3) was central to Maxwell's demonstration that light is an electromagnetic phenomenon. The parameters ϵ and μ are also central to the proof of universality.

By (3) Maxwell's equations in the presence of a material source form a one-parameter family of equations, labeled by μ. The parameter ϵ is connected with Gauss' law for the conservative component of the force field and is given by $4\pi/G$, where G is the strength constant associated with the inverse square law, and is positive for repulsive forces and negative for attractive forces cf. §5. By (3),

$$\mu = \frac{4\pi G}{c^2}. \tag{4}$$

3   Universality of General Relativity

Einstein uses the word *gravitation* in his 1907 paper, but his argument is entirely mathematical and easily extends to any conservative force field. Why then should the Principle of Equivalence apply to the gravitational but not the electromagnetic field? And why should the proposition that the laws of physics be the same in any coordinate system apply to gravity but not to electromagnetism? He says nothing about this, and builds the theory of general relativity on the premise that "gravitation occupies an exceptional position with regard to other forces, particularly the electromagnetic forces . . . " – a premise contradicted by Theorem 2.1.

Einstein [4] shows that invariance, together with the Principle of Equivalence, leads by purely mathematical considerations to field

equations on a semi-Riemannian manifold, with Newton's law of motion replaced by the geodesics. But the energy-momentum tensor is not determined by geometric methods. Einstein introduces that tensor as the source term for the geometric field equations, in analogy with the material source in Maxwell's equations. He discusses two physical examples – the energy-stress tensor of a frictionless adiabatic gas, and the Maxwell stress tensor for the electromagnetic field. As Maxwell's equations are at the time regarded as specific to electromagnetism, while Einstein identifies the geometry of space-time with the gravitational field, he does not propose that tensor in his equations.

The lack of a definitive stress-tensor has led to extensive discussions in the literature of possible energy-momentum tensors, all based on physical arguments, for example [10], [14], [26]. A common approach is to sum the energy-momentum tensors of individual particles. There are extensive discussions in the literature comparing the approaches of Einstein and Hilbert to obtaining the energy-momentum tensor. See especially Chapters 17 and 21 in Misner, Thorne, and Wheeler. Chapter 21 is devoted to a discussion of Hilbert's variational approach. That method is the only mathematical framework known for determining the energy-momentum tensor.

Hilbert's approach to general general relativity [8] is based on an action principle and two invariants, $\mathcal{K}$ and $\mathcal{L}$. The first is the "Riemannische Krümmungsinvariant" (better known as the Ricci scalar curvature R), $\mathcal{K} = \iint K\sqrt{g}\,d\omega$, where

$$ K = g^{jk}K_{jk}, \qquad K_{jk} = \frac{\partial \Gamma^i_{jk}}{\partial x^i} - \frac{\partial \Gamma^i_{ji}}{\partial x^k} + \Gamma^i_{ri}\Gamma^r_{jk} - \Gamma^i_{rk}\Gamma^r_{ji}, \qquad (5) $$

Γ^i_{jk} are the Christoffel symbols; and $\sqrt{g}\,d\omega$ is the invariant volume element, where $g = \det||g_{ij}||$ and $d\omega = dx^1 dx^2 dx^3 dx^4$.

The second is $\mathcal{L} = \iint L\sqrt{g}\,d\omega$, where L is the Lagrangian of a physical system with coordinates q_i and their derivatives $q_{i,j}$, invariant under the group of diffeomorphisms of space-time. Hilbert combines the two invariants as a sum; but in order to conform with Einstein's formulation of the equations, we take $\mathcal{H} = \mathcal{K} - \zeta\mathcal{L}$. (The parameter ζ can be computed up to a pure number by dimensional analysis §8.)

The equations of general relativity are given by the action principle $\delta\mathcal{H} = 0$, where the variations are taken over both the physical variables q_i and the geometric variables consisting of the g_{ij} and the Christoffel symbols. The energy-momentum tensor is obtained from

9

the variational derivative

$$T_{mn} = \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}\, L}{\partial g^{mn}}. \tag{6}$$

Hilbert calculates all the invariants and finds that the simplest case is that in which L is the Lagrangian for Maxwell's equations in free space. Writing $q_i = A_i$, the Maxwell 4-potential, we have

$$L = F_{ij}F^{ij}, \qquad F^{ij} = g^{ik}g^{jl}F_{kl}, \qquad F_{ij} = A_{j,i} - A_{i,j}.$$

The expression L is obtained as an extension to general coordinates of Minkowski's Lagrangian on Minkowski space $\mathfrak{M}^4$ of Maxwell's equations in free space.

Einstein and Hilbert both base their calculations on Minkowski's work; but Einstein extends it to include the source terms, while Hilbert does not. The result is that the energy-momentum tensor obtained by Einstein is specific to the electromagnetic field; while that obtained by Hilbert is specific to Maxwell's equations in free space.

We contend that Maxwell's equations in free space should not be considered specific to the electromagnetic field. When written in the language of Hodge theory, as in §6, they take the form $dF = 0$, $\delta F = 0$ where F is a 2-form on Minkowski four dimensional space-time $\mathfrak{M}^4$. That is, F is a harmonic 2-form, a purely mathematical object. To put this in perspective, take the equation $\Delta \phi = 0$ in $\mathbb{E}^3$. What physical field does the gradient $-\nabla \phi$ represent? We have no information. On the other hand, consider Poisson's equation $\Delta \phi = -\rho/\epsilon$, where ρ is a density and ϵ is the strength parameter associated with Gauss' law. Now $-\nabla \phi$ is the force field induced by ρ.

This ambiguity does not occur in the theory presented this paper, since only Maxwell's equations in the presence of a material source are considered. The associated Lagrangian is (see §7)

$$L = -\tfrac{1}{4}F_{jk}F^{jk} + \mu A_i J^i. \tag{7}$$

The energy-momentum tensor obtained from (6) is calculated explicitly in §8:

$$T_{mn} = \frac{1}{2}(-F_{mj}F_n{}^j + \tfrac{1}{4}F_{ij}F^{ij}g_{mn})$$
$$+ \mu(A_m J_n - \tfrac{1}{2}g_{mn}A_i J^i). \tag{8}$$

Up to a scalar multiple, the first term, quadratic in the F_{ij}, is the classical stress-energy tensor for a general Maxwellian field. (By

(38) μ^2 is implicitly contained in that term.) The Maxwell stress-tensor (without the μ) appears throughout the literature in general relativity, but it is always specific either to the electromagnetic field, or to Maxwell's equations in free space.

