

Gravitation and Special Relativity

D. H. Sattinger
Department of Mathematics
University of Arizona
Tucson, Arizona

Dedicated to Klaus Kirchgässner

July 5, 2012

Abstract

A mathematical derivation of Maxwell's equations for gravitation, based on a mathematical proof of Faraday's Law, is presented. The theory provides a linear, relativistic Lagrangian field theory of gravity in a weak field, and paves the way to a better understanding of the structure of the energy-momentum tensor in the Einstein Field Equations. Hence it is directly relevant to problems in modern cosmology.

The derivation, independent of the perturbation theory of Einstein's equations, puts the gravitational and electromagnetic fields on an equal footing for weak fields, contrary to generally held views. The historical objections to a linear Lorentz invariant field theory of gravity are refuted.

daidsattinger@gmail.com
<http://math.arizona.edu/~dsattinger/>

1 Introduction

In 1893 Oliver Heaviside published a paper [3] entitled “A Gravitational and Electromagnetic Analogy” in which he noted the similarities between the gravitational and electromagnetic fields. “Now, bearing in mind the successful manner in which Maxwell’s localization of electric and magnetic energy in his ether lends itself to theoretical reasoning,” he wrote, “the suggestion is very natural that we should attempt to localize gravitational energy in a similar manner, its density to depend upon the square of the intensity of the force, especially because the law of the inverse squares is involved throughout.” Heaviside’s attempt at a field theory of gravitation was followed by Lorentz (1900) [6] and Poincaré (1905) [11].

Those attempts were abandoned with the success of Einstein’s general theory of relativity, which is geometric in nature, and highly non-linear. Einstein maintained that gravitational forces were inherently non-linear, and could not be described by a linear, relativistic field theory. The historical arguments against a linear field theory of gravity are discussed in Pais [11] Chapter 13, and in the well-known text *Gravitation* by Misner, Thorne, and Wheeler [9], Chapter 7.

On the other hand, the formal linearization of Einstein’s equations at the Minkowski metric gives just such a theory. Chapter 7 of Misner *et.al.*, for example, is entitled “Incompatibility of Gravity and Special Relativity;” while Chapter 18, is entitled “The Linearized Theory of Gravity”. In §18.2, p. 442, we find the statement “The gauge conditions and field equations (18.8a,b) of linearized theory bear a close resemblance to the equations of electromagnetic theory in Lorentz gauge and flat space-time.”

As a result of the objections to a linear theory of gravitation, the early theories of gravity have been reincarnated under the rubric “Gravito-Electromagnetic Analogy,” in which linear field theories of gravity are obtained as a formal perturbation of the Einstein Field Equations in the vicinity of the Minkowski metric. But what is the physical significance of such formal perturbation methods if no *bona fide* linear relativistic field theory of gravitation exists? The perturbative approach is therefore a tacit acknowledgment that a linear field theory of gravity for weak fields must exist; and indeed it does. We shall call it the Maxwell-Heaviside theory, and shall show that the equations are not only mathematically rigorous §5, but utilitarian as well §7; and moreover, that Einstein’s objections to them do not apply

to weak fields §8.

“It is amusing to recall,” write Clark and Tucker [2], “that one of the first theories of post-Newtonian gravitation was formulated by Heaviside in direct analogy with the theory of electromagnetism. . . . It predicted that gravitation, like electromagnetism, was mediated by an independent vector field rather than with a second-degree tensor field associated with the metric of space-time. This difference [implies] that the analogy between weak gravity and electromagnetism is incomplete.”

Maxwell’s derivation of his equations for the electromagnetic field was based on extensive empirical data, especially Faraday’s Law of electromagnetic induction – the mathematical cornerstone of the theory. In the case of gravitation, however, there is as yet no experimental evidence for a gravitational field induced by the motion of mass; but a mathematical proof of Faraday’s Law is given in §5, providing the starting point for a derivation of Maxwell’s equations for gravity. This puts the gravitational and electromagnetic fields on an equal footing for weak fields, thus completing the Heaviside analogy.

One should expect that the linearized Einstein equations coincide with the Maxwell-Heaviside theory, but the matter is not straightforward, as Clark and Tucker demonstrate: “The question of the gauge transformations of the perturbative Einstein equations . . . leads one to contemplate the *most useful way* to define the *gravito-electromagnetic* fields in terms of the perturbed components of the space-time metric. Different choices are often responsible for the location of odd factors of four that permeate the *gravito-electromagnetic* equations compared with Maxwell’s equations.”

Einstein’s General Theory of Relativity models the gravitational field as the geodesic flow of a metric tensor $g^{\mu\nu}$, coupled to the energy-momentum tensor $T^{\mu\nu}$ by a highly nonlinear set of partial differential equations known as the Einstein Field Equations:

$$R^{\mu\nu} - \frac{1}{2}R g^{\mu\nu} = \frac{8\pi G}{c^4}T^{\mu\nu}. \quad (1)$$

Here $R^{\mu\nu}$ is the Ricci curvature tensor of $g^{\mu\nu}$, R is the scalar Ricci curvature, and G is the gravitational constant and c is the speed of light. Roughly speaking, $g^{\mu\nu}$ describes the geometry of space-time, $T^{\mu\nu}$ the physics.

Einstein originally obtained a solution of his equations for a point mass, for which $T^{\mu\nu} = 0$ everywhere except the origin. He obtained

correct results for the advance of the perihelion of Mercury, and the deflection of light by the Sun. His work was followed almost immediately by that of Schwarzschild [16], who obtained the metric tensor for Einstein's solution. There are now many other examples of solutions of Einstein's equations in a vacuum. These include the Kerr black hole, the Neugebauer-Meinell disk, and more generally, a system of integrable equations known as the Ernst equation which includes these two, Lenells [5].

Since roughly 99% of the mass in the solar system lies in the Sun, the Schwarzschild metric is an excellent model for the dynamics of the Solar System as motion in a central force field. The Schwarzschild metric has proved to be one of the most remarkable models of Mathematical Physics; but, due to the distribution of mass and energy in a galaxy, it fails to model the dynamics of galaxies. For this reason, the energy-momentum tensor plays a fundamental role in galactic dynamics, including the questions of dark matter and dark energy (see §7). It is also crucial to the study of gravitational collapse, beginning with the celebrated papers of Tolman [19] and Oppenheimer and Snyder [10]. Yet the energy-momentum tensor is known explicitly only in the case of electrodynamics: it is the Maxwell stress tensor. The fact that gravitation can also be modeled by Maxwell's equations provides significant additional information as to the structure of $T^{\mu\nu}$.

Maxwell's equations constitute potential theory in four dimensional space-time. Differential forms and the Hodge star operation are the natural language of potential theory, and extensive use will be made of them here. There are two orientations of an orientable manifold, and consequently two Hodge star operations; the orientation of space-time is reversed in going from the electromagnetic to the gravitational field. With a view to making this paper self-contained, a brief operational introduction to differential forms and the Hodge star operation is given in §2. A fuller account is given in [9]; and an exposition of the exterior differential calculus along the lines used here is given in Sattinger and Weaver [15]. A review of Maxwell's equations for electrodynamics in the language of differential forms is given in §4.

Minkowski¹ [8] did not use a metric tensor in his famous 1907 construction of space-time. Instead he represented it as \mathbb{E}^4 with $x^4 = ict$; Stratton [17] does the same. It will be denoted here by \mathbb{M}^4 . Minkowski's complex structure permits the use of the Hodge star

¹Minkowski died suddenly in 1909. The paper cited here is a posthumous publication of the article that appeared in the *Göttinger Mathematischen Gesellschaft* in 1907.

operation associated with \mathbb{E}^4 rather than that tied to the Minkowski metric, and simplifies the presentation. The complex structure is inherited by Maxwell's equations; and the Lorentz group is obtained from the rotation group on \mathbb{E}^4 under the transformation $x_4 \rightarrow ict$.

