Astrophysical and Cosmological Consequences of a Gauge Theory of Gravity

AUTHORS
Anthony Lasenby
Chris Doran
Stephen Gull

In N. Sanchez and A. Zichichi, editors
Advances in Astrofundamental Physics, Erice 1994
World Scientific Publishing Co., 1995, p. 359-401

Abstract

A new gauge-theory description of gravity is presented, employing gauge fields in a flat background spacetime. These fields ensure that all physical relations are independent of the position and orientation of the matter fields in this background. The language of 'geometric algebra' best expresses the physical and mathematical content of the theory and is employed throughout. A method of working directly with the physical fields is developed and applied to the case of a radially-symmetric time-varying perfect fluid. A gauge is found in which the physics reduces to a set of Newtonian equations. The insistence on finding global solutions alters the physical picture of the horizon around a black hole, and enables one to discuss the properties of field lines inside the horizon created by a point charge held at rest outside it. Some applications to cosmology are discussed, and a study of the Dirac equation in a cosmological background reveals that the only models consistent with homogeneity are spatially flat.

1 Introduction

The most successful theories of physics created to date are underpinned by the principle of local gauge invariance. This principle is based on the belief that global symmetry transformations, in which the transformation is applied simultaneously at all points in spacetime, are incompatible with the requirements of locality. Instead, any global symmetry must be replaced by a local equivalent in which the symmetry holds independently from point to point. Local symmetry is guaranteed by the existence of fields with certain transformation properties — gauge fields — and these fields are responsible for inter-particle forces. But what of gravity: can general relativity be formulated as a gauge theory? This question has troubled physicists for many years [1, 2, 3, 4]. It was Kibble [2] who, elaborating on work by Utiyama [1], was the first to recover aspects of general relativity (GR) from a gauging argument. Kibble used the 10-component Poincaré group of passive coordinate transformations (consisting of four translations and six rotations) as the global symmetry group. Kibble's use of passive transformations was criticised by Hehl *et al.* [5], who reproduced Kibble's derivation from the standpoint of active transformations of the matter fields. In both cases the authors arrived at a more general version of GR known as a 'spin-torsion' theory, and it is now generally accepted that torsion is an inevitable feature of a gauge theory based on the Poincaré group.

There are a number of conceptual difficulties associated with the approaches adopted by Kibble and Hehl *et al.* For example, Hehl *et al.* state that '*coordinates and frames are regarded as fixed once and for all, while the matter fields are replaced by fields that have been rotated or translated*'. It follows that their derivation can only affect the properties of the matter fields, and not the properties of spacetime itself. Yet, once the gauge fields have been introduced, the authors identify these fields as determining the curvature and torsion of a Riemann-Cartan spacetime. Despite being untouched by the derivation of the gravitational fields, spacetime itself is supposed to have become an active participant in physics! Here we propose an alternative gauge theory of gravity which we believe is free from such conceptual difficulties. Our starting point is the parameterisation of events by *vectors* in a 4-dimensional spacetime. This spacetime is 'flat', though it might more appropriately be referred to as 'structureless'. All matter fields are now parameterised by a vector variable x, but in so doing are we not returning to some Newtonian notion of an absolute space? The answer is no — but only if we insist that all physical laws consist of local relationships between fields, and not between a field and its spacetime position. The location of the matter fields in the background spacetime is then irrelevant. We therefore aim to construct a theory where the relationships between fields are invariant under local changes of the position and orientation of fields in our background space. The result is a gauge theory in which spacetime itself does not play an active role in physics.

Besides the conceptual clarity afforded by the theory outlined here, it contains some considerable practical benefits. In particular, once one chooses a gauge by prescribing how matter fields are parameterised by vectors, one is then free to utilise all the advantages of working in a flat space. These advantages are best exploited using the language of 'geometric algebra' [6] and its spacetime version 'spacetime algebra' [7]. It is well known that many problems in two and three dimensions can be handled most efficiently by representing points with vectors and working in a coordinate-free manner. Spacetime algebra generalises these techniques so that they are applicable to problems in spacetime. Naturally associated with the representation of points by vectors is the vector derivative — the derivative with respect to vector position. The vector derivative plays a central role in the spacetime algebra form of both the Maxwell and Dirac theories. Furthermore, in both cases the derivative appears in a first-order form for which an inverse function can be found (in the form of a Green's function). The resultant first-order propagator theory is highly efficient computationally [6, 8]. The theory developed here enables these same techniques to be employed in gravitational problems, and the gauge structure ensures that all physical predictions are independent of the means by which points were parameterised with vectors. In addition to these 'flatspace' techniques, we have developed a new technique which deals directly with the physically-observable objects. We call this the '*intrinsic*' method. The method consists of manipulating abstract operators satisfying certain bracket relations, and is remarkably powerful in many applications.

This paper is intended to offer an outline review of our theory. The full formal development is contained in [9]. We begin with an introduction to geometric algebra and its application to spacetime physics. We then turn to a discussion of the gauge fields that must be introduced to construct a theory in which all relations between physical fields are independent of their position and orientation in the background space. Having determined the necessary fields, we discuss the appropriate field equations for them. These equations are then studied for the case of a radially-symmetric perfect fluid. We find that a particular choice of time variable leads to some dramatic simplifications and results in a set of equations which are almost entirely Newtonian! A number of applications of this system are then discussed.

Our employment of a background vector space forces us to rethink the physics of horizons and their formation during a collapse process. Not surprisingly, since we work with a background flat spacetime, the possibility of wormholes to new universes does not arise! The example of electromagnetism reveals the importance of finding global solutions to the field equations, which is a crucial point where our theory departs from GR. We then turn to applications in cosmology. The 'intrinsic' method leads quickly to a pair of first-order equations which work directly with the Hubble velocity. The Friedmann equations are easily recovered and a novel treatment of particle horizons is given. Finally, studying the Dirac equation in a cosmological background, we conclude that the only cosmological models consistent with spatial homogeneity are spatially flat. We end with a summary of the main conclusions.

2 Geometric Algebra

The following comprises a brief introduction to the ideas and conventions of geometric algebra. Further details can be found in the series of papers [8, 10, 11, 12], in the books by Hestenes [7, 13] and Hestenes & Sobczyk [6] and in the introductory articles by Hestenes [14, 15] and Vold [16, 17]. The applications to a gauge theory of gravity were first discussed in [18] and [19].

A 'geometric algebra' is a graded vector space with an associative product that is distributive over addition. The grade-0 elements of this space are real scalars and commute with all higher-grade elements (known as 'multivectors'). The grade-1 elements are vectors and are usually given lower-case Roman symbols (a, b). The geometric product is distinguished by the property that the square of any vector is a scalar. It then follows from the identity

$$(a+b)^2 = (a+b)(a+b) = a^2 + (ab+ba) + b^2 \tag{1}$$

that the symmetrised product of two vectors is a scalar. We use this to decompose the geometric product of two vectors into a scalar term

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

and a grade-2 term, called a *bivector*,

$$a\wedge b \equiv \tfrac{1}{2}(ab-ba). \tag{3}$$

Hence we now have

$$ab = a\cdot b + a\wedge b \tag{4}$$

for the full geometric product. The significant feature of the geometric product is that it mixes two different types of object: scalars and bivectors. This is not problematic — the addition implied by (4) is that used when, for example, a real number is added to an imaginary number to form a complex number. Indeed, it was Clifford's desire to extend the complex field that led him to propose the geometric product.

Forming further geometric products of vectors produces higher-grade multivectors. Multivectors in which all elements have the same grade are termed *homogeneous* and are usually written as A_r to show that A contains only a grade-r component. Multivectors inherit an associative product, and the geometric product of a grade-r multivector A_r with a grade-s multivector B_s decomposes into

$$A_r B_s = \langle AB\rangle_{r+s} + \langle AB\rangle_{r+s-2} \ldots + \langle AB\rangle_{|r-s|}. \tag{5}$$

The symbol $\langle M\rangle_r$ denotes the projection onto the grade-r component of M. The projection onto the grade-0 (scalar) component of M is written $\langle M\rangle$. The scalar part of a product of multivectors satisfies the cyclic reordering property

$$\langle A \ldots BC\rangle = \langle CA \ldots B\rangle. \tag{6}$$

The '$\cdot$' and '$\wedge$' symbols are retained for the lowest-grade and highest-grade terms of the series (5), so that

$$\begin{aligned} A_r\cdot B_s &\equiv \langle AB\rangle_{|s-r|} && (7)\\ A_r\wedge B_s &\equiv \langle AB\rangle_{s+r}, && (8) \end{aligned}$$

which are called the interior and exterior products respectively. We also define the commutator product

$$A\times B \equiv \tfrac{1}{2}(AB - BA). \tag{9}$$

The associativity of the geometric product ensures that the commutator product satisfies the Jacobi identity

$$A\times(B\times C) + B\times(C\times A) + C\times(A\times B) = 0. \tag{10}$$

Finally, we introduce an operator ordering convention. *In the absence of brackets,* *inner, outer and scalar products take precedence over geometric products.* Thus $a{\cdot}b\,c$ means $(a{\cdot}b)\,c$, not $a{\cdot}(bc)$. This convention helps to eliminate unwieldy numbers of brackets. Summation convention and natural units ($\hbar = c = \epsilon_0 = G = 1$) are employed throughout.

2.1 The Spacetime Algebra

Of central importance to many applications is the geometric algebra of spacetime, the *spacetime algebra* [7]. To describe the spacetime algebra (STA) it is helpful to introduce a set of four orthonormal basis vectors $\{\gamma_\mu\}$, $\mu = 0 \ldots 3$, satisfying

$$\gamma_\mu{\cdot}\gamma_\nu = \eta_{\mu\nu} = \mathrm{diag}(+ - - -). \tag{11}$$

The vectors $\{\gamma_\mu\}$ satisfy the same algebraic relations as Dirac's γ-matrices, but they now form a set of four independent basis vectors for spacetime, not four components of a single vector in an internal 'spin-space'. The relation between Dirac's matrix algebra and the STA is described in more detail elsewhere [11].

The full STA is spanned by the quantities

$$1, \qquad \{\gamma_\mu\}, \qquad \{\sigma_k,\ i\sigma_k\}, \qquad \{i\gamma_\mu\}, \qquad i, \tag{12}$$

where

$$\sigma_k \equiv \gamma_k\gamma_0, \tag{13}$$

and

$$i \equiv \gamma_0\gamma_1\gamma_2\gamma_3 = \sigma_1\sigma_2\sigma_3. \tag{14}$$

The $\{\sigma_k\}$ form an orthonormal frame of spatial vectors in the space relative to the γ_0 direction. The algebraic properties of the $\{\sigma_k\}$ are the same as those of the Pauli spin matrices, though again they are to be interpreted geometrically and not as components of a vector in spin-space. The highest-grade element (or 'pseudoscalar') is denoted by i. The symbol i is used because the square of i is -1, but the pseudoscalar must not be confused with the unit scalar imaginary employed in quantum mechanics. Since we are working with a space of even dimension, i *anti*commutes with odd-grade elements, and only commutes with even-grade elements.

The split of the six spacetime bivectors into relative vectors $\{\sigma_k\}$ and relative bivectors $\{i\sigma_k\}$ is a frame-dependent operation — different observers determine different relative spaces. This fact is best illustrated with the Faraday bivector F.

The 'spacetime split' [7, 20] of F into the γ_0-system is made by separating F into parts which anticommute and commute with γ_0. Thus

$$F = \boldsymbol{E} + i\boldsymbol{B} \tag{15}$$

where

$$\boldsymbol{E} = \tfrac{1}{2}(F - \gamma_0 F \gamma_0) \tag{16}$$

$$i\boldsymbol{B} = \tfrac{1}{2}(F + \gamma_0 F \gamma_0). \tag{17}$$

Both $\boldsymbol{E}$ and $\boldsymbol{B}$ are spatial vectors in the γ_0-frame, and $i\boldsymbol{B}$ is a spatial bivector. Equation (15) decomposes F into separate electric and magnetic fields, and the explicit appearance of γ_0 in the formulae for $\boldsymbol{E}$ and $\boldsymbol{B}$ shows how this split is observer-dependent. Where required, relative (or spatial) vectors in the γ_0-system are written in bold type to record the fact that in the STA they are actually bivectors. This distinguishes them from spacetime vectors, which are left in normal type. No problems arise for the $\{\sigma_k\}$, which are unambiguously spacetime bivectors, so these are also left in normal type.