Though Einstein had extended Minkowski's treatment to include the source terms, he did not make use of Hilbert's variational method and so did not compute the Lagrangian (7). Consequently, the second term in (8) does not appear in his presentation.

We note a useful variant of Hilbert's original method. The variation $\delta\mathcal{L}$ with respect to the field variables A_i, $A_{i,j}$ constitutes a variation under the constraint that the geometry of space-time is fixed; and the equations so obtained are Maxwell's equations on a fixed curvilinear space-time. The field equations of general relativity are obtained by relaxing this constraint, introducing $\mathcal{K}$, and allowing variations over both field and geometric variables g_{ij} and Γ^i_{jk}.

This approach by-passes the Principle of Equivalence, as well as the postulate that space-time bends in the presence of a force field, by showing that space-time *must* bend in the presence of a Maxwellian field unless it is constrained not to do so. This proves the universality of general relativity, no further postulates being necessary.

We have so far proved the first two statements in Theorem 2.2; the proof of the unified field theory is very simple. Consider a family of Maxwellian fields, indexed by κ, generated by material sources J_κ; A_κ the corresponding Maxwellian potentials; $F_{\kappa,jk}$ the force fields; L_κ the Lagrangians; and ζ_κ the corresponding Lagrange multipliers. Putting $L = \sum_\kappa \zeta_\kappa L_\kappa$ and using (6) we have

$$T_{mn} = \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}\, L}{\partial g^{mn}} = \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}\, \sum_\kappa \zeta_\kappa L_\kappa}{\partial g^{mn}} = \sum_\kappa T_{\kappa,mn}, \qquad (9)$$

where

$$T_{\kappa,mn} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial g^{mn}} \sqrt{g}\, \zeta_\kappa L_\kappa.$$

We assume the linear physical fields interact only via the geometric field. ∎

4 Negative Energy

Maxwell briefly considered the possibility of applying his ideas to gravitation in his original tract, but abandoned it because of the implica-

tions of the negative energy of an attractive field. He had proposed a *luminiferous medium*, the aether, as a mechanical medium to transport the energy of the waves. Since the field energy of an attractive force field is negative, and there is no lower bound on that energy, the aether would have to have an infinite amount of energy, he reasoned, concluding ([13] p. 493)

> "As I am unable to understand in what way a medium can possess such properties, I cannot go any further in this direction in searching for the cause of gravitation."

Many of the arguments against Maxwell's equations for gravitation are based on that brief passage in his treatise. See Pais Chapter 9 [17], and the discussion in [22]). But Einstein scotched Maxwell's aether with his two simple axioms of special relativity; and the negative field energy is neither an obstruction to the derivation of the equations nor to their coupling to the field equations of general relativity. In fact, the negativity of the classical field energy is essential to an understanding of astrophysics.

In [22] §8 we constructed a simple example to show that as mass aggregates under the influence of gravity, the field energy goes to $-\infty$. We considered the family of Gaussian mass distributions

$$\rho_\sigma(x) = \left(\frac{\sigma}{\pi}\right)^{3/2} \exp(-\sigma|x|^2), \qquad x \in \mathbb{E}^3.$$

These densities all have total mass 1, while the total energy,

$$\mathcal{E} = -\iint \frac{\rho(x)\rho(y)}{8\pi|x-y|}dxdy, \qquad x, y \in \mathbb{E}^3,$$

can be computed explicitly and is given by

$$\mathcal{E}_\sigma = -\frac{1}{16\pi^2}\left(\frac{\sigma}{2\pi}\right)^{1/2}.$$

As $\sigma \to \infty$, the mass distribution collapses to the origin ($\rho_\sigma \to \delta(x)$) while $\mathcal{E}_\sigma \to -\infty$

The theory of distributions treats the notion of a point mass rigorously in linear partial differential equations; but the above result shows that the gravitational energy of a point mass is $-\infty$, regardless of the amount. In the construction of the Schwarzschild line element the energy-momentum tensor is assumed to vanish; and this implies that the field source is a Dirac delta function with mass M located at

the origin. Thus a nonlinear system of partial differential equations is being forced by a source with infinite negative energy. Is the resulting singularity, the event horizon, a physical reality, or a simply a spurious artifact of the assumption of a point mass?

A star forms by the aggregation of gases, typically hydrogen, under gravitation. As the volume of the aggregate decreases, its temperature and pressure increase until thermonuclear fusion commences and the contraction is halted, the pressures of the gravitational and electromagnetic fields coming into equilibrium. In the course of contraction, the decrease of energy in the gravitational field is compensated by the increase in thermal energy. This shows that star formation, and more generally gravitational collapse, cannot be modeled by the gravitational field alone; other forms of energy must be included.

There is an extensive body of theoretical literature on gravitational collapse, but it is limited by the fact that only the gravitational field is accommodated. A physically realistic model of the dynamics of a massive body requires the combination of two or more distinct fields – mass and energy for example – a so-called unified field theory, which was never found under the old paradigm.

The complementary processes of gravitational collapse and explosion also manifest themselves in the recent observation of a "gravitational wave" by the Laser Interferometer Gravitational Observatory, announced on February 11, 2016 (www.ligo.caltech.edu):

> "LIGO scientists estimate that the black holes for this event were about 29 and 36 times the mass of the sun, and the event took place 1.3 billion years ago. About 3 times the mass of the sun was converted into gravitational waves in a fraction of a second – with a peak power output about 50 times that of the whole visible universe."