2 Mathematical Preliminaries

We assume the reader is already familiar with the basic operations of wedge product \wedge and exterior derivative d on p forms Λ_p and that the theorems of Green, Gauss, and Stokes are collected in a single theorem, known as Stokes' theorem

$$\iint_{\Omega} d\omega = \int_{\partial\Omega} \omega. \quad (2)$$

Here, $\omega \in \Lambda_p$ has differentiable coefficients, and Ω is a $p + 1$ dimensional, oriented manifold embedded in \mathbb{E}^n , with smooth boundary $\partial\Omega$.

A form $\omega \in \Lambda_p$ is said to be *closed* if $d\omega = 0$ and *exact* if $\omega = d\chi$, where $\chi \in \Lambda_{p-1}$. Since $d^2 = 0$, a p -form is closed if it is exact. In a simply connected region the two conditions are equivalent; and it will be sufficient to restrict ourselves to this case. A necessary and sufficient condition for $\omega \in \Lambda_p$ to be exact in a region U (not necessarily simply connected) is that its integral over every closed p manifold $\Sigma \subset U$ vanish:

$$\iint_{\Sigma} \omega = 0, \quad \text{whenever } \partial\Sigma = \emptyset.$$

In the special case of a 1-form E , Σ is a closed path, and the integral above is a line integral called the *circulation*. If the circulation vanishes for every smooth closed path, then regardless of the topology of the region, E is exact, and there exists a 0-form ϕ (that is, a single-valued function) such that $E = -d\phi$. It is standard convention to normalize the potential to vanish at infinity, so that it is explicitly given by

$$\phi(\mathbf{x}) = \int_{\mathbf{x}}^{\infty} E, \quad \mathbf{x} \in \mathbb{E}^3. \quad (3)$$

The electrostatic and gravitational fields are both conservative, so that their corresponding 1-forms are exact; but the electrostatic potential is positive, while the gravitational potential is negative. This reflects the

fact that the gravitational force is attractive, while the electrostatic force is defined in terms of like charges and so is repulsive.

The Hodge star operation $*$ on differential forms over an n dimensional orientable manifold \mathcal{M} plays a fundamental role in potential theory. It is defined as follows: Given an oriented volume element dv on \mathcal{M} and $\omega \in \Lambda_p$, $*\omega$ is defined as the $n - p$ form for which $\omega \wedge *\omega = dv$. The standard (right-handed) volume element on \mathbb{E}^3 is $dv = dx^1 \wedge dx^2 \wedge dx^3$. The associated Hodge star operation is

$$*dx^i = dx^j \wedge dx^k, \quad *1 = dv, \quad ** = id. \quad (4)$$

(i, j, k in cyclic order.) We shall always denote by $*$ the Hodge star operation on \mathbb{E}^n associated with the standard volume element $dv = dx^1 \wedge \dots \wedge dx^n$.

There is a second volume element on \mathbb{E}^3 , namely the left-handed volume element $d\tilde{v} = dx^3 \wedge dx^2 \wedge dx^1$, and hence a second star operation, denoted by $\tilde{*}$. Every orientable manifold has two orientations, corresponding to the even and odd permutations of the basis 1-forms, and consequently two star operations associated with it. Since $\xi \wedge \tilde{*}\eta = -\xi \wedge *\eta$ for all basis form $\xi, \eta \in \Lambda_p$, it follows that $\tilde{*} = -*$. The two orientations of \mathbb{E}^4 play a fundamental role in distinguishing the gravitational and electromagnetic fields.

For any star operation $*$, an inner product, called the Hodge duality, is defined for $\xi, \eta \in \Lambda_p(\mathbb{E}^n)$ by

$$(\xi, \eta) = \iiint_{\mathbb{E}^n} \xi \wedge *\eta, \quad \xi, \eta \in \Lambda_p \quad (5)$$

The Hodge duality in turn defines a formal adjoint to the exterior derivative d . Called the *coderivative*, it maps Λ_{p+1} to Λ_p , and is defined implicitly by the relation

$$(d\xi, \eta) = (\xi, \delta\eta) \quad \xi \in \Lambda_p, \quad \eta \in \Lambda_{p+1} \quad (6)$$

where ξ, η have compact support on \mathbb{E}^n .

Unless otherwise stated, we shall denote the coderivative associated with the standard Hodge star operation $*$ by δ and that associated with the operation $\tilde{*}$ by $\tilde{\delta}$. The action of δ on Λ_p will be denoted by δ_p .

Proposition 2.1 *We have*

- i) $\delta_p = (-1)^{p-1} *^{-1} d*$ on \mathbb{E}^n for both $*$ operators;
- ii) $\tilde{*} = -*$ on \mathbb{E}^n , hence $\tilde{\delta}_p = \delta_p$;
- iii) $\delta_p = (-1)^{p-1} * d*$ on \mathbb{E}^3 ;
- iv) $** = (-1)^p id$ and $\delta = * d*$ on \mathbb{E}^4 .

Proof: i): Let ξ, η be smooth differential forms with compact support, and integrate the identity

$$\begin{aligned} d(\xi \wedge *\eta) &= d\xi \wedge *\eta + (-1)^p \xi \wedge d*\eta \\ &= d\xi \wedge *\eta - \xi \wedge *(-1)^{p-1} *^{-1} d*\eta = d\xi \wedge *\eta - \xi \wedge *\delta\eta \end{aligned}$$

over any large ball B . By Stokes' theorem, we get on the left side

$$\iiint_B d(\xi \wedge *\eta) = \iint_{\partial B} \xi \wedge *\eta = 0$$

when B is sufficiently large. The right side then gives (6). The same proof holds for $\tilde{*}$.

ii): The assertion for $1 \leq p \leq n-1$ follows from the fact that $\xi \wedge *\eta = -\xi \wedge \tilde{*}\eta$ for all $\xi, \eta \in \Lambda_p$ for such p . For $p=0, n$, note that $\tilde{*}1 = d\tilde{v} = -dv$ while $\tilde{*}d\tilde{v} = 1$ is equivalent to

$$(d\tilde{v}, d\tilde{v})_U = \int_U d\tilde{v} \wedge \tilde{*}d\tilde{v} = \int_U d\tilde{v} = - \int_U dv = -\text{vol}(U).$$

iii): From (4) it follows that $*^{-1} = *$ on \mathbb{E}^3 , and the assertion in iii) follows from that in i).

iv): The Hodge star operation on \mathbb{E}^4 associated with the standard volume element is

$$* dx^j = dx^k \wedge dx^l \wedge dx^4 \qquad * dx^4 = -dv \qquad (7)$$

$$* dx^j \wedge dx^k = dx^l \wedge dx^4 \qquad * dx^j \wedge dx^4 = dx^k \wedge dx^l \qquad (8)$$

$$* dv = dx^4, \qquad * dx^j \wedge dx^k \wedge dx^4 = -dx^l. \qquad (9)$$

It follows that $*^{-1} = (-1)^p *$ on $\Lambda_p(\mathbb{E}^4)$. On the other hand,

$$* : \Lambda_p \rightarrow \Lambda_{n-p}, \quad d : \Lambda_{n-p} \rightarrow \Lambda_{n-p+1}, \quad *^{-1} : \Lambda_{n-p+1} \rightarrow \Lambda_{p-1}.$$

For $n = 4$, therefore, $(-1)^{p-1} *^{-1} = *$; and so, by i),

$$\delta_p = (-1)^{p-1} *^{-1} d * = * d * \text{ on } \mathbb{E}^4 \text{ for all } p \quad (10)$$

This completes the proof of Proposition 2.1.