2.2 Geometric Calculus

Throughout this paper we employ the derivative with respect to a vector argument a. This is written as ∂_a and is defined in terms of its directional derivatives $b \cdot \partial_a$, where

$$b \cdot \partial_a F(a) \equiv \lim_{\tau \mapsto 0} \frac{F(a + \tau b) - F(a)}{\tau}. \tag{18}$$

Then, in terms of an arbitrary vector basis $\{e_k\}$ and reciprocal basis $\{e^k\}$ ($e_j \cdot e^k = \delta_j^k$) the full derivative is defined by

$$\partial_a \equiv e^j e_j \cdot \partial_a. \tag{19}$$

This definition shows how the derivative ∂_a inherits the vector properties of its argument a, as well as a calculus from equation (18). The vector derivative is employed frequently to manipulate linear functions in a frame-independent manner.

The derivative with respect to spacetime position x is of special significance. It is called the *vector derivative*, and is given the symbol

$$\nabla = \nabla_x \equiv \partial_x. \tag{20}$$

In terms of a set of Cartesian coordinates $\{x^\mu\}$, the vector derivative can be written

$$\nabla = \gamma^\mu \partial_\mu, \tag{21}$$

where $\partial_\mu \equiv \partial_{x^\mu}$. Hence, just as the γ-matrices are replaced by vectors in spacetime, objects such as $x^\mu \gamma_\mu$ and $\nabla = \gamma^\mu \partial_\mu$ become frame-free vectors. The usefulness of the geometric product for the vector derivative is illustrated by electromagnetism. In tensor notation, Maxwell's equations take the form

$$\partial_\mu F^{\mu\nu} = J^\nu \qquad \partial_{[\alpha} F_{\mu\nu]} = 0, \tag{22}$$

which have the STA equivalents

$$\nabla \cdot F = J \qquad \nabla \wedge F = 0. \tag{23}$$

But now we can utilise the geometric product to combine these into the single equation

$$\nabla F = J. \tag{24}$$

The significance of the ∇ operator is that it possesses an inverse, so a first-order propagator theory can be developed for it [6, 8]. This is not possible for the individual $\nabla \cdot$ and $\nabla \wedge$ operators.

The vector derivative acts on objects to its immediate right unless brackets are present. So, in the expression ∇AB the ∇ acts only on A, but in the expression $\nabla(AB)$ the ∇ acts on both A and B. If the ∇ is intended to only act on B then this is written as $\dot{\nabla} A \dot{B}$, where the overdot denotes the multivector on which the derivative acts. For example, Leibniz' rule can be written in the form

$$\nabla(AB) = \dot{\nabla} \dot{A} B + \dot{\nabla} A \dot{B}. \tag{25}$$

2.3 Linear Algebra

Geometric algebra offers many advantages over tensor calculus when used for developing the theory of linear functions [6, 21, 15]. A linear function mapping vectors to vectors is written with an underbar $\underline{f}(a)$. The adjoint function is written with an overbar, $\overline{f}(a)$, so that

$$a \cdot \underline{f}(b) = \overline{f}(a) \cdot b, \tag{26}$$

and hence

$$\overline{f}(a) = \partial_b \langle \underline{f}(b) a \rangle. \tag{27}$$

As explained, the use of the derivative operator ∂_b keeps all expressions free from requiring an explicit coordinate frame. Of course, the ∂_b and b vectors can be replaced by the sum over a set of frame vectors and their reciprocals, if desired.

Linear functions extend to act on multivectors via

$$\underline{f}(a\wedge b\wedge \ldots \wedge c) \equiv \underline{f}(a)\wedge \underline{f}(b) \ldots \wedge \underline{f}(c), \tag{28}$$

so that $\underline{f}$ is now a grade-preserving linear function mapping multivectors to multivectors. In particular, since the pseudoscalar of a space, I say, is unique up to a scale factor, we can define

$$\det \underline{f} = \underline{f}(I) I^{-1}. \tag{29}$$

Viewed as linear functions over the entire geometric algebra, $\underline{f}$ and $\overline{f}$ are related by the fundamental formulae

$$\begin{aligned} A_r \cdot \overline{f}(B_s) &= \overline{f}[\underline{f}(A_r) \cdot B_s] \qquad r \leq s \\ \underline{f}(A_r) \cdot B_s &= \underline{f}[A_r \cdot \overline{f}(B_s)] \qquad r \geq s, \end{aligned} \tag{30}$$

which are derived in [6, Chapter 3]. The formulae for the inverse functions are found as special cases of (30),

$$\begin{aligned} \underline{f}^{-1}(A) &= \det(f)^{-1} \overline{f}(AI) I^{-1} \\ \overline{f}^{-1}(A) &= \det(f)^{-1} I^{-1} \underline{f}(IA). \end{aligned} \tag{31}$$

Of particular importance here is the geometric algebra description of rotations. A vector a is rotated to a new vector a' by

$$a' = R a \tilde{R} \tag{32}$$

where R is a *rotor*. In the STA, rotors are elements of the even subalgebra satisfying

$$R\tilde{R} = 1 \tag{33}$$

and any rotor can be written as $\pm \exp\{B\}$, where B is a bivector. The usefulness of rotors in the description of rotations lies in the ease with which they extend to act on multivectors. So, for example,

$$(Ra\tilde{R})\wedge(Rb\tilde{R}) = \tfrac{1}{2}(Rab\tilde{R} - Rba\tilde{R}) = Ra\wedge b\tilde{R} \tag{34}$$

and the generalisation to an arbitrary multivector is simply

$$M \mapsto RM\tilde{R}. \tag{35}$$

Furthermore, this formula is equally valid for boosts as well as spatial rotations. No other mathematical language affords such a simple description of Lorentz transformations.

3 The Gravitational Gauge Fields

In this Section we identify the dynamical variables which describe gravitational interactions. Our aim is to achieve a theory in which all reference to spacetime position is removed, and all that remains are the intrinsic relations between fields at a point (these fields will include the gravitational field). The ‘position’ in spacetime of an object will then be defined *intrinsically* by the values of the fields at that position, and not by a set of arbitrary coordinates. At the same time, we do not want to lose the advantages of representing spacetime positions by vectors. To satisfy both of these demands, we must require that the actual position vector of a field be irrelevant. That is, all equations should be unchanged in form if the fields are moved from one spacetime position to another. Furthermore, these changes of position must be local in character if the physics at a point is to be genuinely free of the position vector of the point. In Section 2.2 we saw how to combine Maxwell’s equations into the single equation $\nabla F = J$. The STA form of the Dirac equation [11, 22] also employs the same differential operator ∇ — the vector derivative. We must therefore focus attention on this operator to see how to make our field equations local.

We start by considering a scalar field $\phi(x)$ and form its vector derivative $\nabla\phi(x)$. Suppose now that $\phi(x)$ is moved around to form a new field $\phi'(x)$,

$$\phi'(x) \equiv \phi(x'), \tag{1}$$

where

$$x' = f(x) \tag{2}$$

and $f(x)$ is an arbitrary (differentiable) map between spacetime position vectors. The map $f(x)$ should not be thought of as a map between manifolds, or as moving points around. The function $f(x)$ is just a rule for relating one position vector to another within a single vector space. If we now consider ∇ acting on the new scalar field ϕ' we form the quantity $\nabla\phi[f(x)]$. To evaluate this we return to the definition of the vector derivative and construct

$$\begin{aligned} a\cdot\nabla\phi[f(x)] &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(\phi f(x+\epsilon a) - \phi f(x)\right) \\ &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big(\phi[f(x)+\epsilon\underline{f}(a)] - \phi f(x)\Big) \\ &= \underline{f}(a)\cdot\nabla_{x'}\phi(x'), \end{aligned} \tag{3}$$

where $\underline{f}(a) = \underline{f}_x(a) = a\cdot\nabla f(x)$ and the subscript on $\nabla_{x'}$ records that the derivative is now with respect to the new vector position variable x'. It follows that

$$\nabla_x = \overline{f}(\nabla_{x'}) \tag{4}$$

and hence that $\nabla\phi'(x) = \overline{f}[\nabla_{x'}\phi(x')]$.

In physics we frequently need to relate the gradient of some scalar to a vector field. In three dimensions, for example, a static electric field can be written as the gradient of the scalar potential ϕ. We wish to move to a situation where all relations are unaffected by arbitrary displacements. To achieve this we must introduce a gauge field which assembles with the vector derivative to form an object which just changes its position dependence under local translations. We construct such an object by replacing ∇ with a new derivative $\overline{h}(\nabla)$, where $\overline{h}(a)$ has an arbitrary position dependence and is a linear function of a. If we wish to make the position dependence explicit we will write $\overline{h}_x(a)$ or $\overline{h}(a,x)$ (recalling that $\overline{h}(a,x)$ is linear on a and non-linear on x). Under local translations the gauge field $\overline{h}(a)$ is defined to transform to the new field $\overline{h}'(a)$, where

$$\overline{h}'(a,x) \equiv \overline{h}(\overline{f}^{-1}(a), f(x)) = \overline{h}_{x'}\overline{f}^{-1}(a) \tag{5}$$

so that

$$\overline{h}_x(\nabla_x) \mapsto \overline{h}_{x'}\overline{f}^{-1}(\nabla_x) = \overline{h}_{x'}(\nabla_{x'}). \tag{6}$$

This transformation law ensures that the vector $\overline{h}(\nabla\phi(x))$ now translates as

$$\overline{h}(\nabla\phi(x)) \mapsto \overline{h}_{x'}[\nabla_{x'}\phi(x')] \tag{7}$$

and hence just changes its vector position under *arbitrary* local translations. This is the type of behaviour we are after, since ultimately all that should matter are the relations between fields at a point, and the actual position vector of that point (and its relation to nearby points) should be irrelevant. The same intrinsic relations should therefore hold if the fields are moved arbitrarily from one vector position to another. The introduction of the $\overline{h}$-field ensures that derivatives can also be moved around arbitrarily. The $\overline{h}$-field is not a connection in the conventional Yang-Mills sense. The coupling to derivatives is different, as is the transformation law (5). This is unsurprising, since the group of arbitrary translations $x \mapsto f(x)$ is infinite-dimensional (if we were considering maps between manifolds then this would form the group of diffeomorphisms). There is no doubt, however, that the $\overline{h}$-field derived from this group embodies the idea of replacing a global symmetry by a local one, so clearly deserves to be called a gauge field.

Henceforth, we refer to any quantity that just changes its position dependence under local translations as behaving *covariantly* under translations. The $\overline{h}$-field enables us to form derivatives of covariant objects which are also covariant under translations. When we come to calculate with this theory, we fix a gauge by choosing a labeling of spacetime points with vectors. In this way we are still free to exploit all the advantages of representing points with vectors. Of course, all the physical predictions of the theory will remain independent of the actual gauge choice.

Having arrived at an operator which transforms covariantly under local translations, we must now consider rotations (*i.e.* Lorentz transformations). Returning to Maxwell's equations in the form $\nabla F = J$, we see that if F is a solution with current J, then

$$F' = RF(x')\tilde{R}, \quad x' = \tilde{R}xR \tag{8}$$

is a rotated solution with current $J' = RJ(x')\tilde{R}$. When the ∇ operator is replaced by $\overline{h}(\nabla)$ we have already taken care of covariance under the displacement from x to x', so the rotation of spacetime position is no longer necessary. All that remains of the transformation is the rotation of F at a point. The reason that this is a symmetry is as follows. Suppose we have two covariant multivector fields A and B satisfying the equation $A(x) = B(x)$. Then the rotated fields $A'(x) = RA(x)\tilde{R}$ and $B'(x) = RB(x)\tilde{R}$ also satisfy the same equation. Hence the intrinsic physical relation between A and B is independent of how we choose to represent their directions in spacetime. This is as it must be if we are to eliminate all spacetime-dependence from the final relations. For the case of the Maxwell equations, the fields F and J are now related by $\overline{h}(\nabla)F = J$. If we rotate the F and J fields, this equation will only remain satisfied if we also rotate the $\overline{h}(a)$ field. This makes sense, because the quantity $\overline{h}(\nabla\phi)$ is also covariant under translations, so we should be free to form relations between this and other covariant objects. We are therefore led to the transformation law for the $\overline{h}(a)$ field under rotations at a point:

$$\overline{h}(a) \mapsto R\overline{h}(a)\tilde{R}, \tag{9}$$

which ensures that the equation $\overline{h}(\nabla)F = J$ is now covariant under a global rotation of $\overline{h}(a)$, F and J.