What is the mechanism that produces the outgoing "gravitational" wave? Here is one way to look at the process. The coalescence of the two black holes results in a reduction of the (already negative) gravitational energy of the system. That decrease is converted into energy, partly the heat, electromagnetic, and kinetic energy of the newly formed black hole, but also into the space-time continuum. (Einstein [4] denoted the energy-momentum tensor of space-continuum by $t_{\mu\nu}$.) Space-time can thus be viewed as the *luminiferous medium* initially proposed by Maxwell – the "mechanical system" that transports energy. We argued earlier that the theory of relativity should not be

seen as specific to the gravitational field. This allows us to view the outgoing wave as one in the fabric of space-time. While the energy source came initially from the gravitational field, the resulting outgoing wave is purely geometric in nature; there are no longer material sources such as charge or mass involved.

5 Maxwellian Fields

In this section we obtain the static equations of a Maxwellian field; the discussion in this section is restricted to $\mathbb{E}^3$. The Hodge star operation is determined by the invariant volume element $dv = dx^1 \wedge dx^2 \wedge dx^3$; the details are given in §9. Thus, $dx^i = dx^j \wedge dx^k$, where i, j, k are in cyclic order. The co-adjoint operator $\delta = *^{-1}d*$ is then the formal adjoint of the exterior derivative d. On $\mathbb{E}^3$, $* = *^{-1}$, so $\delta = *d*$. For proofs and further details concerning Hodge theory, see §9 and [22].

Let $F = \mathbf{F} \cdot d\mathbf{x} = F_i dx^i$ and $J = \mathbf{J} \cdot d\mathbf{x} = J_i dx^i$, where $\mathbf{F}$ is the force field and $\mathbf{J}$ the current in (1) and (2). Then $*F = F_i dx^j \wedge dx^k$. In the stationary case, $\rho_t = 0$, hence $d * J = 0$, and equations (2) take the form

$$dF = *J, \qquad \delta F = \rho. \tag{10}$$

The Hodge decomposition of the Hilbert spaces $\Lambda_p(\mathbb{E}^3)$ of square integrable p-forms (Theorem 9.2) implies that

$$\Lambda_1(\mathbb{E}^3) = \mathfrak{E} \oplus \mathfrak{H}, \qquad \Lambda_2(\mathbb{E}^3) = \mathfrak{B} \oplus \mathfrak{D}, \tag{11}$$

where $\mathfrak{E}$ and $\mathfrak{H}$ are the subspaces of exact and co-exact 1-forms; while $\mathfrak{B}$ and $\mathfrak{D}$ are respectively the subspaces of exact and co-exact 2-forms. By (11) we have $F = E + H$, and $*F = B + D$, where E, B are exact and H, D are co-exact. We may therefore write $E = -d\phi$ (the minus sign is by convention) and $B = dA$; and so from (10) we obtain $dF = *J = dH$, and $\delta F = \rho = *dD$, hence $dD = \rho \, dv$.

The four differential forms, E, H, B and D form the basic structure of the electromagnetic field equations: E and H are the electric and magnetic fields, B the magnetic induction, and D Maxwell's electric displacement. The equation $dH = *J$ is Ampère's law. We use the nomenclature of electromagnetism in the general case as well to orient the reader. But the reader should keep in mind that these results are not specific to the electromagnetic field.

Theorem 5.1 *There exist parameters* $\epsilon = \epsilon(\mathbf{x}, \mathbf{n}, E)$, $\mu = \mu(\mathbf{x}, \mathbf{n}, H)$, $\mathbf{n}$ *a vector situated at* $\mathbf{x} \in \mathbb{E}^3$, *such that*

$$D = \epsilon * E, \qquad B = \mu * H \qquad\qquad (12)$$

These two identities are called constituency laws; in [22] they were taken as postulates.

Proof: Writing $D = \mathbf{D} \cdot d\mathbf{S}$ and $*E = - * d\phi = -\nabla\phi \cdot d\mathbf{S}$ we see that the pair of differential forms are simply the projections of the axial vector field $\mathbf{D}$ and the polar field $-\nabla\phi$ along the normal line to the surface element $d\mathbf{S}$ at $\mathbf{x}$. Hence they are *defacto* scalar multiples of one another, the scalar multiple being a function of $\mathbf{x}$, the normal $\mathbf{n}$ and possibly the field strength E: $\epsilon = \epsilon(\mathbf{x}, \mathbf{n}, E)$.

In the case of B, H write $B = \mathbf{B} \cdot d\mathbf{S}$, $*B = \mathbf{B} \cdot d\mathbf{x}$ and $H = \mathbf{H} \cdot d\mathbf{x}$. Replacing $d\mathbf{x}$ by $\mathbf{n} = d\mathbf{x}/ds$, where s is the arc length along the normal curve to a family of surface elements $d\mathbf{S}$, we see that $*B$ and H are projections of the axial vector $\mathbf{B}$ and polar vector $\mathbf{H}$ along that curve, hence scalar multiples of one another. As above, $*B = \mu H$ where $\mu = \mu(\mathbf{x}, \mathbf{n}, H)$. ∎

The medium is *isotropic* if both ϵ and μ are independent of $\mathbf{n}$, and *homogeneous* if they are independent of $\mathbf{x}$. In free space, they are both, hence they are constant by the assumption of frame invariance.

The dependence of ϵ and μ on the medium account for the interaction of the field with matter. Snell's laws of reflection and refraction in electromagnetic waves, for example, can be derived mathematically when ϵ and μ have jump discontinuities, see Stratton, [24]. If ϵ and μ depend on the fields, as in (12), the field equations are nonlinear; this is precisely the situation in nonlinear optics. Non-local dependence in the case of electromagnetism is necessary to explain such phenomena as chromatic aberration, rainbows, and prisms, since these are manifestations of the dependance of the speed of light on the *frequency* of the field (see (3))

Using (36) the static equations are

$$E = -d\phi, \qquad dD = *\rho, \qquad E = *\frac{1}{\epsilon}D$$

$$\delta E = - * d * d\phi = *d\frac{1}{\epsilon} * D \tag{13}$$

$$B = dA, \qquad dH = *J, \qquad B = \mu * H;$$

$$\delta B = \delta\,dA = *d * dA = *d\,(\mu H). \tag{14}$$

The equations (13) and (14) are in "divergence form"; this form must be used in cases where ϵ and μ are discontinuous.