Maxwell's equations are formulated as a system of partial differential equations for a set of "vector" fields $\mathbf{E}, \mathbf{D}, \mathbf{H}, \mathbf{B}$; but the vector notation is ambiguous. The electric and magnetic fields, \mathbf{E} and \mathbf{H} , are ordinary polar vectors, and are identified with the 1-forms $E = \mathbf{E} \cdot d\mathbf{x}$, $H = \mathbf{H} \cdot d\mathbf{x}$.

The "vectors" \mathbf{B} and \mathbf{D} , the magnetic induction and electric displacement, are axial vectors, associated with surface integrals (fluxes), hence with 2-forms B, D . In \mathbb{E}^3 , we take $B = \mathbf{B} \cdot d\mathbf{S}$ and similarly for D . The expression $d\mathbf{S}$ denotes the vector element of surface area on a 2 dimensional surface S embedded in \mathbb{E}^3 ; it is given by $d\mathbf{S} = (X_u \times X_v) du \wedge dv$, where $X = (x^1(u, v), x^2(u, v), x^3(u, v))$ is a parametrization of a neighborhood of S by local coordinates u, v . We leave it to the reader to verify the identities

$$dx^i \wedge dx^j = \frac{\partial(x^i, x^j)}{\partial(u, v)} du \wedge dv, \quad B = B_j dx^k \wedge dx^l.$$

Here and throughout this article, the expression for B signifies a summation over j, k, l from 1 to 3 in cyclical order.

3 Potential theory in \mathbb{E}^3

There are two physical parameters in Maxwell's theory, the electric and magnetic inductive capacities ϵ and μ . Stratton defines them by the constitutive relations

$$\mathbf{D} = \epsilon \mathbf{E}, \quad \mathbf{B} = \mu \mathbf{H}, \quad (11)$$

relating the axial vectors \mathbf{D} and \mathbf{B} to the polar vectors \mathbf{E} , \mathbf{H} . This logical inconsistency is removed by defining ϵ and μ in terms of the associated differential forms. Accordingly, Gauss' law of electrostatics in an isotropic medium is stated in terms of the Hodge star operation as:

$$\iint_S D = Q_S, \quad D = \epsilon * E. \quad (12)$$

Here, Q_S is the total charge contained inside S , E is the 1-form $E = \mathbf{E} \cdot d\mathbf{x}$, where \mathbf{E} is the electric field, and D is Maxwell's displacement.

For a point charge q at the origin

$$E = \frac{G_e q}{r^2} \hat{\mathbf{r}} \cdot d\mathbf{x} = \frac{G_e q}{r^3} x_j dx^j; \quad D = \frac{q x_j dx^k \wedge dx^l}{4\pi r^3}. \quad (13)$$

where G_e is a physical constant, determined by Coulomb. The first equation is Coulomb's inverse square law. The integral of D over a closed surface S enclosing the origin can be deformed to an integral over a sphere of radius R centered at the origin. By Stokes' theorem

$$\iint_S D = \frac{q}{4\pi R^3} \iint_S x_j dx^k \wedge dx^l = \frac{q}{4\pi R^3} \iiint_{B_R} 3 dv = q,$$

where B_R denotes the interior of the sphere. The second equation in (12), together with the expressions for D and E in (13) imply $\epsilon = 1/4\pi G_e$.

In the case of a continuous charge density, we have

$$D = D_j dx^k \wedge dx^l, \quad dD = \rho dv, \quad \rho = \sum_j \frac{\partial D_j}{\partial x^j}. \quad (14)$$

The differential expression for ρ is valid in general coordinates, but in Cartesian coordinates it can be written in vector form $\operatorname{div} \mathbf{D} = \rho$. If $Q_S > 0$ throughout a region U , then $\rho \geq 0$ in U .

The equations of the electrostatic field are $E = -d\phi$, $D = \epsilon * E$. Eliminating D using Proposition 2.1, iii) we find $\delta E = \rho/\epsilon$. Combining these two equations for E we obtain $\delta d\phi = -\rho/\epsilon$. It is easily verified that on $\Lambda_0(\mathbb{E}^3)$ (functions)

$$\delta d = * d * d = \sum_{j=1}^3 \frac{\partial^2}{\partial x_j^2},$$

and the equation for the electrostatic potential on \mathbb{E}^3 is $\Delta\phi = -\rho/\epsilon$.

In the case of gravitation, the lines of force flow into the region bounded by S , since the force is attractive. If we follow the procedure above, (14) will lead to $\rho = \epsilon \operatorname{div} \mathbf{F} < 0$. A positive mass density can be obtained by reversing the orientation, and using $\tilde{*}$. It is given by

$$\tilde{*} dx^i = dx^j \wedge dx^k, \quad i, j, k \text{ anti-cyclic}; \quad \tilde{*} 1 = d\tilde{v}; \quad \tilde{*}\tilde{*} = id.$$

Gauss' law for the gravitational field then takes the form

$$\iint_S D = M, \quad D = \epsilon_g \tilde{*} F, \quad \epsilon_g = \frac{1}{4\pi G} \quad (15)$$

where F is the 1-form associated with the gravitational field, M the total mass in S , and G the Cavendish gravitational constant. The mass density ρ is non-negative, due to our choice of the left-handed volume element, for now

$$dD = \varepsilon(\operatorname{div} \mathbf{F}) d\tilde{v} = \rho dv, \quad \rho = -\varepsilon \operatorname{div} \mathbf{F} > 0.$$

Summarizing the two cases, we have

Theorem 3.1 *For the electrostatic field in \mathbb{E}^3 , take the Hodge star operation associated with the right-handed volume element. The equations for the electrostatic potential are then*

$$E = -d\phi, \quad \delta E = \operatorname{div} \mathbf{E} = \frac{\rho}{\epsilon}, \quad \Delta \phi = -\frac{\rho}{\epsilon}. \quad (16)$$

For the gravitational field take the Hodge star operation $\tilde{}$ associated with the left-handed volume element. The equations for the gravitational potential are then*

$$F = -d\phi, \quad \tilde{\delta} F = \operatorname{div} \mathbf{F} = -\frac{\rho}{\epsilon_g}, \quad \tilde{\Delta} \phi = \Delta \phi = \frac{\rho}{\epsilon_g}. \quad (17)$$

The proof of the gravitational case is left to the reader.

4 Potential Theory in \mathbb{M}^4 , I

Maxwell's equations of electrodynamics in vector form are [17]

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \quad \operatorname{div} \mathbf{B} = 0; \quad (18)$$

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J}, \quad \operatorname{div} \mathbf{D} = \rho. \quad (19)$$

The system is closed with the two constitutive relations (11). The first pair of equations is the differential form of Faraday's law of magnetic induction, the second Ampère's law modified by Maxwell's introduction of the displacement current \mathbf{D}_t ; ρ denotes the charge density, and

$\mathbf{J} = \rho \mathbf{v}$ the current. Classical potential theory begins with a 1-form F with zero circulation; the corresponding object in Maxwell's equations is the Faraday 2-form

$$F = E \wedge dt + B, \quad (20)$$

where, as above, $E = \mathbf{E} \cdot d\mathbf{x}$ and $B = \mathbf{B} \cdot d\mathbf{S}$.

Theorem 4.1 *Assume \mathbf{E} and \mathbf{B} are smooth vector fields. Then Faraday's Law (18) is a necessary and sufficient condition that the Faraday 2-form (20) be exact. In that case there is a 1-form $A = A_j dx^j$ such that $F = dA$.*

Proof: We calculate $dF = dE \wedge dt + dB$ in Cartesian coordinates:

$$\begin{aligned} dE &= \frac{\partial E_i}{\partial t} dt \wedge dx^i + \left(\frac{\partial E_i}{\partial x_j} - \frac{\partial E_j}{\partial x_i} \right) dx^j \wedge dx^i \\ &= (\nabla \times \mathbf{E})_i dx^j \wedge dx^k - \frac{\partial E_i}{\partial t} dx^i \wedge dt; \\ dB &= \frac{\partial B_i}{\partial t} dx^j \wedge dx^k \wedge dt + \frac{\partial B_i}{\partial x^i} dx^i \wedge dx^j \wedge dx^k \\ &= \frac{\partial B_i}{\partial t} dx^j \wedge dx^k \wedge dt + \operatorname{div} \mathbf{B} dv \end{aligned}$$

Thus $dF = 0$ implies the two equations in (18).