We have now achieved a complete decoupling of rotations and translations, a situation which is not reached in derivations based solely on infinitesimal transformations [2, 5]. But we are still only part way to achieving our goal. The $A(x)$ and $B(x)$ fields above can be rotated by different amounts at each point in space without changing the intrinsic content of the equation $A(x) = B(x)$. Hence we must demand that the equations are invariant under local rotations, where the rotor R is an arbitrary function of position. To see how to achieve this we replace $\overline{h}(\nabla)$ by $\overline{h}(\partial_a)a\cdot\nabla$ and focus attention on the $a\cdot\nabla$ operator. Acting on a multivector $A(x)$ which has been subjected to a position-dependent rotation we find that

$$a\cdot\nabla(RA\tilde{R}) = Ra\cdot\nabla A\tilde{R} + a\cdot\nabla RA\tilde{R} + RAa\cdot\nabla\tilde{R}. \tag{10}$$

Since $R\tilde{R} = 1$ for a rotation we have

$$a\cdot\nabla R\tilde{R} = -Ra\cdot\nabla\tilde{R}, \tag{11}$$

hence $a\cdot\nabla R\tilde{R}$ is equal to minus its reverse and so must be a bivector (an element of the Lie algebra of the rotation group). We can therefore write

$$a\cdot\nabla(RA\tilde{R}) = Ra\cdot\nabla A\tilde{R} + 2(a\cdot\nabla R\tilde{R})\times(RA\tilde{R}). \tag{12}$$

To construct a covariant derivative we must therefore add a 'connection' term to $a\cdot\nabla$ to construct the operator

$$\mathcal{D}_a = a\cdot\nabla + \Omega(a)\times\,. \tag{13}$$

Here $\Omega(a) = \Omega(a, x)$ is a bivector-valued linear function of a with an arbitrary x-dependence. The operation of commuting a multivector with a bivector is grade-preserving. So, even though it is not a scalar operator, $\mathcal{D}_a$ preserves the grade of the multivector on which it acts.

Under local rotations we demand that $\Omega(a)$ transforms as

$$\Omega(a) \mapsto \Omega'(a) = R\Omega(a)\tilde{R} - 2a\cdot\nabla R\tilde{R}. \tag{14}$$

Since $\Omega(a)$ is now an arbitrary function of position, however, it cannot in general be transformed away by the application of a rotor. The equations (13) and (14) ensure that under under a local rotation

$$\mathcal{D}_a{}'(RA\tilde{R}) = R\mathcal{D}_a A\tilde{R}, \tag{15}$$

as required for a covariant derivative. Furthermore, the operator (13) is a *derivation* since it satisfies Leibniz' rule

$$\mathcal{D}_a(AB) = (\mathcal{D}_a A)B + A(\mathcal{D}_a B), \tag{16}$$

as follows from the identity

$$\Omega(a)\times(AB) = (\Omega(a)\times A)B + A(\Omega(a)\times B). \tag{17}$$

Under local translations $\Omega(a)$ must transform in the same way as $a\cdot\nabla R\tilde{R}$, hence

$$\Omega_x(a) \mapsto \Omega_{x'}\underline{f}(a) = \Omega(\underline{f}(a), f(x)), \tag{18}$$

where the subscript is again used to label position dependence. Since we wish to deal with quantities that are covariant under translations, we define the bivectors

$$\omega(a) = \Omega\underline{h}(a), \tag{19}$$

so that $\omega(a,x) \mapsto \omega(a,x')$ under a local translation. In addition, we introduce the directional derivative L_a,

$$L_a \equiv a\cdot\overline{h}(\nabla) \tag{20}$$

and the covariant directional derivative $a\cdot\mathcal{D}$,

$$a\cdot\mathcal{D}A \equiv a\cdot\overline{h}(\nabla)A + \omega(a)\times A. \tag{21}$$

(Note that this notation implies that $\underline{h}^{-1}(a)\cdot\mathcal{D} = \mathcal{D}_a$.) From (21) we define the covariant extension of the vector derivative by

$$\mathcal{D}A = \partial_a a\cdot\mathcal{D}A. \tag{22}$$

As with the vector derivative, $\mathcal{D}$ inherits the algebraic properties of a vector. We can therefore write

$$\mathcal{D}A = \mathcal{D}\cdot A + \mathcal{D}\wedge A, \tag{23}$$

where

$$\mathcal{D}\cdot A \equiv \partial_a\cdot(a\cdot\mathcal{D}A) \tag{24}$$

$$\mathcal{D}\wedge A \equiv \partial_a\wedge(a\cdot\mathcal{D}A). \tag{25}$$

General considerations have led us to the introduction of two new gauge fields, the $\overline{h}(a,x)$ linear function and the $\Omega(a,x)$ bivector-valued linear function, both of which are arbitrary functions of position vector x. This gives a total of $4\times4+4\times6 = 40$ scalar degrees of freedom. The $\overline{h}(a)$ and $\Omega(a)$ fields are incorporated into the vector derivative to form the operator $\mathcal{D} = \overline{h}(\partial_a)\mathcal{D}_a$, which acts covariantly on multivector fields. Thus we can begin to construct equations whose intrinsic content is free from the manner in which we choose to represent spacetime positions with vectors.

4 The Field Equations

Having introduced the $\overline{h}$- and Ω-fields, we look to construct a covariant set of field equations. We start by defining the field-strength via

$$R(a\wedge b)\times A \equiv [\mathcal{D}_a, \mathcal{D}_b]A, \tag{1}$$

$$\implies R(a\wedge b) = a\cdot\nabla\Omega(b) - b\cdot\nabla\Omega(a) + \Omega(a)\times\Omega(b). \tag{2}$$

$R(a\wedge b)$ is a bivector-valued linear function of its bivector argument $a\wedge b$. We write this as $R(B)$, where B denotes an arbitrary bivector. As always, the position dependence can be made explicit by writing $R(B,x)$ or $R_x(B)$.

The definition (2) ensures that under local rotations $R(B)$ transforms as

$$R(B) \mapsto R'(B) = RR(B)\tilde{R}, \tag{3}$$

and under local translations $R(B)$ transforms as

$$R(B) \mapsto R'(B) = R(\underline{f}(B), f(x)). \tag{4}$$

A covariant quantity can therefore be constructed by defining

$$\mathcal{R}(B) \equiv R\underline{h}(B). \tag{5}$$

Under local translations and rotations $\mathcal{R}(B)$ has the following transformation laws:

$$\begin{array}{rrcl} \text{Translations:} & \mathcal{R}'(B,x) & = & \mathcal{R}(B,x') \\ \text{Rotations:} & \mathcal{R}'(B,x) & = & R\mathcal{R}(\tilde{R}BR)\tilde{R}. \end{array} \tag{6}$$

Any linear function with transformation laws of this type is referred to as a covariant tensor. We have begun to employ a notation which is very helpful for the theory developed here. Covariant quantities such as $\mathcal{R}(B)$ and $\mathcal{D}$ are written with calligraphic ('curly') symbols. This helps keep track of the true covariant quantities, so we can easily read off the 'intrinsic' physical relations between fields.

It is useful to extend (2) to a form that gives $\mathcal{R}(B)$ directly. This is achieved by writing

$$\mathcal{R}(a\wedge b) = L_a\omega(b) - L_b\omega(a) + \omega(a)\times\omega(b) - \omega(c), \tag{7}$$

where

$$c = \underline{h}^{-1}(L_a\underline{h}(b) - L_b\underline{h}(a)). \tag{8}$$

From $\mathcal{R}(B)$ we define the following contractions:

$$\begin{array}{rrcl} \text{Ricci Tensor:} & \mathcal{R}(b) & = & \partial_a\cdot\mathcal{R}(a\wedge b) \\ \text{Ricci Scalar:} & \mathcal{R} & = & \partial_a\cdot\mathcal{R}(a) \\ \text{Einstein Tensor:} & \mathcal{G}(a) & = & \mathcal{R}(a) - \frac{1}{2}a\mathcal{R}. \end{array} \tag{9}$$

The use of the symbol $\mathcal{R}$ for all of the Riemann tensor, Ricci tensor and Ricci scalar does not pose a problem, since the argument always specifies which is required. Both $\mathcal{R}(a)$ and $\mathcal{G}(a)$ are covariant tensors, since they inherit the same transformation properties as $\mathcal{R}(B)$.

The natural way to proceed now is to construct an invariant action integral and vary this with respect to the $\overline{h}$- and ω-functions. This is the approach adopted in [9] and here we simply quote the necessary results. In the full gauge-theoretic treatment, gravity is introduced by minimally coupling the $\overline{h}$- and ω-fields to the Dirac action. The only action integral which results in a self-consistent theory is the Hilbert action (together with a possible cosmological term)

$$S = \int |d^4x| \, \det \underline{h}^{-1}(\tfrac{1}{2}\mathcal{R} + \Lambda - \kappa\mathcal{L}_m), \tag{10}$$

where $\mathcal{L}_m$ describes the matter content, Λ is the cosmological term and $\kappa = 8\pi G$. In the absence of spin, the equations obtained from varying this equation are

$$\mathcal{D}\wedge\overline{h}(a) = \overline{h}(\nabla\wedge a) \tag{11}$$

and

$$\mathcal{G}(a) - \Lambda a = \kappa\mathcal{T}(a), \tag{12}$$

where $\mathcal{T}(a)$ is the matter stress-energy tensor. The latter equation shows that our gauge theory is very close to standard GR. Indeed, one can recover the field equations of GR by introducing a set of coordinates $\{x^\mu\}$ with an associated coordinate frame $\{e_\mu\}$. A metric tensor is then defined by

$$g_{\mu\nu} = \underline{h}^{-1}(e_\mu)\cdot\underline{h}^{-1}(e_\nu). \tag{13}$$

But crucial differences do exist between the gauge theory developed here and conventional GR. One of these is immediately apparent from our use of a flat background, which is that the differential equations (11) and (12) have equivalent integral forms. This is particularly significant for the treatment of singularities in our theory [9]. Further differences will emerge when we consider some applications in Section 6.

The significance of equation (11) becomes clearer once one forms

$$\begin{aligned} \overline{h}(\dot{\nabla})\wedge\dot{\overline{h}}(c) &= -\partial_d\wedge(\omega(d)\cdot\overline{h}(c)) \\ \Longrightarrow \quad \langle b\wedge a\overline{h}(\dot{\nabla})\wedge\dot{\overline{h}}(c)\rangle &= -\langle b\wedge a\partial_d\wedge(\omega(d)\cdot\overline{h}(c))\rangle \\ \Longrightarrow \quad [\dot{L}_a\dot{\underline{h}}(b) - \dot{L}_b\dot{\underline{h}}(a)]\cdot c &= [a\cdot\omega(b) - b\cdot\omega(a)]\cdot\overline{h}(c) \end{aligned} \tag{14}$$

where, as always, the overdots determine the scope of a differential operator. The left-hand side of equation (14) contains a term that appears in the commutator of L_a and L_b, which suggests that we form

$$\begin{aligned} [L_a, L_b] &= [L_a\underline{h}(b) - L_b\underline{h}(a)]\cdot\nabla \\ &= [\dot{L}_a\dot{\underline{h}}(b) - \dot{L}_b\dot{\underline{h}}(a)]\cdot\nabla + (L_a b - L_b a)\cdot\overline{h}(\nabla) \\ &= [a\cdot\omega(b) - b\cdot\omega(a) + L_a b - L_b a]\cdot\overline{h}(\nabla). \end{aligned} \tag{15}$$

We can therefore write

$$[L_a, L_b] = L_c \tag{16}$$

where

$$c = a\cdot\omega(b) - b\cdot\omega(a) + L_a b - L_b a. \tag{17}$$

This 'bracket' structure summarises the content of (11). It also simplifies the definition of $\mathcal{R}(a \wedge b)$ (7) since the vector c defined by equation (8) can now be found from equation (17).