In a homogeneous medium, where ϵ and μ are constant, the equations simplify to

$$\delta E = - * d * d\phi = \frac{\rho}{\epsilon}, \qquad \delta B = *d * dA = \mu J. \tag{15}$$

Proposition 5.2 *The relation $D = \epsilon * E$ is equivalent to Gauss' Law together with Newton's inverse square law for the exact component of the field. The field is repulsive or attractive according as ϵ is positive or negative.*

Proof: The displacement D due to a point source q at the origin is [22]

$$D = \frac{q}{4\pi}\frac{x_j}{r^3}\,dx^k \wedge dx^l. \tag{16}$$

By the first relation in (12), it follows that

$$E_i = \frac{q}{4\pi\epsilon}\frac{x_i}{r^3}.$$

Writing $\hat{\mathbf{r}} = \mathbf{r}/r$ the force on a point source of strength q' at $\mathbf{r}$ is then

$$\mathbf{E} = \frac{Gqq'}{r^2}\hat{\mathbf{r}}, \qquad G = \frac{1}{4\pi\epsilon}. \tag{17}$$

The "lines of force" exit or enter the region bounded by S, according as $\mathbf{E} \cdot \mathbf{n}$ is positive or negative. If ρ (hence q) is taken to be non-negative in both cases, then ϵ is positive or negative according as the force is repulsive or attractive. Equation (17) is the classical inverse square law when G is the negative of the Cavendish constant in the case of gravity, and the Coulomb constant in the case of the electrostatic field.

6 The Dynamical Equations

Special relativity is efficiently encoded by formulating the dynamical equations on 4 dimensional Minkowski space-time $\mathfrak{M}^4$, obtained from $\mathbb{E}^4$ by setting $x^4 = ict$. The Lorentz group is obtained directly from the rotation group $SO(4)$ on $\mathbb{E}^4$; and the star operation is obtained by replacing x_4 with ict in (19), (20), (36) in §9. The conventions for Maxwell's equations on $\mathfrak{M}^4$ in the general case are identical with those in Stratton's classic treatise on electromagnetism [24].

By Theorem 9.2 the Hodge star operation on $\mathfrak{M}^4$ interchanges exact p forms with co-exact $4 - p$ forms; hence it maps Λ_2 to itself. The Hodge duality is

$$(A, B) = \frac{1}{ic} \int_{\mathfrak{M}^r} A \wedge *B. \tag{18}$$

We begin by combining the density ρ and current $\mathbf{J}$ into a single source term, the 1-form $J = J_i dx^i$, $1 \leq i \leq 4$ on $\mathfrak{M}^4$. Since the entries in the Lorentz matrices are pure numbers, all components of J must have the same dimension. That, plus the conservation law (1), implies that $J_4 = ic\rho$ and $\delta J = 0$

Similarly, we combine the potentials ϕ and $\widehat{A} = \mathbf{A} \cdot d\mathbf{x}$ from the static case into a single 1-form $A = A_i dx^i \in \Lambda_1(\mathfrak{M}^4)$. By the Hodge decomposition of $\Lambda_1(\mathfrak{M}^4)$, we may write $A = d\psi + \delta\Phi$, but since the field is obtained from dA, the term $d\psi$ has no effect; hence we may assume that $\delta A = 0$. We take $A_4 = i\phi/c$ for reasons that will be explained shortly.

The Hodge star operation on $\mathbb{E}^4$ associated with the oriented volume element $dv \wedge dx^4$ is

$$* dx^j = dx^k \wedge dx^l \wedge dx^4 \qquad\qquad * dx^4 = -dv \tag{19}$$

$$* dx^j \wedge dx^k = dx^l \wedge dx^4 \qquad * dx^j \wedge dx^4 = dx^k \wedge dx^l \tag{20}$$

$$* dv = dx^4, \qquad\qquad * dx^j \wedge dx^k \wedge dx^4 = -dx^l. \tag{21}$$

Here and below j, k, l run from 1 to 3 in cyclic order. Note that $** = (-1)^p id$ on $\mathbb{E}^4$.

The Hodge star operation on $\mathfrak{M}^4$ is then given by replacing x^4 by ict in the relations above. The volume element $dv \wedge dx^4$ is denoted by

17

$d\omega$. Maxwell's equations in the language of Hodge theory then take the form of a potential problem on $\mathfrak{M}^4$,

$$F = dA, \quad \delta F = -\mu J. \tag{22}$$

In the specific case of the electromagnetic field, F is the Faraday 2-form $E \wedge dt + B$. Equations (22) (without the μ) are obtained in [14] starting with Maxwell's equations in vectorial form.

The first equation in (22) is equivalent to Faraday's Law of electromagnetic induction (see [22]), and therefore cannot be taken as an axiom in the general case. We proceed instead as follows. (Recall that E and B are respectively general 1- and 2-forms, not specific to the electromagnetic field.)

Lemma 6.1 *Every exact 2-form can be written as*

$$F = \frac{E}{ic} \wedge dx^4 + B, \tag{23}$$

where $E = E_j dx^j$, $B = B_j dx^k \wedge dx^l$ is exact, and (j, k, l) run from 1 to 3 in cyclic order.

Proof: The exact forms in Λ_2 are given by $F = dA$, where $A = A_j dx^j$, and

$$dA = \sum_{j < k \leq 3} \left(\frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k} \right) dx^j \wedge dx^k$$

$$+ \sum_{j=1}^{3} \left(\frac{\partial A_4}{\partial x^j} - \frac{\partial A_j}{\partial x^4} \right) dx^j \wedge dx^4$$

The result follows by putting

$$B_i = \frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k}, \qquad \frac{E_j}{ic} = \left(\frac{\partial A_4}{\partial x^j} - \frac{\partial A_j}{\partial x^4} \right). \qquad \blacksquare \tag{24}$$

The reason for the choice of A_4 is now apparent: the second equation in (24) gives $E_j = -\partial\phi/\partial x^j$ in the special case when the solenoidal component vanishes.