Conversely, Faraday's Law implies that F is closed; and since \mathbb{M}^4 is simply connected, a closed form is exact. In short, Faraday's Law is equivalent to the existence of a 1-form A such that $F = dA$. ■

We shall call the class of all 1-forms $A = A_\mu dx^\mu$ for which $\partial_\mu A_\mu = 0$ Lorentzian 4-potentials, and we denote them by \mathcal{P} . Throughout this paper we consider only such 4-potentials; and we restrict the discussion to smooth forms and vector fields.

The 4-current J is defined to be $J = J_j dx^j = \mathbf{J} \cdot d\mathbf{x} + J_4 dx^4$. Under the transformation $x_4 = ict$, $J_4 = ic\rho$, $J = \rho v_j dx^j - c^2 \rho dt$ (see Stratton, §1.21) The conservation of charge in rectangular coordinates is $\operatorname{div} \mathbf{J} + \partial_t \rho = 0$. This can be written as $\partial_\mu J_\mu = 0$; hence the 4-current also belongs to the class \mathcal{P} .

By (8) the pair of constitutive relations (11) can be written as

$$D = \epsilon * E \wedge dx^4, \quad B = \mu * H \wedge dx^4. \quad (21)$$

Theorem 4.2 *The Maxwell-Ampère law is equivalent to the equation $dG = *J$, where $G = ic(H \wedge dt - D)$ and J is the 4-current. The necessary and sufficient condition for the solvability of this equation is $d * J = 0$.*

*Assume the Faraday 2-form is exact, and that the constitutive laws (21) hold. Then $*F = \mu G$ if and only if $\epsilon\mu c^2 = 1$; and in that case Maxwell's equations take the form*

$$F = dA, \quad \delta F = -\mu J, \quad \square A = \mu J, \quad (22)$$

where

$$\delta dA = * d * dA = -\square A, \quad \square = \Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}. \quad (23)$$

Proof: Observe that G , with $H = H_j dx^j$ and $D = D_j dx^k \wedge dx^l$, represents a general 2-form on \mathbb{M}^4 . By direct calculation

$$dG = \left(\frac{\partial H_l}{\partial x^k} - \frac{\partial H_k}{\partial x^l} - ic \frac{\partial D_j}{\partial x^4} \right) dx^k \wedge dx^l \wedge dx^4 - ic \sum_{j=1}^3 \frac{\partial D_j}{\partial x^j} dv \quad (24)$$

In Cartesian coordinates,

$$\frac{\partial H_l}{\partial x^k} - \frac{\partial H_k}{\partial x^l} = (\nabla \times \mathbf{H})_j, \quad \sum_{j=1}^3 \frac{\partial D_j}{\partial x^j} = \operatorname{div} \mathbf{D}. \quad (25)$$

By (7) $*J = (J_j dx^k \wedge dx^l \wedge dx^4 - ic \rho dv)$; hence in Cartesian coordinates (24) is precisely the Maxwell-Ampère equation (19). Note that the conservation of charge implies that $d * J = 0$.

Now

$$\begin{aligned} *F &= * \left(\frac{E}{ic} \wedge dx^4 + B \right) = \frac{1}{ic\epsilon} D + \mu H \wedge dx^4 \\ &= \mu \left(H \wedge dx^4 + \frac{1}{ic\epsilon\mu} D \right) = \mu ic \left(H \wedge dt + \frac{1}{(ic)^2 \epsilon \mu} D \right) = \mu G, \end{aligned}$$

if and only if $\mu\epsilon c^2 = 1$. By Proposition 2.1, iv):

$$\delta F = (* d *) * \mu G = \mu * dG = \mu ** J = -\mu J.$$

The third equation in (22) follows from the relation $\delta d = * d * d = -\square$. ■

5 Potential Theory in \mathbb{M}^4 , II

In the case of gravitation there is presently no experimental analog of Faraday's Law of magnetic induction; but it turns out that Faraday's Law is a mathematical artifact of the geometry of space-time. Though the following is general, we retain the notation of electrodynamics for clarity.

Theorem 5.1 *Let $E = E_j(x, t)dx^j$ be any space-like 1-form on \mathbb{M}^4 . Then there is a family of Lorentzian 4-potentials \mathcal{P} and an exact 2-form $B = B_j(x, t) dx^k \wedge dx^l$ such that $F = E \wedge dt + B = dA$ for all $A \in \mathcal{P}$.*

Proof: The 4-potentials $A \in \mathcal{P}$ are obtained by inverting the linear hyperbolic system of partial differential equations

$$\begin{pmatrix} \partial_t & 0 & 0 & \partial_1 \\ 0 & \partial_t & 0 & \partial_2 \\ 0 & 0 & \partial_t & \partial_3 \\ \partial_1 & \partial_2 & \partial_3 & c^{-2}\partial_t \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ \phi \end{pmatrix} + \begin{pmatrix} E_1 \\ E_2 \\ E_3 \\ 0 \end{pmatrix} = 0, \quad (26)$$

where $\partial_j = \partial/\partial x^j$ and $\phi = -icA_4$.

A quick calculation shows that for $A = A_j dx^j$,

$$\begin{aligned} dA &= \sum_{j < k \leq 3} \left(\frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k} \right) dx^j \wedge dx^k \\ &\quad + \sum_{j=1}^3 \left(\frac{\partial A_4}{\partial x^j} - \frac{\partial A_j}{\partial x^4} \right) dx^j \wedge dx^4. \end{aligned}$$

Putting $F = dA$ we obtain

$$B_i = \frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k}, \quad \frac{E_j}{ic} = \frac{\partial A_4}{\partial x^j} - \frac{\partial A_j}{\partial x^4}. \quad (27)$$

The first set of equations implies that $\partial_i B_i = 0$ no matter the choice of A . Adding the Lorentz condition $\partial_\mu A_\mu = 0$ and putting $\phi(x, t) = -icA_4$, we arrive at (26).

The system (26) constitutes a first order hyperbolic system. Eliminating the A_j , we obtain

$$\Delta\phi - \frac{1}{c^2}\phi_{tt} + \rho = 0, \quad \rho = \frac{\partial E_i}{\partial x_i}, \quad \Delta\phi = \sum_{j=1}^3 \partial_j^2 \phi. \quad (28)$$

Given a solution ϕ , set $\psi(x, t) = \int^t \phi(x, s) ds$. Then

$$A_j = - \int^t E_j ds - \frac{\partial \psi}{\partial x^j}, \quad j = 1, 2, 3; \quad A_4 = - \frac{\partial \psi}{\partial x^4}. \quad (29)$$

Any two solutions of (26) differ by a homogeneous solution – i.e. $E = 0$, hence $\rho = 0$ and ϕ is a solution of the homogeneous wave equation. By (29), the 4-potential of a homogeneous solution is then $-d\psi$. But dA is unchanged under the gauge transformation $A \rightarrow A - d\psi$; hence B does not depend on the choice of solution. ■

Remark: The 2-form B , though independent of the choice of $A \in \mathcal{P}$, is not uniquely determined. Consider a divergence-free, stationary vector field $\mathbf{V} = (V^1, V^2, V^3)$, and put $V = V_j dx^j$. Then V is a Lorentzian 4-potential and under the gauge transformation $A \rightarrow A + V$, B goes to $B + B_V$, where B_V is obtained from V via (27). In the electromagnetic case, B_V is a magnetostatic field, generated by a steady current, while B corresponds to the induced field, generated by fluctuations in the electric field.