The general technique we use for studying the field equations is to let $\omega(a)$ contain a set of arbitrary functions, and then use (17) to find relations between them. In doing so we lose some of the information contained in the 'wedge' equation (11). This information is recovered by employing the symmetry properties of $\mathcal{R}(B)$ and the Bianchi identity. The symmetry of $\mathcal{R}(B)$ is summarised by the single equation

$$\partial_a\wedge\mathcal{R}(a\wedge b). \tag{18}$$

As $\mathcal{R}(B)$ maps bivectors to bivectors it has, at most, 36 degrees of freedom. Equation (18) gives a set of 16 scalar equations, reducing the number of degrees of freedom in $\mathcal{R}(B)$ to 20 — the expected number for the Riemann tensor. It is notable how easy this calculation is in geometric algebra! The Bianchi identity follows from a simple application of the Jacobi identity, and can be written

$$\partial_a\wedge(a\cdot\mathcal{D}\mathcal{R}(B) - \mathcal{R}(a\cdot\mathcal{D}B)) = 0. \tag{19}$$

In terms of the L_a and $\omega(a)$, equation (19) becomes

$$\partial_a\wedge[L_a\mathcal{R}(B) - \mathcal{R}(L_a B) + \omega(a)\times\mathcal{R}(B) - \mathcal{R}(\omega(a)\times B)] = 0. \tag{20}$$

The contracted Bianchi identity is

$$\partial_a\cdot[L_a\mathcal{G}(b) - \mathcal{G}(L_a b) + \omega(a)\times\mathcal{G}(b) - \mathcal{G}(\omega(a)\times b)] = 0, \tag{21}$$

and it follows from (12)) that the matter stress-energy tensor must satisfy the same equation. This is the covariant version of conservation of the stress-energy tensor.

The final relation we need before applying this formalism is to invert (11) to give $\omega(a)$ in terms of $\overline{h}(a)$. We first write equation (11) as

$$\overline{h}(\dot{\nabla})\wedge\dot{\overline{h}}\,\overline{h}^{-1}(a) + \partial_b\wedge(\omega(b)\cdot a) = 0. \tag{22}$$

On defining

$$B(a) \equiv -\overline{h}(\dot{\nabla})\wedge\dot{\overline{h}}\,\overline{h}^{-1}(a) = \overline{h}(\dot{\nabla}\wedge\dot{\overline{h}}^{-1}(a)) \tag{23}$$

equation (4) becomes

$$\partial_b\wedge(\omega(b)\cdot a) = B(a). \tag{24}$$

We solve this equation by first 'protracting' to give

$$\partial_a\wedge\partial_b\wedge(\omega(b)\cdot a) = 2\partial_b\wedge\omega(b) = \partial_b\wedge B(b) \tag{25}$$

and then dotting with a, leaving

$$\omega(a) - \partial_b\wedge(a\cdot\omega(b)) = \tfrac{1}{2}a\cdot(\partial_b\wedge B(b)). \tag{26}$$

Hence, using equation (24) again, we find that

$$\omega(a) = -B(a) + \tfrac{1}{2}a\cdot(\partial_b\wedge B(b)). \tag{27}$$

This completes the required set of equations.

5 Spherically-Symmetric Matter Distributions

We now apply the formalism of Section 4 to a time-dependent radially-symmetric perfect fluid. We temporarily drop the cosmological term from equation (12). It will be replaced when we come to consider cosmological applications. The first step is to introduce a set of polar coordinates. In terms of the fixed $\{\gamma_\mu\}$ frame we define:

$$\begin{aligned} t &\equiv x\cdot\gamma_0 \\ r &\equiv \sqrt{(x\wedge\gamma_0)^2} \\ \cos\theta &\equiv x\cdot\gamma^3/r \\ \tan\phi &\equiv (x\cdot\gamma^2)/(x\cdot\gamma^1). \end{aligned} \tag{1}$$

The associated coordinate frame is

$$\begin{aligned} e_t &\equiv \gamma_0 \\ e_r &\equiv x\wedge\gamma_0\,\gamma_0/r \\ e_\theta &\equiv r\cos\theta(\cos\phi\,\gamma_1 + \sin\phi\,\gamma_2) - r\sin\theta\,\gamma_3 \\ e_\phi &\equiv r\sin\theta(-\sin\phi\,\gamma_1 + \cos\phi\,\gamma_2) \end{aligned} \tag{2}$$

and the dual-frame vectors are denoted as $\{e^t, e^r, e^\theta, e^\phi\}$. We will also frequently employ the unit vectors $\hat{\theta}$ and $\hat{\phi}$ defined by

$$\begin{aligned} \hat{\theta} &\equiv e_\theta/r && (3)\\ \hat{\phi} &\equiv e_\phi/(r\sin\theta). && (4)\end{aligned}$$

Rotational symmetry is imposed on $\overline{h}(a)$ as follows. If B is a constant spatial bivector ($e_t\cdot B = 0$) and

$$\begin{aligned} R &= e^{B/2} && (5)\\ x' &= \tilde{R}xR, && (6)\end{aligned}$$

then rotating $\overline{h}(a)$ and translating it to the back-rotated position x' should leave $\overline{h}(a)$ unchanged. Hence rotational symmetry requires that

$$R\overline{h}_{x'}(\tilde{R}aR)\tilde{R} = \overline{h}(a). \tag{7}$$

This places a strong restriction on $\overline{h}(a)$, and the most general form that $\overline{h}(a)$ can take is given by

$$\left.\begin{aligned} \overline{h}(e^t) &= f_1 e^t + f_2 e^r\\ \overline{h}(e^r) &= g_1 e^r + g_2 e^t\\ \overline{h}(e^\theta) &= \alpha e^\theta\\ \overline{h}(e^\phi) &= \alpha e^\phi, \end{aligned}\right. \tag{8}$$

where f_1, f_2, g_1, g_2 and α are all functions of t and r only. Rotational symmetry alone does not fix all the gauge freedom in the $\overline{h}$-function because we are free to reparameterise t and r without altering the functional form of (8). Once we have identified the intrinsic quantities, however, a natural gauge choice will emerge in which $\overline{h}(a)$ is expressed solely in terms of physically-determined variables.

Conventional GR would now proceed by constructing the line element from (8). This has the general form

$$ds^2 = g_{00}(t,r)\,dt^2 + g_{01}(t,r)\,dt\,dr + g_{11}(t,r)\,dr^2 + g_{22}(t,r)(d\theta^2 + \sin^2\theta\,d\phi^2). \tag{9}$$

From (9) the connection coefficients are then found, and the Riemann and Einstein tensors computed. This procedure results in a set of complicated non-linear second-order equations which are difficult to solve. Moreover, the gauge-transformation properties of the quantities involved are hard to recover. Here we follow a different and, to our knowledge, entirely new approach. We use equation (27) to find what terms must be present in $\omega(a)$ for the $\overline{h}$-field specified by (8), but leave the individual coefficients as arbitrary functions. The only terms from $\overline{h}(a)$ which are kept in $\omega(a)$ are those which are undifferentiated. These arise from frame derivatives and the algebra is usually simplified if these terms are left explicit in $\omega(a)$. Applying this procedure to (8) we are led to the following form for $\omega(a)$:

$$\begin{aligned}\omega(e_t) &= G e_r e_t \\ \omega(e_r) &= F e_r e_t \\ \omega(\hat{\theta}) &= S\hat{\theta} e_t + (T - \alpha/r) e_r \hat{\theta} \\ \omega(\hat{\phi}) &= S\hat{\phi} e_t + (T - \alpha/r) e_r \hat{\phi},\end{aligned} \tag{10}$$

where G, F, S and T are functions of t and r only.

The bracket relations obtained from (10) are as follows:

$$\begin{aligned}[L_t, L_r] &= G L_t - F L_r & \qquad [L_r, L_{\hat{\theta}}] &= -T L_{\hat{\theta}} \\ [L_t, L_{\hat{\theta}}] &= -S L_{\hat{\theta}} & [L_r, L_{\hat{\phi}}] &= -T L_{\hat{\phi}} \\ [L_t, L_{\hat{\phi}}] &= -S L_{\hat{\phi}} & [L_\theta, L_\phi] &= 0.\end{aligned} \tag{11}$$

Here L_t, L_r, L_θ and L_ϕ are abbreviations for the associated coordinate frame derivatives, so $L_\theta = e_\theta \cdot \overline{h}(\nabla)$ *etc.* For the unit vectors we of course have $L_{\hat{\theta}} = \hat{\theta} \cdot \overline{h}(\nabla)$ and $L_{\hat{\phi}} = \hat{\phi} \cdot \overline{h}(\nabla)$. Use of the unit vectors $\hat{\theta}$ and $\hat{\phi}$ eliminates the need to calculate some coordinate derivatives which cancel out of the final results.

We now apply equations (7) and (17) to compute the Riemann tensor and find that

$$\begin{aligned}\mathcal{R}(e_r e_t) &= (L_r G - L_t F + G^2 - F^2) e_r e_t \\ \mathcal{R}(\hat{\theta} e_t) &= (-L_t S + GT - S^2)\hat{\theta} e_t + (-L_t T + SG - ST) e_r \hat{\theta} \\ \mathcal{R}(\hat{\phi} e_t) &= (-L_t S + GT - S^2)\hat{\phi} e_t + (-L_t T + SG - ST) e_r \hat{\phi} \\ \mathcal{R}(e_r \hat{\theta}) &= (L_r T - FS + T^2) e_r \hat{\theta} + (L_r S - FT + ST)\hat{\theta} e_t \\ \mathcal{R}(e_r \hat{\phi}) &= (L_r T - FS + T^2) e_r \hat{\phi} + (L_r S - FT + ST)\hat{\phi} e_t \\ \mathcal{R}(\hat{\theta}\hat{\phi}) &= (-S^2 + T^2 - (\alpha/r)^2)\hat{\theta}\hat{\phi}.\end{aligned} \tag{12}$$

In arriving at (12) the bracket structure is used to simplify the derivatives of α/r. Since $\alpha/r = L_{\hat{\theta}}\theta$, we have

$$L_t(\frac{\alpha}{r}) = L_t L_{\hat{\theta}}\theta = [L_t, L_{\hat{\theta}}]\theta = -S\frac{\alpha}{r} \tag{13}$$

and

$$L_r(\frac{\alpha}{r}) = L_r L_{\hat{\theta}}\theta = [L_r, L_{\hat{\theta}}]\theta = -T\frac{\alpha}{r}. \tag{14}$$

To proceed, we must now decide what form the matter stress-energy tensor should take. We assume that the matter is modelled by an ideal fluid so we can write

$$\mathcal{T}(a) = (\rho + p)a \cdot v v - pa, \tag{15}$$

where ρ is the energy density, p is the pressure and v is the covariant fluid velocity ($v^2 = 1$). Radial symmetry means that v can only be in the e_t and e_r directions, so v must be of the form $(\cosh u \, e_t + \sinh u \, e_r)$. But, in the $\overline{h}$-function defined by equation (8), we are free to perform an arbitrary radial boost as this does not change the functional form of $\overline{h}(a)$. It follows that we are free to apply any radial boost to v. We use this freedom to set $v = e_t$ so that the matter stress-energy tensor becomes

$$\mathcal{T}(a) = (\rho + p)a \cdot e_t e_t - pa. \tag{16}$$

Contracting (12) to form the Ricci and Einstein tenors and equating with (16) we obtain the following set of equations:

$$\begin{aligned}
L_t S &= 2A + GT - S^2 - 4\pi p && (17)\\
L_t T &= S(G - T) && (18)\\
L_r S &= T(F - S) && (19)\\
L_r T &= -2A + FS - T^2 - 4\pi\rho && (20)\\
L_r G - L_t F &= F^2 - G^2 + 4A + 4\pi(\rho + p) && (21)
\end{aligned}$$

where

$$A \equiv \tfrac{1}{4}(-S^2 + T^2 - (\alpha/r)^2). \tag{22}$$

Feeding these relations back into (12) we find that the Riemann tensor can be written concisely as

$$\mathcal{R}(B) = (A + \frac{2\pi}{3}\rho)(B + 3e_r e_t B e_r e_t) + 4\pi[(\rho + p)B \cdot e_t e_t - \frac{2}{3}\rho B]. \tag{23}$$

The first term in this expression for the Riemann tensor is the Weyl tensor, and the second is the term that contracts to give the Ricci tensor. This latter contribution to $\mathcal{R}(B)$ does not appear to have been given a name in the literature. It is perhaps best thought of as the 'source' term.

The remaining equations are contained in the Bianchi identities. It turns out that, in this case, the full Bianchi identities are satisfied using only the above equations and the contracted identity (21). The contracted Bianchi identity produces the pair of equations

$$\mathcal{D}\cdot(\rho e_t) + p\mathcal{D}\cdot e_t = 0 \tag{24}$$

$$(\rho + p)(e_t\cdot\mathcal{D}e_t)\wedge e_t - (\mathcal{D}p)\wedge e_t = 0, \tag{25}$$

which reduce to

$$L_t\rho = -(F + 2S)(\rho + p) \tag{26}$$

$$L_r p = -G(\rho + p). \tag{27}$$

Equations (13), (14), (17)–(21), (26) and (27) and the bracket relation

$$[L_t, L_r] = GL_t - FL_r \tag{28}$$

form the complete set of intrinsic equations. The structure is closed, in that it is easily verified that the bracket relation (28) is consistent with the known derivatives. There is a remarkable (pseudo)duality between these equations, in that the structure is almost completely unchanged under the simultaneous interchange $F \leftrightarrow G$, $S \leftrightarrow T$, $\rho \leftrightarrow p$ and $L_t \leftrightarrow L_r$. The duality is not exact because of the α/r term (though this term can be viewed as an artefact of the r coordinate) and the matter terms. Elsewhere [9] we exhibit a similar duality in static rotating systems where, for the matter-free case, the duality is exact.