In [22] Faraday's Law was reformulated as a theorem, essentially that given any $E = E(x, t)$ there is an exact 2-form B such that $E \wedge dt + B$ is exact. But two additional axioms were necessary for the derivation of Maxwell's equations: (3) and the extension of (12)

obtained by formally replacing the star operation on $\mathbb{E}^3$ by that on $\mathfrak{M}^4$:

$$*D = \epsilon\, E \wedge dx^4, \qquad *B = \mu H \wedge dx^4. \tag{25}$$

These dualities, together with (3), were taken as axioms and used to prove that

$$*F = \mu G, \qquad dG = *J, \tag{26}$$

where F was the Faraday 2-form. The duality (26) shows that G is co-exact, and implies (22).

We now show that (3) and (25) can be derived from the static constitutive laws (12) alone, which were obtained mathematically. The condition $\delta J = 0$ implies that $d*J = 0$, hence that there is a 2-form G such that $dG = *J$ locally. We write $G = (H \wedge dx^4 - icD)$, where $H = H_j dx^j$ and $D = D_j dx^k \wedge dx^l$ are chosen so that $dG = *J$. In the language of vector analysis this equation gives the Maxwell-Ampère equation of the electromagnetic field [22].

There is no obstruction to solving the equations for G, but the 2-form is not necessarily co-exact.

Lemma 6.2 *Let F be an exact 2-form, and let G satisfy $dG = *J$. Then $*F = \mu G$ if and only if both (3) and (25) are satisfied.*

Proof: If both (3) and (25), then

$$* F = * \left(\frac{E}{ic} \wedge dx^4 + B \right) = \frac{1}{ic\epsilon} D + \mu H \wedge dx^4$$

$$= \mu \left(H \wedge dx^4 + \frac{1}{ic\epsilon\mu} D \right) = \mu\, G.$$

Conversely, if $*F = \mu G$ then

$$* \left(\frac{E}{ic} \wedge dx^4 + B \right) = \mu(H \wedge dx^4 - icD);$$

hence $*B = \mu H \wedge dx^4$ and $*E \wedge dx^4 = \mu c^2 D$, and (25) are established. It follows that

$$E_j dx^j \wedge dx^4 = E \wedge dx^4 = * * E \wedge dx^4 = \mu c^2 * D$$

$$= \mu c^2 * D_j dx^k \wedge dx^l = \mu c^2 D_j dx^j \wedge dx^4.$$

When the fields are stationary, $D_j = \epsilon E_j$ by (12); and (3) follows. $\blacksquare$

We can now complete the proof of Theorem 2.1. Given a compactly supported source J, we first solve the system of equations $\Box A = -\mu J$, $\delta A = 0$. The equations $\Box A_i = -\mu J_i$ are obtained globally without difficulty, since the equations are hyperbolic and J is compactly supported. The Lorentz condition, $\delta A = 0$, can always be obtained by a gauge transformation if necessary, that is, by replacing A by $A + d\psi$ where ψ is determined by the hyperbolic equation $\Box\psi = -\delta A$.

Now put $F = dA$ and $G = \mu^{-1}*dA$. Then F is exact, G is co-exact, and $dG = *J$. Moreover $*F = \mu G$; hence the dynamic constitutive laws (26) as well as (3) follow directly from Lemma 6.2; and Maxwell's equations (22) are obtained. ∎

7 Hodge to Minkowski

In addition to universality, Maxwell's equations have the remarkable property that the Lagrangian (7) is invariant with respect to the entire group of general coordinate transformations. These two properties, universality and invariance, are the basis underlying Theorem 2.2. Hodge theory, the basis for the proof of universality, is restricted to transformation groups leaving the measure $d\omega$ invariant; but since (7) is tensorial, it suffices to derive it in the case of special relativity.

The Lagrangian for Maxwell's equations (22) is [22]

$$S = \frac{1}{ic}\iiint \frac{1}{2}F \wedge *F + \mu\, A \wedge *J = \frac{1}{2}(F,F) + \mu(A,J), \qquad (27)$$

where (F,G) is the Hodge duality (18). In fact, denoting the variation of S by $\dot{S}$, we get $\dot{S} = ic[(\dot{F},F) + \mu\,(\dot{A},J)]$. Noting that $(\dot{F},F) = (d\dot{A},F) = (\dot{A},\delta F)$, we have $\dot{S} = ic\,(\dot{A},\delta F + \mu J)$. Letting $\dot{A}$ vary over all admissible variations, and assuming these to be a dense set, we get $\delta F + \mu J = 0$, which, together with the relation $F = dA$, comprise Maxwell's equations.

From (23) we have

$$F \wedge *F = \left(\frac{E}{ic} \wedge dx^4 \right) \wedge * \left(\frac{E}{ic} \wedge dx^4 \right) + B \wedge *B$$

$$= \left(\sum_j B_j B_j - \frac{E_j}{c} \frac{E_j}{c} \right) d\omega, \qquad (d\omega = dv \wedge dx^4),$$

$$= - \left(B_j B^j + \frac{1}{c^2} E_j E^j \right) d\omega,$$

where $B^j = -B_j$ and $E^j = E_j$ are the contravariant tensors obtained by raising the indices using the Minkowski metric tensor $\eta = \operatorname{diag}(-1, -1, -1, 1)$.

From (23) we obtain $F_{kl} = B_j$, $F_{k4} = E_k/ic$. Replacing x_4 by x_4/i and noting that $B^j = F^{kl}$, etc. we find

$$(F, F) = -\frac{1}{2} F_{jk} F^{jk}.$$

Since F_{jk} is a covariant tensor, its contraction is an invariant. The source term (A, J) is transformed in the same way, so (7) follows from (27).