The inductive capacity ϵ_g for gravity was given in (15). The recognition that the mass current, analogous to the charge current in electrodynamics, is the momentum, is due to Heaviside [3]. The momentum 1-form for a particle in Special Relativity is ([4] §9):

$$P = p_\mu dx^\mu = \mathbf{p} \cdot d\mathbf{x} + ic\rho dx^4 = \rho v_j dx^j + ic\rho dx^4, \quad (30)$$

where $\rho = \rho_0(1 - v^2/c^2)^{-1/2}$ with ρ_0 the rest mass density.

As in the case of static potential theory on \mathbb{E}^3 , we need to reverse orientations in going from the electromagnetic to the gravitational field. For gravitation we use the Hodge star operator associated with the oriented volume element $dx^4 \wedge dv$:

$$\tilde{*} dx^j = -dx^4 \wedge dx^k \wedge dx^l \qquad \tilde{*} dx^4 = dv \quad (31)$$

$$\tilde{*} dx^j \wedge dx^k = dx^4 \wedge dx^l \qquad \tilde{*} dx^4 \wedge dx^j = dx^k \wedge dx^l \quad (32)$$

$$\tilde{*} dv = -dx^4, \qquad \tilde{*} dx^4 \wedge dx^j \wedge dx^k = dx^l. \quad (33)$$

We use the same 2-forms F and G that we used in the electromagnetic case, but replace $*$ by $\tilde{*}$ in the computations. This time we have to solve

$$dG = \tilde{*}P, \qquad \tilde{*}P = -p_j dx^4 \wedge dx^k \wedge dx^l + ic\rho dv.$$

The vector form of the Maxwell-Heaviside equation – the analog of the Maxwell-Ampère equation for the gravitational field – is then

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = -\mathbf{p}, \quad \operatorname{div} \mathbf{D} = -\rho. \quad (34)$$

By Proposition 2.1 ii) we have $*\tilde{F} = -*F = -\mu G$; but we may also check this by a direct computation using the equations for $\tilde{*}$ above:

$$\begin{aligned} \tilde{*}F &= \tilde{*} \left[\frac{E_j}{ic} dx^j \wedge dx^4 + B_j dx^k \wedge dx^l \right] \\ &= - \left[\frac{D_j}{ic\epsilon_g} dx^k \wedge dx^l + ic\mu H_j dt \wedge dx^j \right] \\ &= -ic\mu \left[H \wedge dt + \frac{D_j}{(ic)^2 \mu_g \epsilon_g} dx^k \wedge dx^l \right] = -\mu G. \end{aligned}$$

Therefore $\tilde{\delta}F = \tilde{*}d\tilde{*}F = -\mu\tilde{*}dG = -\mu\tilde{*}\tilde{*}P = \mu P$. The other equations follow as before.

The Maxwell-Heaviside equations are given in the following:

Theorem 5.2 *Let F and G be the 2-forms in Theorem 4.2 and let P the mass current (30). The necessary and sufficient condition for the solvability of the equation $dG = \tilde{*}P$ is $\tilde{\delta}P = 0$. Let $\mu_g = 4\pi Gc^{-2}$. Then the Maxwell-Heaviside equations for gravity are*

$$F = dA, \quad \tilde{\delta}F = \mu_g P, \quad \tilde{\square}A = -\mu_g P. \quad (35)$$

In terms of the $$ operation, the equations are*

$$F = dA, \quad \delta F = \mu_g P, \quad \square A = -\mu_g P, \quad (36)$$

Remarks: i) The hypothesis $\mu_g = 4\pi Gc^{-2}$ is equivalent to the assumption that gravitational waves travel with the speed of light. This relationship was known experimentally in the theory of the electromagnetic field and was cited explicitly by Maxwell [7] (pp. 577-580). In the theory of Special Relativity, c appears as a universal constant of nature that determines not only the speed of light, but the speed of gravitational waves as well in a vacuum. In Einstein's general theory of relativity, it comes out of the analysis of weak gravitational waves [9].

ii) The value of μ_g is

$$\frac{4\pi G}{c^2} = 9.31 \times 10^{-28} \frac{cm \ sec^2}{gm}$$

The extreme smallness of this parameter indicates why the gravitational field induced by the motion of mass is so difficult to detect.

6 The Lagrangian

Theorems 4.2 and 5.2 can be merged into a single statement.

Theorem 6.1 *Let P be the 4-momentum (30), e be the electric charge, and $*$ the Hodge star corresponding to $dv \wedge dx^4$. Then Maxwell's equations are*

$$J = (-e)^\sigma P, \quad F = E \wedge dt + B, \quad (37)$$

$$F = dA, \quad \delta F = \mu_\sigma J_\sigma, \quad \square A = -\mu_\sigma J_\sigma \quad (38)$$

where $\sigma = 1$ for the Maxwell-Ampère equations, and 0 for the Maxwell-Heaviside equations, and $\mu_\sigma = \mu_g$ for $\sigma = 0$ and μ , the magnetic permittivity, for $\sigma = 1$.

The action integral in terms of the Minkowski metric is given in Landau and Lifshitz, §27 for the Maxwell-Ampère equations, and in Misner *et.al.* Exercise 7.2, for both cases. The action can also be written in terms of the Hodge duality on \mathbb{M}^4 . On \mathbb{M}^4 the volume element is $dv \wedge dx^4 = ic \, dv \wedge dt$, so to get a real symmetric inner product on real p -forms we take

$$(\xi, \eta) = \frac{1}{ic} \iiint_{\mathcal{K}} \xi \wedge *\eta, \quad \xi, \eta \in \Lambda_p(\mathcal{K}), \quad \mathcal{K} \subset \mathbb{M}^4. \quad (39)$$

Since Maxwell's equations form a hyperbolic system, we take the region of integration \mathcal{K} to be the backward ray cone from a point in \mathbb{M}^4 . The action for Maxwell's equations is then

$$S = ic \left[\frac{1}{2} (F, F) + \mu (A, J) \right] = \iiint_{\mathcal{K}} \frac{1}{2} F \wedge *F + \mu A \wedge *J. \quad (40)$$

The action is imaginary, and its critical points are obtained via the Principle of Stationary Phase. Denoting the variation of S by \dot{S} , we get $\dot{S} = ic[(\dot{F}, F) + \mu(\dot{A}, J)]$, where $(\dot{F}, F) = (d\dot{A}, F) = (\dot{A}, \delta F)$. Thus $\dot{S} = ic(\dot{A}, \delta F + \mu J)$. 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 (38). (The third equation in (38) is a consequence of the first two.)

The Lagrangian L , which is real, is obtained by setting $x^4 = ict$ and putting

$$L dv \wedge dt = \frac{1}{ic} \left[\frac{1}{2} F \wedge *F + \mu A \wedge *J \right].$$

In the electrodynamic case, we find

$$\begin{aligned} A \wedge *J &= ice(\mathbf{A} \cdot \mathbf{p} + \phi\rho) dv \wedge dt; \\ F \wedge *F &= ic \left(\mathbf{B} \cdot \mathbf{B} - \frac{1}{c^2} \mathbf{E} \cdot \mathbf{E} \right) dv \wedge dt \\ &= ic\mu (\mu \mathbf{H} \cdot \mathbf{H} - \epsilon \mathbf{E} \cdot \mathbf{E}) dv \wedge dt. \end{aligned}$$

Hence

$$L = \mu (\mu \mathbf{H} \cdot \mathbf{H} - \epsilon \mathbf{E} \cdot \mathbf{E}) + e(\mathbf{A} \cdot \mathbf{p} + \phi\rho). \quad (41)$$

Turning now to the gravitational case, note that $\bar{S} = -S$ is also an action for Maxwell's equations, and that

$$\bar{S} = \iiint_{\mathcal{K}} \frac{1}{2} F \wedge \tilde{*}F + \mu A \wedge \tilde{*}J. \quad (42)$$

Thus \bar{S} is the action for the Maxwell-Heaviside equations (35); and the associated Lagrangian is the negative of that for the electromagnetic field:

$$L = -\mu_g (\mu_g \mathbf{H} \cdot \mathbf{H} - \epsilon_g \mathbf{E} \cdot \mathbf{E}) - (\mathbf{A} \cdot \mathbf{p} + \phi\rho). \quad (43)$$

In particular, the energy of the gravitational field is negative, as noted in [9] Exercise 7.2.