To simplify this structure we start by forming the derivatives of A. From equations (13), (14) and (17)–(21) we find that

$$L_tA + 3SA = 2\pi Sp \tag{29}$$

$$L_rA + 3TA = -2\pi T\rho. \tag{30}$$

The above results for the derivatives of A, and equations (18) and (19), suggest that we should look for an integrating factor for the $L_t + S$ and $L_r + T$ operators. Such a function, X say, should have the properties that

$$L_tX = SX \tag{31}$$

$$L_rX = TX. \tag{32}$$

A function with these properties can only be found if the derivatives are consistent with the bracket relation (28). We check this by forming

$$\begin{aligned}[L_t, L_r]X &= L_t(TX) - L_r(SX) \\ &= X(L_tT - L_rS) \\ &= X(SG - FT) \\ &= GL_tX - FL_rX, \end{aligned} \tag{33}$$

so the properties of X are consistent with (28) and we can assume the existence of such a function. Once a particular form of $\overline{h}(a)$ has been chosen X can be found from S and T by partial integration and self-consistency of this procedure is guaranteed by (33).

Now that X is available to us we can simplify many of the above equations. We start by defining

$$M \equiv -2X^3A \tag{34}$$

so that

$$\begin{aligned} L_tM &= -4\pi SX^3p \\ L_rM &= 4\pi TX^3\rho. \end{aligned} \tag{35}$$

We next find that

$$L_t(X\frac{\alpha}{r}) = 0, \qquad L_r(X\frac{\alpha}{r}) = 0 \tag{36}$$

from which it follows that

$$X\frac{\alpha}{r} = c, \tag{37}$$

where c is a constant. Equations (31) and (32) only define X up to a constant multiple, however, so we can absorb the constant c into X and write

$$\frac{\alpha}{r} = \frac{1}{X}. \tag{38}$$

It is now clear that X plays the role of an intrinsic distance variable. A natural gauge choice is therefore to set the distance scale so that

$$r = X \tag{39}$$

which implies that

$$\alpha = 1. \tag{40}$$

In making this gauge choice we ensure that the distance scale of the background vector space matches the scale defined physically by the gravitational fields.

From the form of $\overline{h}(a)$ (8) and equations (31) and (32) we now see that

$$g_1 = L_r r = Tr \tag{41}$$

$$g_2 = L_t r = Sr \tag{42}$$

and it follows that M is given by

$$M = \frac{r}{2}({g_2}^2 - {g_1}^2 + 1). \tag{43}$$

In addition, we now find that G and F satisfy

$$L_t g_1 = G g_2 \tag{44}$$

$$L_r g_2 = F g_1. \tag{45}$$

The 'Newtonian' Gauge

To complete the solution we need to make a gauge choice for the t coordinate. Unlike the distance variable X, no natural time coordinate has yet emerged. We must therefore look for some additional criteria to motivate a choice of time coordinate. Returning to the derivatives of M (35) we find that these formulae can be inverted to yield

$$\frac{\partial M}{\partial t} = \frac{-4\pi g_1 g_2 r^2(\rho + p)}{f_1 g_1 - f_2 g_2} \tag{46}$$

$$\frac{\partial M}{\partial r} = \frac{4\pi r^2(f_1 g_1 \rho + f_2 g_2 p)}{f_1 g_1 - f_2 g_2}. \tag{47}$$

The second equation reduces to a simple classical relation if we choose $f_2 = 0$ as we then obtain

$$\partial_r M = 4\pi r^2 \rho, \tag{48}$$

which says that $M(r,t)$ is determined by the amount of mass-energy in a sphere of radius r. There are other reasons for choosing the time variable such that $f_2 = 0$. For example, we can then use the bracket structure to solve for f_1. With $f_2 = 0$ we have

$$L_t = f_1 \partial_t + g_2 \partial_r \tag{49}$$

$$L_r = g_1 \partial_r \tag{50}$$

and the bracket relation (28) implies that

$$\begin{aligned} L_r f_1 &= -G f_1 \\ \Longrightarrow \quad \partial_r f_1 &= \frac{G}{g_1} f_1 \\ \Longrightarrow \quad f_1 &= \epsilon(t) \exp\{-\int^r \frac{G}{g_1} dr\}. \end{aligned} \tag{51}$$

The function $\epsilon(t)$ can be absorbed by a further rescaling of t, so we are left with the simple result that

$$f_1 = \exp\{-\int^r \frac{G}{g_1} dr\}. \tag{52}$$

Another reason why $f_2 = 0$ is a natural gauge choice is seen when the pressure is zero. In this case equation (27) forces G to be zero, and equation (52) then sets $f_1 = 1$. A free-falling particle with $v = e_t$ (*i.e.* comoving with the fluid) then has

$$\dot{t} e_t + \dot{r} e_r = e_t + g_2 e_r, \tag{53}$$

where the dots denote differentiation with respect to the affine parameter. Since $\dot{t} = 1$ the time coordinate t matches the proper time of all observers comoving with the fluid. So, in the absence of pressure, we are able to recover a global 'Newtonian' time on which all observers can agree (provided all clocks are correlated initially, which is not hard to do). Furthermore, it is also clear from (53) that g_2 is the velocity of the particle. Hence equation (46), which reduces to

$$\frac{\partial M}{\partial t} = -4\pi g_2 r^2 \rho \tag{54}$$

in the absence of pressure, has a simple Newtonian interpretation — it equates the work with the rate of flow of energy density. The equation for M (43) in the form

$$\tfrac{1}{2} {g_2}^2 - \frac{M}{r} = \tfrac{1}{2}({g_1}^2 - 1) \tag{55}$$

is also now familiar from Newtonian physics — it is a Bernoulli equation for zero pressure and total (non-relativistic) energy $({g_1}^2 - 1)/2$. For these various reasons we refer to $f_2 = 0$ as defining the 'Newtonian' gauge. This is the gauge choice we employ for many of the applications to follow.

The equations of 'intrinsic fluid dynamics' in the Newtonian gauge are summarised in Table 1.

The $\overline{h}$-function

$$\begin{aligned}\overline{h}(e^t) &= f_1 e^t\\ \overline{h}(e^r) &= g_1 e^r + g_2 e^t\\ \overline{h}(e^\theta) &= e^\theta\\ \overline{h}(e^\phi) &= e^\phi\end{aligned}$$

The ω-function

$$\begin{aligned}\omega(e_t) &= G e_r e_t\\ \omega(e_r) &= F e_r e_t\\ \omega(\hat{\theta}) &= g_2/r\,\hat{\theta} e_t + (g_1-1)/r\, e_r\hat{\theta}\\ \omega(\hat{\phi}) &= g_2/r\,\hat{\phi} e_t + (g_1-1)/r\, e_r\hat{\phi}\end{aligned}$$

Directional derivatives

$$\begin{aligned}L_t &= f_1\partial_t + g_2\partial_r\\ L_r &= g_1\partial_r\end{aligned}$$

Equations relating the $\overline{h}$- and ω-functions

$$\begin{gathered}L_t g_1 = G g_2\\ L_r g_2 = F g_1\\ f_1 = \exp\{\textstyle\int^r -G/g_1\,dr\}\end{gathered}$$

Definition of M

$$M \equiv \tfrac{1}{2} r({g_2}^2 - {g_1}^2 + 1)$$

Remaining derivatives

$$\begin{aligned}L_t g_2 &= G g_1 - M/r^2 - 4\pi r p\\ L_r g_1 &= F g_2 + M/r^2 - 4\pi r\rho\end{aligned}$$

Matter derivatives

$$\begin{aligned}L_t M &= -4\pi g_2 r^2 p \qquad & L_t\rho &= -(2g_2/r + F)(\rho+p)\\ L_r M &= 4\pi g_1 r^2\rho & L_r p &= -G(\rho+p)\end{aligned}$$

Riemann tensor

$$\begin{aligned}\mathcal{R}(B) \;=\; & 4\pi[(\rho+p)B\cdot e_t e_t - 2\rho/3\, B]\\ & -\tfrac{1}{2}(M/r^3 - 4\pi\rho/3)(B + 3e_r e_t B e_r e_t)\end{aligned}$$

Fluid stress-energy tensor

$$\mathcal{T}(a) = (\rho+p)a\cdot e_t e_t - pa$$

Table 1: *Equations governing a radially-symmetric perfect fluid.*

6 Applications

In the following sections we develop a number of applications of the equations summarised in Table 1. We start by treating the well-known case of a static, radially-symmetric star.

6.1 Static Matter Distributions

For a static matter distribution, ρ and p are functions of r only. It follows from the derivatives of M that

$$M(r) = \int_0^r 4\pi r'^2 \rho(r')\, dr' \tag{1}$$

and

$$L_t M = g_2 4\pi r^2 \rho = -g_2 4\pi r^2 p. \tag{2}$$

But, for any physical matter distribution, ρ and p must both be positive, in which case equation (2) can only be satisfied if

$$\begin{aligned} g_2 &= 0 \quad (3) \\ \implies F &= 0. \quad (4) \end{aligned}$$

Since $g_2 = 0$ we see that g_1 is given simply in terms of M by

$${g_1}^2 = 1 - 2M/r, \tag{5}$$

which recovers contact with the standard line element for a static, radially-symmetric field.

The remaining equation of use is that for $L_t g_2$ which now gives

$$G g_1 = M/r^2 + 4\pi r p. \tag{6}$$

Equations (5) and (6) now combine with that for $L_r p$ to yield the famous Oppenheimer-Volkov equation

$$\frac{\partial p}{\partial r} = -\frac{(\rho + p)(M(r) + 4\pi r^3 p)}{r(r - 2M(r))}. \tag{7}$$

By this point we have successfully recovered all the usual equations governing a non-rotating star. The work involved in recovering these equations from the full time-dependent case is minimal, and the final form of $\overline{h}(a)$ is very simple (it is a diagonal function). What's more, the meaning of the t and r coordinates is clear, since they have been defined operationally. The final equations found here do not differ from those of GR, however, so no observable differences can be expected for the case of a stationary star.

6.2 Point-Source Solutions — Black Holes

The next solution of interest is obtained when the matter is concentrated at a single point ($r = 0$). For such a solution, $\rho = p = 0$ everywhere away from the source, and the matter equations reduce to

$$\left.\begin{array}{l} L_t M = 0 \\ L_r M = 0 \end{array}\right\} \implies M = \text{constant}. \tag{8}$$

Maintaining the symbol M for this constant we now find that the equations reduce to

$$\begin{aligned} L_t g_1 &= G g_2 && (9) \\ L_r g_2 &= F g_1 && (10) \end{aligned}$$

and

$$g_1{}^2 - g_2{}^2 = 1 - 2M/r. \tag{11}$$

There are no further equations which yield new information. In the vacuum, therefore, we have an under-determined system of equations and some additional gauge-fixing is needed to choose an explicit form of $\overline{h}(a)$. The reason for this is that in the vacuum region the Riemann tensor reduces to

$$\mathcal{R}(B) = -\frac{M}{2r^3}(B + 3e_r e_t B e_r e_t). \tag{12}$$

This tensor is now *invariant* under boosts in the $e_r e_t$ plane. In the presence of matter the source contribution to $\mathcal{R}(B)$ breaks this invariance since the e_t vector is transformed to a different vector under a radial boost. This new symmetry of $\mathcal{R}(B)$ appears as soon as ρ and p vanish and manifests itself as a new freedom in the choice of $\overline{h}$-function.