8   The Energy-Momentum Tensor

In this section we obtain the explicit form of the energy-momentum tensor (8) directly from (6) and (7). Since the the action depends only on the g_{ij} and not their derivatives, the calculation reduces to

$$T_{mn} = \mathcal{C}_1(g)(-\tfrac{1}{4} F_{ij} F_{kl}) + \mathcal{C}_2(g)(\mu A_i J_j),$$

where

$$\mathcal{C}_1 = \frac{1}{4} \left(\frac{1}{\sqrt{g}} \frac{\partial}{\partial g^{mn}} \sqrt{g} g^{ik} g^{jl} \right), \quad \mathcal{C}_2 = \left(\frac{1}{\sqrt{g}} \frac{\partial}{\partial g^{mn}} \sqrt{g} g^{ij} \right).$$

These two coefficients can be simplified using a standard identity from matrix theory:

$$\left(\frac{\partial}{\partial g_{ij}} \right) \log g = g^{ji}, \qquad \left(\frac{\partial}{\partial g^{ij}} \right) \log g = -g_{ji}. \tag{28}$$

21

The second identity follows from the first by replacing the matrix g_{ij} by its inverse g^{ij}. The first follows from the Laplace expansion of the determinant $g = \sum_i g_{ij}\Delta^{ij}$, where Δ^{ij} is the cofactor of g_{ij}. Since Δ^{ij} is independent of g_{ij}, $\partial g/\partial g_{ij} = \Delta^{ij} = g^{ji}g$, which is equivalent to the first identity.

Using (28) we obtain

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial g^{mn}}\left(\sqrt{g}g^{ij}\right) = \frac{\partial g^{ij}}{\partial g^{mn}} + \frac{1}{2}g^{ij}\frac{\partial \log g}{\partial g^{mn}} = \delta_m^i\delta_n^j - \frac{1}{2}g^{ij}g_{mn}. \quad (29)$$

Similarly

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial g^{mn}}\left(\sqrt{g}g^{ik}g^{jl}\right) = \delta_m^i\delta_n^k g^{jl} + \delta_m^j\delta_n^l g^{ik} - \frac{1}{2}g^{ik}g^{jl}g_{mn}. \quad (30)$$

Equation (8) then follows.

The field equations of general relativity obtained by Hilbert's variational approach are therefore

$$K_{jk} - \tfrac{1}{2}Kg_{jk} = \zeta T_{jk}, \quad (31)$$

where T_{jk} is given by (8). The parameter ζ is necessary to match the differing dimensions of the two sides of (31). It can be determined up to a numerical constant by dimensional analysis.

Let ℓ, τ, m, q denote respectively the units of length, time, (inertial) mass and source, with ℓ and τ restricted so that $c = \ell/\tau$. Then $[dx^i] = \ell$, and $[d] = 1$, that is $[d\eta] = [\eta]$ for any differential form. Also, $[\delta] = [*^{-1}d*] = [*^{-1}][d][*] = 1$.

To compute the dimension of T_{jk} we first observe that the two terms in (8) are of the same dimension: from (22) it follows that $[F] = [A] = [\mu][J]$; and therefore that $[F]^2 = [A]^2 = [\mu AJ]$, hence that $[F_{jk}^2] = [\mu A_i J^i] = [T_{jk}]$. Also $[J] = [J_i dx^i] = [\rho][v]\ell = [q](\ell\tau)^{-1}$, $[J_i] = q/\ell^2\tau$. From the inverse square law (17) we deduce that $[G] = m\ell^3/\tau^2 q^2$. Hence by (4) it follows that $[\mu] = m\ell/q^2$. From the equation $\Box A_i = -\mu J_i$ we obtain $[A_i] = [\mu J_i]\ell^2 = m\ell/q\tau$.

Since the g_{ij} are dimensionless, the left side of (31) is of dimension ℓ^{-2} by (5); hence

$$\frac{1}{\ell^2} = \zeta[\mu A_i J^i] = \zeta\frac{m^2}{q^2\tau^2}, \qquad \zeta = \left(\frac{q}{m}\right)^2\frac{1}{c^2}. \quad (32)$$

Einstein calculates ζ directly by a perturbation method, based on the presumption that the dynamics of a slowly moving particle in a weak field in general relativity is approximated by Newtonian mechanics (§21, Newton's Theory as a Weak Approximation).

9 Appendix: Hodge Theory

Vector analysis is today's *lingua franca* of applied mathematics and physics; but differential forms already made a cameo appearance in Maxwell's treatise, for example,

$$\left(\frac{dH}{dy} - \frac{dG}{dz}\right) dy dz$$

is "the number of lines of magnetic force which pass through the area $dy dz$."

Here is his introduction of the electric displacement and its currents:

> "(55) Electrical displacement consists in the opposite electrification of the sides of a molecule or particle of a body which may or may not be accompanied with transmission through the body. Let the quantity of electricity which would appear on the faces $dy.dz$ and an element dx, dy, dz cut from the body be $f.dy.dz$, then f is the component of the electric displacement parallel to x. We shall use f, g, h to denote the electric displacements parallel to x, y, z respectively.
>
> The variation of the electrical displacement must be added to the currents p, q, r to get the total motion of electricity . . . " [13] p.480

The exterior calculus of differential forms, developed by Grassman, Poincaré [19] (see Chapter 22) and later Cartan [1], is comprised of two operations, the exterior derivative d, and the wedge product $\wedge$, acting on the linear spaces Λ_p of p-forms – completely anti-symmetric covariant tensors of rank p. Both operations are tensorial – that is, they are the same in all coordinate systems [23] §§18,19. The exterior calculus is the natural language of integration theory.

On $\mathbb{E}^n$ the Hodge star operation $* : \Lambda_p \to \Lambda_{n-p}$ associated with the standard volume element $dv_n = dx^1 \wedge \cdots \wedge dx^n$ is obtained by defining its action on the basis forms $\eta = dx^{i_1} \wedge \cdots \wedge dx^{i_p}$ so that $\eta \wedge *\eta = dv_n$. An inner product on Λ_p, called the Hodge duality, is then given by

$$(\xi, \eta) = \iiint_{\mathbb{E}^n} \xi \wedge *\eta, \qquad \xi, \eta \in \Lambda_p. \tag{33}$$

This duality implicitly defines a formal adjoint of d, called the *co-derivative*, $\delta : \Lambda_{p+1} \to \Lambda_p$, defined by

$$(d\,\xi, \eta) = (\xi, \delta\eta) \qquad \xi \in \Lambda_p, \ \eta \in \Lambda_{p+1}, \tag{34}$$

where ξ, η are smooth and compactly supported on $\mathbb{E}^n$. The co-derivative acting on Λ_p is given by

$$\delta_p = (-1)^p *^{-1} d *. \tag{35}$$

(This result was proved in [22], Proposition 2.1. The first statement there must be corrected by replacing δ_p by δ_{p+1} on the left side, and then replacing p by $p-1$ on both sides.)