7 Dark Matter: Red Herring?

The failure of the Schwarzschild metric to model the dynamics of spiral galaxies has been demonstrated dramatically by extensive observations by Rubin et.al. [13][14].

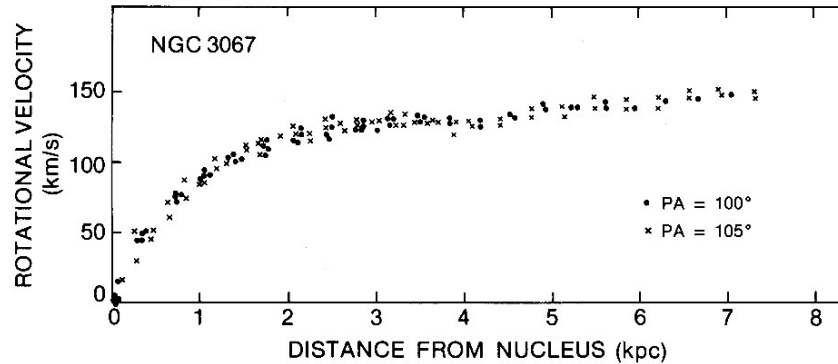


FIG. 2. Rotation velocities in the plane of NGC 3067, as a function of the distance from the center projected to p.a. = 101°5, from $H\alpha$ and $[N\ III]$ $\lambda\ 6583$ lines. Circles are from spectrum in p.a. = 100°, ×'s are from spectrum in p.a. = 105°.

The graph above, called the rotation curve, shows increasing speeds of hydrogen and nitrogen atoms as measured by shifts in the emission lines $H\alpha$ (hydrogen) and $N\ III$ (nitrogen) going out from the core of the galaxy. The rising curve violates Kepler's third law of planetary motion in a central force field, according to which the velocities should drop off as $r^{-1/2}$, where r is the mean distance from the central mass. The distance is measured in kiloparsecs, 1 kpc being approximately 3,260 light years. The particle velocities are small compared with the speed of light, but since the gravitational field propagates with the speed of light, it would require 3,260 years for a disturbance to travel 1 kpc.

The rotation curve, as well as other anomalies in galactic dynamics, have led to a hypothesis of "dark matter," matter which is not visible – even yet-to-be discovered forms of matter² which do not interact with the electromagnetic field – in order to explain the observations. By definition, any "evidence" for dark matter in the universe is necessarily indirect, taking the form of dynamics which cannot be explained using current mathematical models. The dark matter hypothesis is sometimes presented as an alternative: 1. Either Newton's Laws of

²On July 4, 2012 physicists at CERN announced the discovery of a heavy boson with properties similar to the long-sought Higgs boson. Such a discovery is significant for the standard model of elementary particles in quantum physics, but the particle is an artifact of high-energy physics, and decays very rapidly. Such a theory is relevant to the early stages of the universe, when the energy density was very high; but its relevance to galactic dynamics, which is a problem in low energy physics, has yet to be established.

motion are not valid; or 2. There is additional “nonluminous” matter.

Current research focusses almost entirely on the second proposition, despite the fact that Newton’s theory of gravity posits “action at a distance”, and cannot hold on a galactic scale. The discovery of some exotic form of “non-luminous” matter would thus not resolve the dynamical issues involved. The one conclusion that can definitively be drawn from the data obtained by Rubin and her co-workers, dark matter or not, is that the energy-momentum tensor does not vanish in the outer reaches of galaxies.

The general theory of the energy-momentum tensor is presented in the monograph by Landau and Lifshitz [4] §32, §35, where it is obtained as the Euler-Lagrange derivative of a Lorentz invariant Lagrangian. Mass does not appear explicitly in T , but T^{00} represents energy density. It includes energy of all types: mass, gravitational energy, electromagnetic energy, thermal energy, and kinetic energy, including rotational kinetic energy. The general form of the tensor for macroscopic bodies (§35) is given as the strikingly simple formula

$$T^{\mu\nu} = (p + \varepsilon)u^\mu u^\nu - pg^{\mu\nu}, \quad (44)$$

where ε is the energy density, p is the “pressure”, and u^ν is the 4-velocity of the material flow. There are no specifics about the composition of the energy density.

Equation (44) is derived assuming that Pascal’s law of hydrostatics is valid in a co-moving frame, that is, a reference frame moving with the material point. In that case, the classical energy-momentum tensor takes the form

$$T^{\mu\nu} = \begin{pmatrix} \varepsilon & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}$$

The form (44) then follows by a field of Lorentz transformations depending on the point in question.

Pascal’s Law states that the pressure throughout an incompressible fluid at rest is constant. To suggest it holds in the interior of a star or in a turbulent star-forming region such as the Orion nebula is, to say the least, a bit of a stretch. Finally, $T^{\mu\nu}$ is supposed to be the gradient of a Lorentz invariant functional; yet none is given.

The dark matter hypothesis was originally put forward by Zwicky [21]. In §5 of that paper, Zwicky makes some rough statistical calculations about the Coma cluster in which he compares estimates of

its mass obtained by two different methods. The Coma cluster is comprised of some 800 galaxies at a distance of 45 million light years. Zwicky measures the velocity dispersion of the individual galaxies and, using classical statistical arguments, gets an estimate of the total kinetic energy of the cluster. Assuming the 800 galaxies form a “gas” in statistical equilibrium, he uses the classical virial theorem, which implies that the mean kinetic and potential energies of the system are equal, to estimate the total mass of the cluster. By this means he estimates the average density of the cluster “to be at least 400 times larger than that derived on the grounds of observations of luminous matter. If this would be confirmed we would get the surprising result that dark matter is present in much greater amount than luminous matter.”

Classical statistical mechanics, based on Newton’s theory, plays a central role in Zwicky’s analysis; yet the Coma cluster is about one million light years across, calling instead for a relativistic statistical mechanics based on the Maxwell-Heaviside equations. Zwicky’s model of the Coma cluster as a “gas” of point particles raises questions of its own. Galaxies are not point masses; they have internal structure, especially rotational kinetic energy.

In his 1937 paper [22], Zwicky returns to the problem of dark matter, listing several caveats. In the abstract he states

“Present estimates of the masses of nebulae are based on observations of the *luminosities* and *internal rotations* of nebulae. It is shown that both these methods are unreliable; that from the observed luminosities of extragalactic systems only lower limits for the values of their masses can be obtained, and that from internal rotations alone no determination of the masses of nebulae is possible.”

He goes on

“In order to derive trustworthy values of the masses of nebulae from their absolute luminosities, detailed information on the following three points is necessary.

1. According to the mass-luminosity relation, the conversion factor from absolute luminosity to mass is different for different types of stars. The same holds true for any kind of luminous matter. In order to determine the conversion factor for a nebula as a whole, we must know in what proportions all the possible luminous components are represented in this nebula.

2. We must know how much dark matter is incorporated in nebulae in the form of cool and cold stars, macroscopic and microscopic solid bodies, and gases.
3. Finally, we must know to what extent the apparent luminosity of a given nebula is diminished by the internal absorption of radiation because of the presence of dark matter.”