Given that this new freedom exists, we should look for a choice of g_1 and g_2 which simplifies the physics as far as possible. If we attempt to reproduce the Schwarzschild solution we would set $g_2 = 0$, but then we immediately run into difficulties with g_1, which is not defined for $r < 2M$. We must therefore look for an alternative gauge choice. We will see in the following section that for collapsing dust g_1 controls the energy of infalling matter at $r = \infty$. A sensible gauge choice is therefore to set

$$g_1 = 1 \tag{13}$$

so that

$$
\begin{aligned}
g_2 &= -\sqrt{2M/r} && (14)\\
G &= 0 && (15)\\
F &= -M/(g_2 r^2) && (16)\\
f_1 &= 1. && (17)
\end{aligned}
$$

In this gauge the $\overline{h}$-function takes the remarkably simple form

$$\overline{h}(a) = a - \sqrt{2M/r}\, a \cdot e_r e_t, \tag{18}$$

which only differs from the identity through a single term. The geodesic equation for a radially-infalling particle (with unit mass) reduces to

$$
\begin{aligned}
\dot{r}^2 &= 2M/r + \sinh^2 u_0 && (19)\\
(1 - 2M/r)\dot{t} &= \cosh u_0 + \dot{r}\sqrt{2M/r}. && (20)
\end{aligned}
$$

The $\dot{r}$ equation shows immediately that $\ddot{r} = -M/r^2$ and the constant $\sinh^2 u_0$ can be identified with twice the initial kinetic energy of the infalling particle. At the horizon ($r = 2M$) $\dot{r} = -\cosh u_0$, so there is no pole in $\dot{t}$ (20). All particles cross the horizon and reach the singularity in a finite coordinate time. Some possible trajectories are illustrated in Figure 6.2.

In the case where the particle is dropped from rest at $r = \infty$ equations (19) and (20) reduce to

$$\dot{r} = -\sqrt{2M/r}, \quad \dot{t} = 1, \tag{21}$$

and we recover an entirely Newtonian description of the motion. The properties of a black hole are so simple in the gauge defined by (18) that it is astonishing that it is almost never seen in the literature. This is presumably because the line element associated with (18) does not look as natural as the $\overline{h}$-function itself, and hides the underlying simplicity of the system. In the Newtonian gauge one hardly needs to modify classical reasoning at all to understand the processes involved — all particles just cross the horizon and fall into the singularity in a finite coordinate time. And the horizon is located at $r = 2M$ precisely because we can apply Newtonian arguments! The only departures from Newtonian physics lie in relativistic corrections to the proper-time taken for infall, and in modifications to the equations for angular motion which lead to the familiar results for orbital

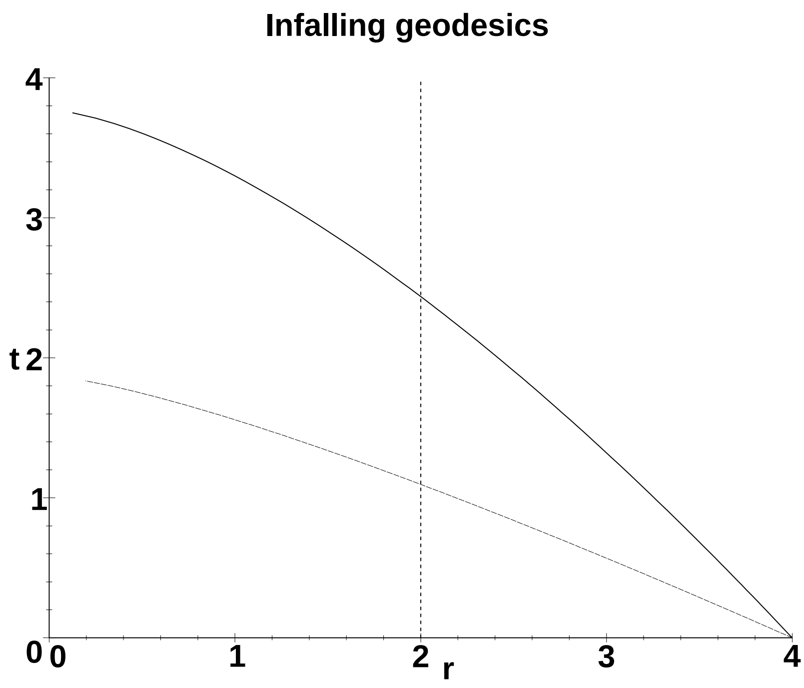

Figure 1: Possible particle trajectories for radially-infalling particles in the Newtonian gauge. The upper curve is for a particle released with $\dot{r} = -\sqrt{2M/r}$ at $r = 4M$, while the lower curve is for a photon. The vertical dotted line indicates the horizon.

precession.

Horizons and Birkhoff's Theorem

The bulk of the literature on black holes works with the Schwarzschild solution, which is obtained by setting $g_2 = 0$. In this case

$$g_1 = \sqrt{1 - 2M/r} \tag{22}$$

which is only defined for $r > 2M$. GR starts to invoke coordinate transformations at this point, usually resulting in the advanced or retarded Eddington-Finkelstein form of the solution. In our gauge theory, however, the picture is different; $g_2 = 0$ is simply *not* an allowed solution since it does not result in a globally-defined $\overline{h}$-function. It is over such questions regarding the global nature of fields that differences between our 'flatspace' gauge theory and GR start to emerge. Here the differences manifest themselves quite clearly. We are forced to a solution in which $g_2 \neq 0$, so $\overline{h}(a)$ is not diagonal and hence not time-reverse symmetric. Time reversal is achieved by combining the translation

$$f(x) = -e_t x e_t = x' \tag{23}$$

with the reflection

$$\overline{h}'(a) = -e_t \overline{h}(a) e_t. \tag{24}$$

Hence the time-reversed solution is given by

$$\overline{h}^*(a) = e_t \overline{h}_{x'}(e_t a e_t) e_t. \tag{25}$$

Applying this to the solution (18) we obtain

$$\overline{h}^*(a) = a + \sqrt{2M/r}\, a{\cdot}e_r e_t, \tag{26}$$

which has changed the sign of the off-diagonal term. The result is a solution in which particles inside the horizon are swept out. Once outside, the force is still attractive but particles cannot re-enter back through the horizon. Despite the fact that the Riemann tensor is unchanged by time-reversal, it is impossible to find an $\overline{h}$-function that is also time-reverse symmetric. In our theory the black hole has more memory about its formation than simply its mass M. It also remembers that it was formed in a particular time direction. The appearance of a horizon is thus associated with the onset time asymmetry, which is very satisfying from a physical viewpoint.

Another form of the Schwarzschild solution that runs into difficulties in our theory is that obtained by using Kruskal coordinates. These coordinates introduce a 'double-cover' of spacetime so that each value of r determines two distinct hypersurfaces. This is clearly not possible in our theory without some radical redefinition of how r is viewed as a function of spacetime position x. In GR the Kruskal-Szerkes solution is the maximal continuation of the Schwarzschild metric and is viewed as giving the complete description of a radially-symmetric black hole [23]. The fact that it is ruled out of our gauge theory means that our allowed solutions are not 'maximal' and forces us to address the issue of geodesic incompleteness. For the solution (18) geodesics exist which cannot be extended into the past for all values of their affine parameter. But, if we adopt the view that the black hole must have formed in the past from some collapse process, then there must have been a time before which the horizon did not exist. Any geodesic must therefore have come from a region in the past where no horizon was present, so there is no question of the geodesics being incomplete. A model of such a collapse process is discussed in the following section. We therefore arrive at a consistent picture in which the formation of a horizon retains information about the direction of time in which collapse occurred, and all geodesics from the past emanate from a period before the horizon formed. This picture is very different from GR, which is happy to deal with eternal, time-symmetric black holes. This shift to a picture with a fixed time direction is typical of the transition from a second-order theory to a first-order one [8].

It is finally worth commenting on how the above affects Birkhoff's theorem. One form of Birkhoff's theorem is that the gravitational fields outside any radially-symmetric distribution of matter are necessarily static. This is seen immediately from equation (8) which shows that the Riemann tensor is a function of r only. However, Birkhoff's theorem is frequently used to argue that the line element outside a radially-symmetric body can always be brought to the form [24]

$$ds^2 = (1 - 2M/r)\, dt^2 - (1 - 2M/r)^{-1}\, dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{27}$$

As we have seen, this ceases to be correct in our theory if a horizon is present. In this case $\bar{h}(a)$ is independent of t, but cannot be time-reverse symmetric.

6.3 Collapsing Dust

The simplest model for collapsing matter is one in which the pressure is set to zero so that the situation describes collapsing dust. In this case $G = 0$, $f_1 = 1$ and t is the time measured by freely-falling observers (from ∞). The equations of Table 1 reduce to

$$\begin{aligned} F &= \partial_r g_2 && (28)\\ M(r,t) &= \int_0^r 4\pi r'^2 \rho(r',t)\, dr', && (29)\end{aligned}$$

which define F and M on a time slice, together with the update equations

$$\begin{aligned} \partial_t g_2 + g_2 \partial_r g_2 &= -M/r^2 && (30)\\ \partial_t M + g_2 \partial_r M &= 0. && (31)\end{aligned}$$

As g_2 is the velocity of a particle comoving with the fluid, equations (30) and (31) provide a completely Newtonian description of the fluid. Equation (30) is the Euler equation with an inverse-square gravitational force, and (31) is the equation for conservation of mass. It is also worth noting that the L_t derivative now plays the role of the 'matter' or 'comoving' derivative for the fluid.

Equations (30) and (31) enable ρ and M to be propagated from one time-slice to the next so, given suitable initial conditions, one can propagate into the future. The system of equations is in fact soluble analytically [9], but here we just consider a numerical simulation. From an initial density profile $\rho(r, t_0)$ and velocity profile $g_2(r, t_0)$, equation (29) is used to obtain $M(r, t_0)$. Equations (30) and (31) are then used to update M and g_2. On any given time-slice, ρ and F can be recovered using equations (28) and (29), and g_1 can be recovered from g_2 and M. The results of such a simulation are displayed in Figure 6.3.

Any particle on a radial path has a covariant velocity of the form

$$v = \cosh u \, e_t + \sinh u \, e_r. \tag{32}$$

The underlying trajectory has $\dot{x} = \underline{h}(v)$, so the radial motion is determined by

$$\dot{r} = \cosh u \, g_2 + \sinh u \, g_1. \tag{33}$$

Since g_2 is negative for collapsing matter, the particle can only achieve an outward velocity if ${g_1}^2 > {g_2}^2$. A horizon therefore forms at the point where

$$2M(r,t)/r = 1. \tag{34}$$

The appearance of this horizon is illustrated in Figure 6.3. It is conventional to extend the horizon back in time along the past light-cone to the origin ($r = 0$), since any particle inside this surface could not have reached the point where $2M/r - 1$ first drops to zero, and hence is also trapped [25].

If pressure is included the pressure gradient causes the internal clock carried by the fluid to run at a different rate from t. The overall picture is similar, however. The horizon forms in a finite external coordinate time and matter has no difficulty crossing the horizon and falling onto the central singularity. The resulting solution is 'maximal' and the end result is a solution of the type discussed in the preceding Section. Again, the 'Newtonian' gauge makes the physics so simple that it surely deserves a prominent place in the description of black hole physics.

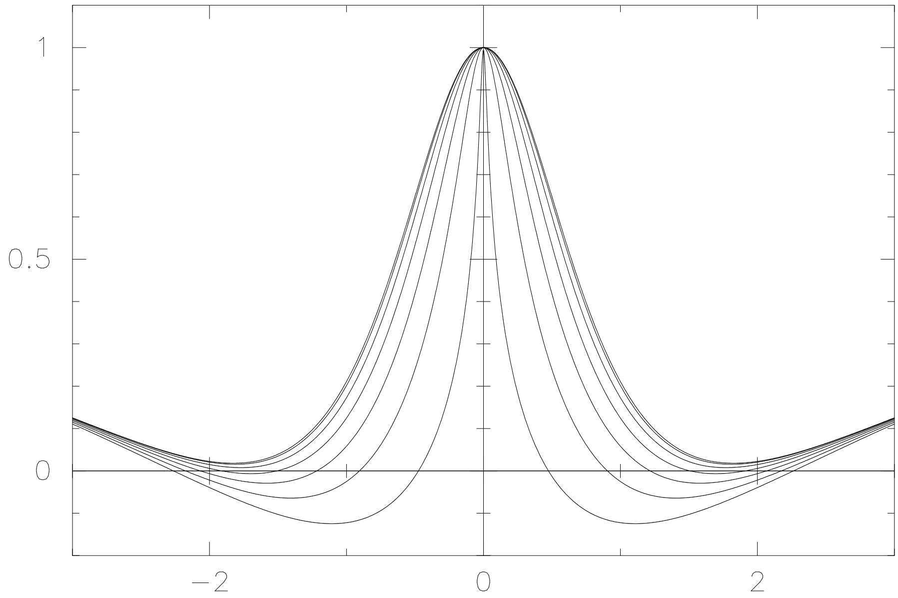

Figure 2: Simulation of collapsing dust in the Newtonian gauge. Successive time slices for the horizon function $(1 - 2M(r,t)/r)$ *versus* radius are shown, with the top curve corresponding to $t = 0$ and lower curves to successively later times. The initial density profile is of the form $\rho = \rho_0/(1 + (r/r_0)^2)^2$, and the initial velocity is everywhere zero. There is no horizon initially, but a trapped region quickly forms, since in regions where $1 - 2M/r < 0$, photons can only move inwards.