Using those results for δ_p we find

$$\delta_p = (-1)^p * d * \ \text{ on } \ \mathbb{E}^3; \ \text{ and } \ \delta = * d * \text{ on } \mathbb{E}^4. \tag{36}$$

Under a smooth coordinate transformation $x^i \to x'^i$ on $\mathbb{E}^n$, $dv_n \to dv'_n = J\,dv_n$, where J is the Jacobian $\partial(x'^1, \ldots x'^n)/(x^1, \ldots x^n)$. The group of volume preserving transformations on $\mathbb{E}^n$ are those for which $J = 1$. Since the Hodge star operation is defined by its action on the basis forms dx^j, $dx^j \wedge dx^k, \ldots$, an invariant Hodge star operation is defined by the relation $*\xi' = (*\xi)'$. The Hodge star operation on $\mathbb{E}^n$ is thus invariant under the group of volume preserving transformations.

The same arguments apply on $\mathfrak{M}^4$. Moreover, d and $\wedge$ are tensorial – they are invariant under the entire diffeomorphism group [23]. Accordingly, by (22) we see that Maxwell's equations are invariant under the entire group of volume preserving transformations, not just the Lorentz group.

It is well known that Maxwell's equations can be written in curvilinear coordinates on a *fixed* geometry of space-time. (See [10]) We can extend the above discussion to this situation. We begin by noting that the star operation is an artifact of integration theory on differentiable manifolds, independent of its geometry. (After all, the initial applications of Hodge theory were to algebraic topology.)

On a general Riemannian manifold the volume form $\sqrt{g}\,dv_n$, where $g = \det\|g_{ij}\|$ is the determinant of the metric tensor, is invariant under the entire diffeomorphism group. In the specific case of general relativity, $n = 4$ and this volume element was used by both Einstein and Hilbert. It is invariant under the subgroup which preserves $d\omega$ (the group of volume preserving transformations on $\mathfrak{M}^4$); hence $\sqrt{g}$ is also invariant under this subgroup, and therefore so is $\sqrt{g}\,d\omega$.

In curvilinear coordinates on $\mathfrak{M}^4$ the invariant star operation $*_g$ and its inverse are

$$*_g = \sqrt{g} * \qquad (*_g)^{-1} = \frac{1}{\sqrt{g}} *^{-1}, \tag{37}$$

where $*$ is that on $\mathbb{E}^4$.

Maxwell's equations in curvilinear coordinates are obtained by replacing δ by δ_g

$$F = dA, \qquad \delta_g F = \frac{1}{\sqrt{g}} *^{-1} d \sqrt{g} * F = -\mu J.$$

These equations are valid when the geometry of space-time is fixed (up to the group of volume-preserving coordinate transformations). Maxwell's equations in general coordinates are

$$F_{ij} = A_{j,i} - A_{i,j}, \qquad \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}\, F^{ij}}{\partial x^j} = F^{ij}{}_{;j} = -\mu J^i. \tag{38}$$

In the case of the vector form of Maxwell's equations in a vacuum there is an extensive literature on additional geometrical symmetries that go beyond the Poincaré group and include the conformal group. Fushchych and Nikitin [6] have extended this class to additional non-geometric symmetries; and Pohjanpelto [18] has shown that the first order generalized symmetries of the equations form a closed Lie algebra isometric to $so(6, C)_R \otimes \mathfrak{g}$, where $\mathfrak{g}$ is an abelian algebra.

A differential form ξ is *exact* if $\xi = d\phi$, and *co-exact* if $\xi = \delta\psi$. The Hodge star interchanges exact and co-exact forms. Similarly ξ is *closed* if $d\xi = 0$ and *co-closed* if $\delta\xi = 0$. A differential form is called harmonic if it is both closed and co-closed, i.e. $d\xi = \delta\xi = 0$.

The Hodge decomposition on a compact manifold states that every (smooth) p-form ω can be decomposed as $\omega = d\xi + \delta\eta + \alpha$, where α is harmonic. Since $(d\xi, \delta\eta) = (\xi, \delta^2\eta) = 0$, etc. it is clear that the subspaces of exact, co-exact, and harmonic forms are mutually orthogonal; hence the Hodge decomposition can also be written as

$$\Lambda_p = [\operatorname{im} d] \oplus [\operatorname{im} \delta] \oplus [\ker d \cap \ker \delta]. \tag{39}$$

We now denote by Λ_p the Hilbert space of all measureble p-forms ξ for which $(\xi, \xi) < +\infty$. The formulation of Hodge's theorem above leads to a simple proof of its extension to these Hilbert spaces based on the well-known orthogonal decomposition theorem.

Proposition 9.1 *Let $\mathcal{H}$ be a Hilbert Space, $\mathcal{E}$ a subspace, and $\mathcal{E}^\perp$ its orthogonal complement. Then every element $F \in \mathcal{H}$ can be uniquely decomposed as $F = E + G$, where $E \in \mathcal{E}$ and $G \in \mathcal{E}^\perp$. We write $\mathcal{H} = \mathcal{E} \oplus \mathcal{E}^\perp$.*

To account for the lack of differentiability of the elements of Λ_p, the decomposition must incorporate the notion of weak derivatives. We say that $dF = G$ in the weak sense if $(dF, \eta) = (G, \delta\eta)$ for all smooth differential forms $\eta \in \Lambda_{p+1}$ with compact support. A similar definition applies to the equation $\delta F = G$. Note that if $dF = G$ in the weak sense and F is C^1, then G is continuous and $dF = G$ in the strong sense.