Zwicky’s three points illustrate the complexity inherent in the dark matter mystery, a complexity that is monotonically increasing as a function of technology:

“In 2005, the Advanced Camera for Surveys instrument of the Hubble Space Telescope finished capturing the most detailed image of the nebula yet taken. The image was taken through 104 orbits of the telescope, capturing over 3,000 stars down to the 23rd magnitude, including infant brown dwarfs . . . A year later, scientists working with the HST announced the first ever masses of a pair of eclipsing binary brown dwarfs. . . . in the Orion Nebula [having] approximate masses of $0.054M_{\odot}$ and $0.034M_{\odot}$ respectively, with an orbital period of 9.8 days. Surprisingly, the more massive of the two also turned out to be the less luminous.” Orion Nebula, Wikipedia.

Rubin expresses support for a dynamical approach to the dark matter mystery:

Currently, the theory of dark matter is the most popular candidate for explaining the galaxy rotation problem. The alternative theory of MOND (Modified Newtonian Dynamics) has little support in the community. Rubin, however, supports the MOND approach, stating “If I could have my pick, I would like to learn that Newton’s laws must be modified in order to correctly describe gravitational interactions at large distances. That’s more appealing than a universe filled with a new kind of sub-nuclear particle.” Vera Rubin, Wikipedia

The Maxwell-Heaviside equations of gravitation constitute a linear, relativistic correction to Newton’s equations of motion; they interpolate between Newton’s and Einstein’s theories of gravitation, and are therefore a natural mathematical model on which to build a dynamical theory of galactic structures.

8 Maxwell’s Enigma

In 1905 Poincaré [12] observed that the transformations used by Lorentz in his theory of the electron constituted a transformation group. He believed that Lorentz covariance was a fundamental fact of physics; and at the end of his paper, in a section entitled *Hypothèses sur la Gravitation*, he proposed a rudimentary Lorentz-covariant form of the gravitational field that included motion. He speculated about “l’onde gravifique, . . . étant supposée se propager avec la vitesse de la lumière.” “La force totale”, he wrote “peut se partager en trois composantes, la première une vague analogie avec la force mécanique due au champ électrique, les deux autres avec la force mécanique due au champ magnétique.”

When Einstein turned to the problem of deriving a relativistic theory of gravity, he introduced the Principle of Equivalence as a fundamental axiom in addition to Lorentz covariance. This states that an observer cannot distinguish between an accelerated reference frame and a gravitational force. By 1907, he had begun to realize that the Principle of Equivalence is incompatible with invariance under Lorentz transformations. Indeed, the Poincaré group (the Lorentz group plus space-time translations) is the symmetry group of the Minkowski metric (see [20], §2.1 for a direct proof); coordinate transformations beyond those in the Poincaré group lead to more general metric tensors, and to the curvature of space-time.

Poincaré died in 1912, just as the battle for the Holy Grail of Mathematical Physics of the era, a relativistic theory of gravitation, was beginning to heat up. Minkowski had already passed away in 1909. The banner of special relativity was taken up by three Knights Errant, Gustav Mie, Gunnar Nordstöm, and Max Abraham in the years 1912-1914.

There is a lingering perception that gravitation cannot be described by special relativity due to the fact that the energy of the gravitational field is negative (Pais [11], chapter 13). The issue was prompted by a remark of Maxwell in his paper *A Dynamical Theory of the Electromagnetic Field* ([7], pp. 570,571). In the section, “Note on the Attraction of Gravitation,” Maxwell compares the energy fields of gravitational attraction and magnetic repulsion of two like magnetic poles. He gives a calculation showing that the “gravitational energy” due to the two bodies is negative, and, though he does not say so explicitly, that it is unbounded below. Maxwell’s argument is cryptic

from today’s perspective, but his conclusion is not:

“The assumption, therefore, that gravitation arises from the action of the surrounding medium in the way pointed out, leads to the conclusion that every part of this medium possesses, when undisturbed, an enormous intrinsic energy, and that the presence of dense bodies influences the medium so as to diminish this energy wherever there is a resultant attraction.

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.”

Abraham [1], citing Maxwell’s comment, showed that “Vector Theories” of gravity are inherently unstable. His analysis is couched in terms of partial differential equations, hence more understandable in modern terms. He takes the partial differential equations of a static gravitational field to be

$$\operatorname{div} \mathbf{F}^g = -\rho, \quad \mathbf{F}^g = -\operatorname{grad} \phi, \quad (45)$$

where ρ is the mass density and ϕ the gravitational potential. He then states that in Newton’s theory the energy of the resulting gravitational field is

$$\mathcal{E} = \frac{1}{2} \iiint \rho \phi \, dv. \quad (46)$$

(This statement is correct; the proof is similar to that in the electrostatic case [4], §37.) Now equations (45) combine to give Poisson’s equation $\Delta\phi = \rho$. Assuming $\rho \in L^1(\mathbb{E}^3)$, the solution is given by the convolution integral

$$\phi(x) = - \int \frac{\rho(y)}{4\pi|x-y|} dy, \quad x, y \in \mathbb{E}^3;$$

and since $\rho \geq 0$, the integral in (46) is negative. It can be written as

$$\mathcal{E} = - \iint \frac{\rho(x)\rho(y)}{8\pi|x-y|} dx dy, \quad x, y \in \mathbb{E}^3. \quad (47)$$

The issue raised by Maxwell and Abraham is succinctly illustrated by a simple example: consider the family of Gaussians

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

These densities have total mass 1, and the energy (47) can be computed explicitly

$$\begin{aligned}\mathcal{E}_\sigma &= - \iiint \frac{dz}{8\pi|z|} \iiint \rho_\sigma(y)\rho_\sigma(y+z)dy \\ &= \iiint \frac{dz}{8\pi|z|} \left(\frac{\sigma}{2\pi}\right)^{3/2} \exp\left(-\frac{\sigma|z|^2}{2}\right) dz = -\frac{1}{4\pi} \left(\frac{\sigma}{2\pi}\right)^{1/2}.\end{aligned}$$

In fact, we have

$$\begin{aligned}\iiint \rho_\sigma(y)\rho_\sigma(y+z)dy &= \left(\frac{\sigma}{\pi}\right)^3 \exp(-\sigma|z|^2) \iiint \exp(-2\sigma(|y|^2 + 2y \cdot z)) dy \\ &= \left(\frac{\sigma}{\pi}\right)^3 \exp\left(-\frac{\sigma|z|^2}{2}\right) \iiint \exp\left(-2\sigma\left|y + \frac{z}{2}\right|^2\right) dy \\ &= \left(\frac{\sigma}{\pi}\right)^3 \exp\left(-\frac{\sigma|z|^2}{2}\right) \left(\int_{-\infty}^{\infty} \exp(-2\sigma t^2) dt\right)^3 \\ &= \left(\frac{\sigma}{2\pi}\right)^{3/2} \exp\left(-\frac{\sigma|z|^2}{2}\right).\end{aligned}$$

This example shows that solutions to Poisson's equation are dynamically unstable: as $\sigma \rightarrow \infty$, the total mass remains the same, but concentrates at the origin, while $\mathcal{E}_\sigma \rightarrow -\infty$. Since matter naturally aggregates under the force of gravity, this implies that matter, in the absence of forces other than gravity, collapses to a point.

Pais suggests that the negative energy issue was the stumbling block to a Lorentz invariant field theory of gravity:

“Maxwell's wise words were not generally heeded, not even by physicists of great stature. Oliver Heaviside discussed the gravitational-electromagnetic analogy without mentioning the negative energy difficulty. So, remarkably, did Lorentz in one of his rare speculative papers . . . As late as 1912, it was still necessary to show that all these vector theories made no sense because of Maxwell's negative energy difficulty.”

The negative energy of the gravitational field was cited by Abraham in his dispute with Einstein over Lorentz invariance as a fundamental principle of physics, and had gotten entangled in the fray. But the negative energy is neither an obstacle to a field theory of gravitation – classical potential theory is a counterexample – nor is it resolved by the General Theory. Indeed the issue is central to modern cosmology. It led Einstein to introduce his cosmological constant in order to “stabilize” the universe. The same instability underlies the theory of black holes, which finds its genesis in the singularity of the Schwarzschild metric.