6.4 Electromagnetism in a Black-Hole Background

Maxwell's equations in a gravitational background can be written in a number of different forms [9]. From the vector potential A we define the field strength $F \equiv \nabla \wedge A$. The covariant field strength is obtained from F by defining

$$\mathcal{F} \equiv \overline{h}(\nabla\wedge A). \tag{35}$$

In terms of $\mathcal{F}$, Maxwell's equations (in the absence of torsion) become [9]

$$\mathcal{D}\mathcal{F} = \mathcal{J}. \tag{36}$$

This first-order form of Maxwell's equations is very useful for problems related to the propagation of electromagnetic waves, for example. An equivalent form of the equations is obtained by defining

$$G \equiv \underline{h}\overline{h}(F) \det \underline{h}^{-1} \tag{37}$$

and

$$J = \underline{h}(\mathcal{J}) \det \underline{h}, \tag{38}$$

so that equation (36) decomposes into the separate equations

$$\nabla\wedge F = 0, \qquad \nabla\cdot G = J. \tag{39}$$

Finally, performing a spacetime split with respect to the γ_0 direction, as at equation (15), and defining

$$\begin{aligned} F &= \boldsymbol{E} + i\boldsymbol{B} && (40)\\ G &= \boldsymbol{D} + i\boldsymbol{H} && (41)\\ J\gamma_0 &= \rho + \boldsymbol{J}, && (42) \end{aligned}$$

equations (39) take the conventional form

$$\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{B} &= 0 & \qquad \boldsymbol{\nabla}\cdot\boldsymbol{D} &= \rho \\ \boldsymbol{\nabla}\times\boldsymbol{E} &= -\frac{\partial\boldsymbol{B}}{\partial t} & \qquad \boldsymbol{\nabla}\times\boldsymbol{H} &= \boldsymbol{J} + \frac{\partial\boldsymbol{D}}{\partial t} \end{aligned} \tag{43}$$

where, for spatial vectors, $\boldsymbol{a} \times \boldsymbol{b} \equiv -i\boldsymbol{a} \wedge \boldsymbol{b}$. The form (43) offers some insight into how the gravitational field interacts with the electromagnetic field. The tensor $\det \underline{h}^{-1}\underline{h}\overline{h}$ is a generalized permittivity/permeability tensor and defines the properties of the space through which the electromagnetic field propagates. For example, the bending of light by the sun can be easily understood in terms of the properties of the dielectric defined by the $\overline{h}$-field exterior to it.

The problem of interest here is to find the fields around a point source at rest outside the horizon of a radially-symmetric black hole. The $\overline{h}$-function in this case can be taken as that of equation (18). The solution to this problem can be found by adapting the work of Copson [26] and Linet [27] to the present gauge choices. Assuming that the charge is placed at a distance $a > 2M$ along the z-axis, the vector potential can be written in terms of a single scalar potential $V(r,\theta)$ as

$$A = V(r,\theta)(e_t + \frac{\sqrt{2Mr}}{r-2M}e_r), \tag{44}$$

so that

$$\mathcal{F} = -\frac{\partial V}{\partial r}e_r e_t - \frac{1}{r-2M}\frac{\partial V}{\partial \theta}\,\hat{\theta}(e_t + \sqrt{2M/r}e_r) \tag{45}$$

and

$$G = \boldsymbol{D} = -\frac{\partial V}{\partial r}e_r e_t - \frac{1}{r-2M}\frac{\partial V}{\partial \theta}\hat{\theta}e_t. \tag{46}$$

The Maxwell equations now reduce to the single partial differential equation

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{r(r-2M)}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial \theta}\right) = -\rho \tag{47}$$

where $\rho = q\delta(\boldsymbol{x} - \boldsymbol{a})$ is a δ-function at $z = a$. This was the problem originally tackled by Copson [26] who obtained a solution that was valid locally in the vicinity of the charge, but contained an additional pole at the origin. Linet [27] modified Copson's solution by removing the singularity at the origin to produce a potential $V(r,\theta)$ whose only pole is on the z-axis at $z = a$. Linet's solution is

$$V(r,\theta) = \frac{q}{ar}\frac{(r-M)(a-M) - M^2\cos^2\theta}{D} + \frac{qM}{ar}, \tag{48}$$

where

$$D = [r(r-2M) + (a-M)^2 - 2(r-M)(a-M)\cos\theta + M^2\cos^2\theta]^{1/2}. \tag{49}$$

The novel feature we wish to stress here is that once (48) is inserted back into (45) the resultant $\mathcal{F}$ is both finite and continuous at the horizon. Working in the Newtonian gauge enables us to discuss the *global* properties of the solution, rather than having to resort to the 'membrane paradigm' [28] to exclude the region $r < 2M$. One simple way to illustrate the global properties of $\mathcal{F}$ and G is to plot the streamlines of $\boldsymbol{D}$, which are conserved by equation (43). The streamlines should therefore spread out from the charge and cover all space. Furthermore, since the distance scale r was chosen to agree with the gravitationally-defined distance, the streamlines of $\boldsymbol{D}$ convey genuine intrinsic information. Hence the plots are completely unaffected by our choice for the g_1 or g_2 functions, or indeed our choice of t-coordinate. Figure 6.4 shows streamline plots for charges held at different distances above the horizon. Similar plots were first obtained by Hanni & Ruffini [29], though they were unable to extend their plots through the horizon. The physical interpretation of these plots differs somewhat from that advanced in the 'Membrane Paradigm' [28]. These differences are dealt with in detail elsewhere [9].

An interesting feature of the above solution is the existence of a repulsive 'polarisation' force [30], the effect of which is that a smaller force is needed to keep a charged particle at rest outside a black hole than an uncharged one. In their derivation of this force, Smith & Will [30] employed a complicated energy argument which involved renormalising various divergent integrals. In fact, as we show elsewhere [9], this repulsion is due entirely to the term that Linet added to Copson's formula. This is the second term in equation (48), and produces an outward-directed force on the charge of magnitude q^2M/a^3 — the same magnitude as found in [30]. This is an example of the importance of finding global solutions. The polarisation force is felt outside the horizon, yet the correction term that led to it was motivated by the properties of the field at the origin.

6.5 Cosmology

The equations of Table 1 are sufficiently general to deal with cosmology as well as astrophysics. In recent years, however, it has again become fashionable to include a cosmological constant in the field equations. The derivation of Section 4 is largely unaffected by the inclusion of the cosmological term, and only a few modifications to Table 1 are required. The full set of equations with a cosmological constant incorporated are summarised in Table 2.

In cosmology we are interested in homogeneous solutions to the equations of Table 2. Such solutions are found by setting ρ and p to functions of t only, so it

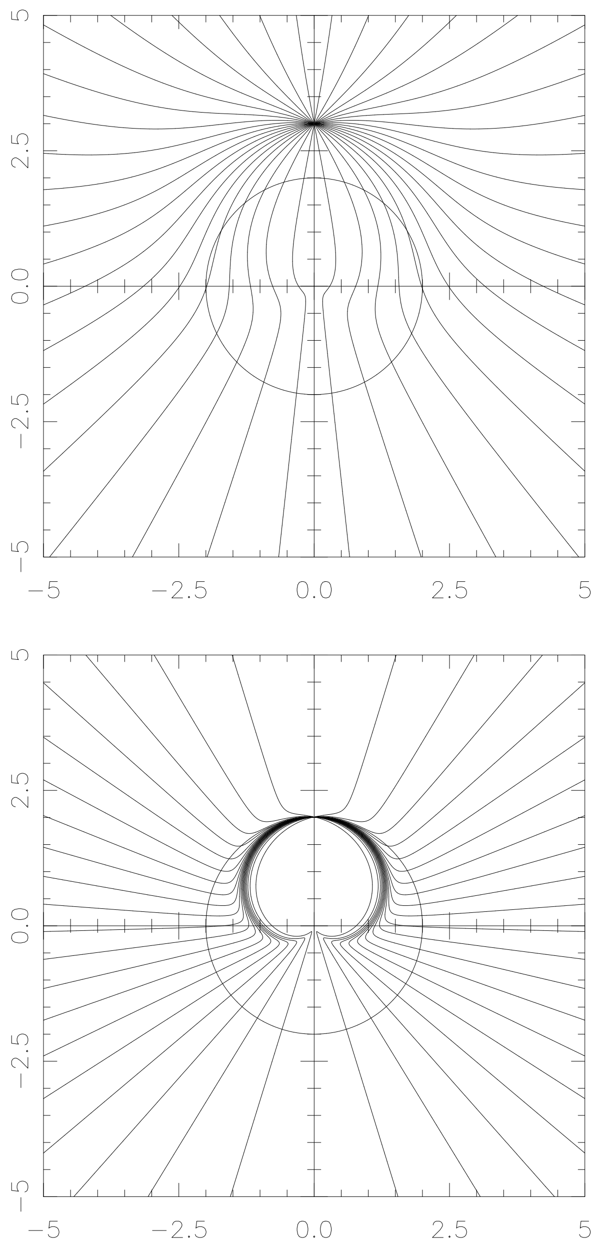

Figure 3: Streamlines of the $\boldsymbol{D}$ field. The horizon is at $r = 2$ and the charge is placed on the z-axis. In the top diagram the charge is held at $z = 3$, and in bottom diagram it is at $z = 2.1$. Note how the streamlines are attracted towards the origin but never actually meet it. Had Linet's additional term not been included some of the streamlines would have terminated on the origin, which is not permitted.

The $\overline{h}$-function	The ω-function
$\overline{h}(e^t) = f_1 e^t$ $\overline{h}(e^r) = g_1 e^r + g_2 e^t$ $\overline{h}(e^\theta) = e^\theta$ $\overline{h}(e^\phi) = e^\phi$	$\omega(e_t) = G e_r e_t$ $\omega(e_r) = F e_r e_t$ $\omega(\hat{\theta}) = g_2/r\, \hat{\theta} e_t + (g_1 - 1)/r\, e_r \hat{\theta}$ $\omega(\hat{\phi}) = g_2/r\, \hat{\phi} e_t + (g_1 - 1)/r\, e_r \hat{\phi}$
Directional derivatives	$L_t = f_1 \partial_t + g_2 \partial_r$ $L_r = g_1 \partial_r$
Equations relating the $\overline{h}$- and ω-functions	$L_t g_1 = G g_2$ $L_r g_2 = F g_1$ $f_1 = \exp\{\int^r -G/g_1\, dr\}$
Definition of M	$M \equiv \frac{1}{2} r({g_2}^2 - {g_1}^2 + 1 - \Lambda r^2/3)$
Remaining derivatives	$L_t g_2 = G g_1 - M/r^2 + r\Lambda/3 - 4\pi r p$ $L_r g_1 = F g_2 + M/r^2 - r\Lambda/3 - 4\pi r \rho$
Matter derivatives	$L_t M = -4\pi g_2 r^2 p$ $L_t \rho = -(2g_2/r + F)(\rho + p)$ $L_r M = 4\pi g_1 r^2 \rho$ $L_r p = -G(\rho + p)$
Riemann tensor	$\mathcal{R}(B) = 4\pi(\rho + p) B \cdot e_t e_t - \frac{1}{3}(8\pi\rho + \Lambda) B - \frac{1}{2}(M/r^3 - 4\pi\rho/3)(B + 3 e_r e_t B e_r e_t)$
Fluid stress-energy tensor	$\mathcal{T}(a) = (\rho + p) a \cdot e_t e_t - p a$

Table 2: *Equations governing a radially-symmetric perfect fluid — case with a non-zero cosmological constant* Λ*. The shaded equations differ from those of Table 1.*

follows immediately from the $L_r p$ equation that

$$G = 0 \quad \Longrightarrow \quad f_1 = 1. \tag{50}$$

For homogeneous fields the Weyl component of the Riemann tensor must vanish since this contains directional information through the e_r vector. The vanishing of this term requires that

$$M(r,t) = \frac{4}{3}\pi r^3 \rho, \tag{51}$$

which is consistent with the $L_r M$ equation. The $L_t M$ and $L_t \rho$ equations now reduce to

$$F = g_2/r \tag{52}$$

and

$$\dot{\rho} = -3g_2(\rho + p)/r. \tag{53}$$

But we know that $L_r g_2 = F g_1$, which can only be consistent with (52) if

$$F = c(t), \qquad g_2(r,t) = rc(t). \tag{54}$$

The $L_t g_2$ equation now reduces to a simple equation for $\dot{c}$,

$$\dot{c} + c^2 - \Lambda/3 = -\frac{4\pi}{3}(\rho + 3p). \tag{55}$$

Finally, we are left with the following pair of equations for g_1:

$$L_t g_1 = 0 \tag{56}$$

$$L_r g_1 = ({g_1}^2 - 1)/r. \tag{57}$$

The latter equation yields ${g_1}^2 = 1 + r^2\phi(t)$ and the former reduces to

$$\dot{\phi} = -2c(t)\phi. \tag{58}$$

Hence g_1 is given by

$${g_1}^2 = 1 - kr^2 \exp\{-2\int^t c(t')\, dt'\}, \tag{59}$$

where k is an arbitrary constant of integration. It is straightforward to check that (59) is consistent with the equations for $\dot{c}$ and $\dot{\rho}$. The full set of equations describing a homogeneous perfect fluid are summarised in Table 3.