Theorem 9.2 *The Hilbert space $\Lambda_p(\mathbb{E}^n)$ decomposes into the direct sum of exact and co-exact forms: $\Lambda_p = [d\Lambda_{p-1}] \oplus [\delta\Lambda_{p+1}]$, where $[d\Lambda_{p-1}]$ denotes the L^2 closure of the linear set $\{dA : A \in \Lambda_{p-1}\}$, etc. and by default, $[d\Lambda_n] = [\delta\Lambda_0] = 0$. Thus every differential form $F \in \Lambda_p$ can be written as $F = dA + \delta\Phi$, where $A \in \Lambda_{p-1}$ and $\Phi \in \Lambda_{p+1}$.*

The Hodge star operation interchanges exact and co-exact differential forms. Specifically, since $** = \pm id$,

$$*[d\Lambda_p] = [\delta\Lambda_{n-p}]. \tag{40}$$

Proof: A harmonic form A satisfies $dA = \delta A = 0$, hence $\Delta A = 0$, where $\Delta = d\delta + \delta d$. Writing $A = A_\alpha \wedge dx^\alpha$ where $\alpha = \{i_1 < \ldots i_p\}$, it follows that $\Delta A_\alpha = 0$ as well. Therefore each A_α is a harmonic function in the ordinary sense on $\mathbb{E}^n$. In order that they be square integrable they must also be bounded, and in that case they are constant, by Liouville's theorem. Those constants must vanish if A if is to be square integrable on $\mathbb{E}^n$; so there are no harmonic components in the Hodge decomposition of $\Lambda_p(\mathbb{E}^n)$. The orthogonal decomposition (39) therefore follows immediately from Proposition 9.1. The identity (40) follows very simply:

$$* [d\Lambda_p] = \{*d\alpha : \alpha \in \Lambda_p\} =$$

$$\{*d * \beta : \beta = (-1)^{n+1} * \alpha \in \Lambda_{n-p}\} = [\delta\Lambda_{n-p}].$$

The assumption that the material source is compactly supported implies that F is square integrable on $\mathbb{E}^n$. By (2), F is harmonic outside a sufficiently large ball B; and we may assume without loss of

generality that $\delta A = 0$. By the argument above, $\Delta A_\alpha = 0$ outside B, and each $A_\alpha = O(r^{2-n})$ at infinity. It follows without difficulty that $F = dA \in \Lambda_{p-1}(\mathbb{E}^n)$ $\blacksquare$.

References

[1] É. Cartan. Sur certaines expressions différentielles et le probéme de Pfaff. *Annales scientifiques de l'É.N.S.* 3^3 *série*, **16**:239–332, 1899.

[2] A. Einstein. Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 1905. Translated as "On the Electrodynamics of Moving Bodies", in The collected papers of Albert Einstein. Vol. 2, Princeton University Press, Princeton, New Jersey, 1989; Anna Beck translator.

[3] A. Einstein. Über das Relativitätsprinzip und die aus demselben gezogene Folgerungen. *Jahrbuch der Radioaktivität und Elektronik*, 1907. Translated as "On the relativity principle and the conclusions drawn from it", *op. cit.*

[4] A. Einstein. Die Grundlagen der Allgemeine Theorie der Relativität. *Annalen der Physik*, 1916. Translated as "The Foundation of the General Theory of Relativity", *op. cit.*

[5] Feynman, R., R. Leighton and M. Sands. *Lectures on Physics, vol.2*. Addison Wesley, Reading, Mass., 1963.

[6] W.I. Fucshchych and A.G. Nikitin. *Symmetries of Maxwell's Equations*. D. Reidel Publishing Co., Dordrecht, Netherlands, 1987.

[7] O. Heaviside. A gravitational and electromagnetic analogy. *The Electrician*, 31:281–282, 1893.

[8] Hilbert, D. *David Hilbert's Lectures on the Foundations of Physics, 1915-1927*, volume 5 of *David Hilbert's Lectures on the Foundations of Mathematics and Physics, 1891-1933*. Springer-Verlag, Berlin, Heidelberg, 2009.

[9] T. Kaluza. Zum Unitätsproblem in der Physik. *Sitzungsber. Preuss. Akad. Wiss. Berlin, (Math-Phys).*, 15:966–972, 1921.

[10] L.D. Landau and E. M. Lifshitz. *The Classical Theory of Fields.* Elsevier, Amsterdam, 1975.

[11] T. Levenson. *The Hunt for Vulcan*. Random House, New York, 2015.

[12] H.A. Lorentz. Considerations on gravitation. *Proc. Royal Netherlands Academy of Arts and Sciences*, 2:559–574, 1900.

[13] J.C. Maxwell. A dynamical theory of the electromagnetic field. *Phil. Trans. Royal Soc. London*, pages 459–512, 1865.

[14] Misner, C., K. Thorne and J.A. Wheeler. *Gravitation*. Freeman, New York, 1973.

[15] Nordström, G. Über die Möglichkeit, das elektromagnetishce Feld und das Gravitationsfeld zu vereinigen. *Physik. Zeitschr.*, **15**:504–506, 1914.

[16] Nordström, G. On the energy of the gravitational field in Einstein's theory. *Proc. Kon. Ned. Akad. Wet.*, **20**:1238–1245, 1918.

[17] A. Pais. *Subtle is the Lord, The Science and Life of Albert Einstein*. Oxford University Press, Oxford, 1982.

[18] J. Pohjanpelto. First order symmetries of Maxwell's equations. *Phys, Lett. A*, 129:148–150, 1988.

[19] H. Poincaré. *Les Méthodes Nouvelles de la Mécanique Céleste, Tome 3*. Gauthiers-Villars, Paris, 1892.

[20] H. Poincaré. Sur la dynamique de l'électron. *Rendiconti del Circolo Matematico di Palermo*, **21**:129–176, 1906.

[21] H. Reissner. Über die Eigengravitation des elektrischen Felds nach der Einsteinshen Theorie. *Ann. Physik*, **50**:106–120, 1916.

[22] D.H. Sattinger. Gravitation and special relativity. *Journal of Dynamics and Differential Equations*, **27**:1007–1025, 2015.

[23] D.H. Sattinger and O.L. Weaver. *Lie groups and Algebras with Applications to Physics, Geometry, and Mechanics*. Springer-Verlag, New York, Berlin, 1986.

[24] J.A. Stratton. *Electromagnetic Theory*. McGraw-Hill, New York and London, 1941.

[25] R. Wald. *General Relativity*. University of Chicago Press, Chicago, 1984.

[26] S. Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. John Wiley & Sons, New York, 1972.