Einstein’s adherence to the Principle of Equivalence forced him to abandon the Lorentz group in favor of a more general class of observers; while Abraham, Mie, and Nordström continued to hold that Lorentz invariance was a fundamental hypothesis of nature. (see the discussion in Thorne [18], p. 115.) But the issue, like the negative energy issue, is not so clear cut.

If the Principle of Equivalence knocks out the Maxwell-Heaviside gravitational theory, it knocks out Maxwell’s electromagnetic theory as well – something no one, not even Einstein himself, would claim. Strict Lorentz invariance would not allow for force or acceleration in electrodynamics; yet a fundamental deduction from Maxwell’s equations is the existence of electromagnetic waves, generated by accelerated charges.

Accelerated reference frames are in fact implicit in Maxwell’s equations: here are two arguments. First, Maxwell’s equations form a first order hyperbolic system. If the first order system is reduced to a linear second order equation, as Maxwell did, one finds that it contains second order time derivatives, a.k.a. acceleration. Second, the time independent Maxwell equations describe the static fields and are themselves invariant under Lorentz transformations. All inertial observers see a static field, with only the relative strength, position and direction of the electric and magnetic fields changed. Thus, the dynamic and static Maxwell’s equations are each invariant under Lorentz transformations, so to go from one to another requires a transformation to an accelerated reference frame.

The second proof is motivated by a “Machian” *Gedankenexperiment*. Consider a binary star, rotating in the $x - y$ plane, with the center of gravity of the two stars at the origin. Consider two observers on the z -axis, one whose coordinate system is fixed to the x, y, z axes; the other whose coordinate system is tied to the circling stars. The

first sees the two stars rotating about each other; the second sees them sitting motionlessly, suspended in space; but to go from one to the other requires a transformation to an accelerated reference system.

The atmosphere, moving with the Earth, is in an accelerated reference frame and experiences a force, the Coriolis force. The Coriolis force in geophysics is thus a familiar example of Einstein's Principle of Equivalence in Newtonian mechanics. In short, the entire subject of dynamics, inasmuch as it involves force and acceleration, is incompatible with inertial reference frames. Accelerated reference frames are indeed synonymous with the curvature of space-time, but by Einstein's Equivalence Principle, space-time is virtually flat if the acceleration is not too great, and linear theories suffice.

Finally, Einstein's computation of the precession of the perihelion of Mercury was a signature achievement. It validated the General Theory and demonstrated the fundamental nonlinear character of the gravitational field; but it does not invalidate the Maxwell-Heaviside theory in weak fields. Thorne ([18], p. 95) points out that the total advance of the perihelion is 1.38 seconds of arc per revolution, 1.28 of which can be accounted for by Newtonian theory. So Newton's theory is in fact remarkably good! No one is suggesting that Newton's linear, non-relativistic theory be abandoned because of this one small lapse; so why should the Maxwell-Heaviside theory, taken as a linear approximation in a weak field, be rejected on these same grounds?

Abraham's paper contains a comprehensive bibliography of the literature devoted to vector models of gravitation. "This paradoxical conclusion [the negative energy of the field] is characteristic of the vector theory of the gravitational field," writes Abraham. He goes on to discuss the gravitational theories of Lorentz, Poincaré and Minkowski in some detail, concluding that "the laws of interaction of attracting masses drawn up by these pioneers of relativity are accordingly very well reconciled with the vector theory of gravitation sketched above."

9 Acknowledgements

I would like to express my thanks to the Editorial Board for their invitation to submit this paper to the memorial volume for Klaus Kirchgässner. I first visited Klaus in 1970 when he was at the Ruhr-Universität in Bochum. We became instant friends, and we enjoyed many discussions over the years, not only about mathematics, but

about history, music, German Romanticism, people (Kepler, Einstein, Newton, Mozart, etc.) Klaus was a prolific mathematician, with high intellectual standards, and his work, especially his contributions to the theory of nonlinear waves, has set the standard. A mentor, a gracious host, an organizer, a promoter of science and mathematics, he was always interested in what everyone else was doing. He had a wonderful sense of humor. He will always be with us.

Everyone associated with Klaus shared his interest in the applications of mathematics to physical problems. All of us have worked in some area or another of dynamical systems or continuum mechanics, with analysis and differential equations as a central tool. The techniques we acquired are directly applicable to problems in modern Cosmology.

I thank a number of people for useful comments and stimulating conversations which have been incorporated into this article: Jacek Szmigielski, Adrian Korpel, Jim Wheeler, Paul Faris, Joel Smoller, and Jim Rauch.

References

- [1] M. Abraham. Neuere Gravitationstheorien. *Jahrbuch der Radioaktivität und Elektronik*, 11(4):470–505, 1915.
- [2] S.J. Clark and R. W. Tucker. Gauge symmetry and gravitomagnetism. *Classical and Quantum Gravity*, 17:4125–4157, 2000.
- [3] O. Heaviside. A gravitational and electromagnetic analogy. *The Electrician*, 31:281–282, 1893.
- [4] L.D. Landau and E. M. Lifshitz. *The Classical Theory of Fields*. Elsevier, Amsterdam, 1975.
- [5] J. Lenells. Boundary Value Problems for the Stationary Axisymmetric Einstein Equations: A Disk Rotating Around a Black Hole. *Communications of Mathematical Physics*, 304(3):585–635, 2011.
- [6] H.A. Lorentz. Considerations on gravitation. *Proc. Royal Netherlands Academy of Arts and Sciences*, 2:559–574, 1900.
- [7] James Clerk Maxwell. *The Scientific Papers I*, chapter A Dynamical Theory of the Electromagnetic Field. Cambridge University Press, Cambridge, 1890. editor W.D. Niven.

- [8] H. Minkowski. Das Relativitätsprinzip. *Annalen der Physik* 352, 15:927–938, 1915.
- [9] Misner, C., K. Thorne, and J.A. Wheeler. *Gravitation*. Freeman, New York, 1973.
- [10] J.R. Oppenheimer and H. Snyder. On continued gravitational contraction. *Physical Review*, **56**:455–459, 1939.
- [11] A. Pais. *Subtle is the Lord, The Science and Life of Albert Einstein*. Oxford University Press, Oxford, 1982.
- [12] H. Poincaré. Sur la Dynamique de l'Électron. *Rendiconti del Circolo Matematico di Palermo*, **21**:129–176, 1906.
- [13] Rubin, Ford, and Thonnard. Rotational properties of 21 Sc galaxies with a large range of luminosities and radii. *The Astrophysical Journal*, 238:471–187, 1980.
- [14] Rubin, Thonnard, and Ford. NGC 3067: additional evidence for nonluminous matter? *The Astronomical Journal*, 87:477–85, 1982.
- [15] D.H. Sattinger and O.L. Weaver. *Lie groups and Algebras with Applications to Physics, Geometry, and Mechanics*. Springer-Verlag, New York, Berlin, 1986.
- [16] K. Schwarzschild. Über das Gravitationsfeld eines Massenpunktes nach der Einstein'schen Theorie. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, **1**:189–196, 1915.
- [17] J.A. Stratton. *Electromagnetic Theory*. McGraw-Hill, New York and London, 1941.
- [18] K. Thorne. *Black Holes and Time Warps, Einstein's Outrageous Legacy*. Norton, New York, 1994.
- [19] R.C. Tolman. *Proc. Nat. Acad. Sci*, 20, 1934.
- [20] S. Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. John Wiley & Sons, New York, 1972.
- [21] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helv. Phys. Acta*, 6:110–127, 1933.
- [22] F. Zwicky. On the masses of nebulae and clusters of nebulae. *The Astrophysical Journal*, 86:217–246, 1937.