At first sight, the equations of Table 3 do not resemble the usual Friedmann equations. The Friedmann equations are recovered straight-forwardly, however, by setting

$$c(t) = \frac{\dot{S}(t)}{S(t)}. \tag{60}$$

The $\overline{h}$-function	$\overline{h}(a) = a + a{\cdot}e_r[(g_1 - 1)e^r + c(t)re^t]$ ${g_1}^2 = 1 - kr^2 \exp\{-2 \int^t c(t')\, dt'\}$
The ω-function	$\omega(a) = c(t)a{\wedge}e_t - (g_1 - 1)/r\, a{\wedge}(e_r e_t)e_t$
The density	$\frac{8\pi}{3}\rho = c(t)^2 - \Lambda/3 + k \exp\{-2 \int^t c(t')\, dt'\}$
Dynamical equations	$\dot{c} + c^2 - \Lambda/3 = -4\pi/3\,(\rho + 3p)$ $\dot{\rho} = -3c(t)(\rho + p)$

Table 3: *Equations governing a homogeneous perfect fluid.*

With this substitution we find that

$$ {g_1}^2 = 1 - kr^2/S^2 \tag{61} $$

and the $\dot{c}$ and density equations become

$$ \frac{\ddot{S}}{S} - \frac{\Lambda}{3} = -\frac{4\pi}{3}(\rho + 3p) \quad \text{and} \quad \frac{\dot{S}^2 + k}{S^2} - \frac{\Lambda}{3} = \frac{8\pi}{3}\rho, \tag{62} $$

recovering the Friedmann equations in their standard form [31]. The intrinsic treatment has therefore led us to work directly with the 'Hubble velocity' $c(t)$, rather than the 'distance' scale $S(t)$. There is a good reason for this. Once the Weyl tensor is set to zero, the Riemann tensor reduces to

$$ \mathcal{R}(B) = 4\pi(\rho + p)B{\cdot}e_t e_t - \frac{1}{3}(8\pi\rho + \Lambda)B, \tag{63} $$

and we have now lost contact with an intrinsically-defined distance scale. We can therefore rescale the radius variable r with an arbitrary function of t (or r) without altering the Riemann tensor. The Hubble velocity, on the other hand, is intrinsic and it is therefore unsurprising that our treatment has led directly to equations for this.

Amongst the class of radial rescalings a particularly useful one is to rescale r to

$r' = S(t)r$. This is achieved with the transformation

$$f(x) = x\cdot e_t e_t + Sx\wedge e_t e_t \tag{64}$$

so that, on applying equation (5), the transformed $\overline{h}$-function is

$$\overline{h}'(a) = a\cdot e_t e_t + \frac{1}{S}[(1-kr^2)^{1/2} a\cdot e_r e^r + a\wedge(e_r e_t)e_t e_t]. \tag{65}$$

The function (65) reproduces the standard line element used in cosmology. We can therefore use the transformation (64) to move between the 'Newtonian' gauge developed here and the traditional 'static' gauge. The differences between these gauges can be understood by considering geodesic motion. A particle at rest with respect to the cosmological frame (defined by the cosmic microwave background) has $v = e_t$. In the standard 'static' gauge such a particle is not moving in the flatspace background (the distance variable r is equated with the comoving coordinate of GR). For this reason we refer to this gauge as being static, even though the associated line element is usually thought of as defining an expanding spacetime [19]. In the Newtonian gauge, on the other hand, comoving particles are moving outwards radially at a velocity $\dot{r} = c(t)r$, though this expansion centre is not an intrinsic feature. Of course, attempting to distinguish these pictures is a pointless exercise, since all observables must be gauge invariant. All that is of physical relevance is that, if two particles are at rest with respect to the cosmological frame (defined by the cosmic microwave background), then the light-travel time between these particles is an increasing function of time and light is redshifted as it travels between them. The value of this redshift is independent of the gauge in which it is calculated [19], and attempting to assign the redshift to an expansion of spacetime, or a change in the speed of light, or any other property of the background space is a gauge-dependent description. The fact that there are alternative gauge choices is often ignored in modern treatments of cosmology, which invariably employ the 'expanding universe' description of redshifts.

Dust Models and Horizons

As an illustration of the utility of the Newtonian gauge in cosmology, we consider horizons in a dust model ($p = 0$). Setting p to zero implies that

$$c(t) = -\frac{\dot{\rho}}{3\rho} \tag{66}$$

so

$$g_1{}^2 = 1 - kr^2\rho^{2/3} \tag{67}$$

and

$$c(t) = \left(\frac{8\pi}{3}\rho - k\rho^{2/3} + \frac{\Lambda}{3}\right)^{1/2}. \tag{68}$$

We are therefore left with a single first-order differential equation for ρ. Explicit solutions of this equations are rarely needed, as we can always parameterise time by the density $\rho(t)$.

Now consider radial null geodesics with a trajectory

$$x(\tau) = t(\tau)e_t + r(\tau)e_r. \tag{69}$$

For these trajectories

$$v = \underline{h}^{-1}(\dot{x}) = \dot{t}(e_t - \frac{rc(t)}{g_1}\, e_r) + \frac{\dot{r}}{g_1}\, e_r \tag{70}$$

and the condition that $v^2 = 0$ reduces to

$$\begin{aligned} \frac{\dot{r}}{g_1} &= \dot{t}(\frac{rc(t)}{g_1} \pm 1) && (71)\\ \Longrightarrow \quad \frac{dr}{dt} &= rc(t) \pm g_1. && (72) \end{aligned}$$

But, as $c(t)$ is positive for a cooling universe, there is a critical radius beyond which $rc(t) - g_1$ is positive and all photons from this point onwards must follow an outward trajectory. So, if we are at the origin, we cannot receive signals from beyond this critical radius and a particle horizon exists for us at

$$r_c c(t) = g_1 \tag{73}$$

$$\Longrightarrow r_c = \left(\frac{8\pi\rho}{3} + \frac{\Lambda}{3}\right)^{-1/2}. \tag{74}$$

As the density drops this radius increases, though the inclusion of the cosmological term means that r_c never exceeds $\sqrt{3/\Lambda}$.

6.6 The Dirac Equation in a Cosmological Background

As a final application of the theory outlined here we consider the Dirac equation in a cosmological background. At this point it is slightly easier to work with the $\overline{h}$-function in the more conventional diagonal form of (65). But clearly there is a problem here if k is positive, since the square root in (65) is ill-defined for $r > k^{-1/2}$. This can be viewed as an artefact of the means by which this solution maps a 3-sphere onto Euclidean space [19]. We can remove the difficulty by using a stereographic projection to achieve a globally-defined $\overline{h}$-function, though this is not necessary for the application discussed here. (The stereographic projection is useful in other applications, however, for example it shows up that in a $k > 0$ universe every charge must have an image charge. It is somewhat surprising that this bizarre feature of a $k > 0$ universe is considered acceptable.)

An advantage of the STA formalism is that there is no need to introduce matrices or column vectors to represent spinors. Instead, a spinor can be represented as an element of the 8-dimensional even subalgebra [11]. The even subalgebra is closed under left multiplication by rotors, so can be used to represent a spin-space. In the STA the action of the $\hat{\gamma}^\mu$ matrices on a column spinor $|\psi>$ is replaced by the operation $\psi \mapsto \gamma^\mu \psi \gamma_0$, and multiplication by the unit imaginary is replaced $\psi \mapsto \psi i\sigma_3$. The free-field Dirac equation can then be written

$$\nabla \psi i \sigma_3 = m \psi \gamma_0. \tag{75}$$

This equation just employs the vector derivative ∇, so the same gauging arguments as used in Section 3 can be employed to couple in the gravitational field. The only difference is that spinors transform single-sidedly under rotations, so the ω-function only enters on the left-hand side of ψ. The resulting equation is

$$\partial_a[L_a + \tfrac{1}{2}\omega(a)]\psi i\sigma_3 = m\psi\gamma_0, \tag{76}$$

which can also be derived from an action integral [9]. It is easily verified that equation (76) is covariant under local translations and rotations.

Observables are formed by double-sided application of the spinor ψ on one of the fixed $\{\gamma_\mu\}$ frame vectors. In particular, the charge current is $\mathcal{J} = \psi\gamma_0\tilde{\psi}$ and the spin current is $\mathcal{S} = \psi\gamma_3\tilde{\psi}$. Using equation (76) it is not hard to show that $\mathcal{D} \cdot \mathcal{J} = 0$, as is required by charge conservation.

Our purpose here is to find solutions to (76) in the background of a dust model for cosmology. With the $\overline{h}$-function in the form of (65) the relevant equation is

$$\Big(e^t\partial_r + \frac{1}{S}[(1-kr^2)^{1/2}e^r\partial_r + e^\theta\partial_\theta + e^\phi\partial_\phi]\Big)\psi i\sigma_3 \\ +\tfrac{1}{2}(3c(t)e_t - 2[(1-kr^2)^{1/2}-1]e_r)\psi i\sigma_3 = m\psi\gamma_0. \tag{77}$$

The question we wish to address is this: can we find solutions to (77) such that the observables are homogeneous? There is clearly no difficulty if $k=0$, since equation (77) is solved by

$$\psi = \rho^{1/2}e^{-i\sigma_3 mt} \tag{78}$$

and the observables are fixed vectors which just scale as $\rho(t)$ in magnitude. But what happens when $k \neq 0$? It turns out that the solution (78) must now be modified to

$$\psi = \frac{\rho^{1/2}}{1+\sqrt{1-kr^2}}e^{-i\sigma_3 mt}. \tag{79}$$

But this is no longer homogeneous, and the observables obtained form (79) all contain additional r-dependence as well as scaling as $\rho(t)$. In principle, therefore, one could determine the origin of this space with measurements of the current density. This clearly violates the principle of homogeneity, though it is not necessarily inconsistent with current experiments. The consequences of this fact for self-consistent solutions of the Einstein-Dirac equations are discussed in [19] (see also [32]).

So how can it be that classical phenomena do not see this 'preferred' direction in $k \neq 0$ models, but the quantum spinor ψ does? The answer lies in the gauge structure of the theory. The 'minimal-coupling' procedure couples Dirac spinors directly to the ω-function [9], which in this case is not homogeneous. Dirac spinors can therefore probe the theory directly at the level of the ω-field, whereas classical quantities only interact at the level of the covariant quantities obtained from the gravitational fields (which are homogeneous). Whilst the above result is not conclusive, it does strongly suggest that cosmological models with $k \neq 0$ are inconsistent with the assumption of homogeneity. This does not rule out spatially-flat universes with a non-zero cosmological constant.

7 Conclusions

We have seen how a theory of gravity can be developed as a gauge theory within the framework of a flat background spacetime. The resulting theory is conceptually simple. It employs neither differential geometry nor manifold theory, and the role of the gauge fields can be clearly understood. The theory is also easier to calculate with than GR, and the 'intrinsic' technique outlined here offers the hope of real progress with problems such as finding the gravitational fields inside a rotating star [9]. By studying a radially-symmetric perfect fluid we have found a new gauge which dramatically simplifies the physics, reducing it essentially to a set of classical equations. It has also emerged that, where the global nature of the solutions is important, our theory departs from GR. This is seen most clearly in the description of the horizon outside a black hole. The existence of a global solution enables us to produce fieldline plots for a charge held outside a horizon which continue smoothly across the horizon. The structure of the fieldlines inside the horizon shows up the 'dielectric' nature of the interaction between gravity and electromagnetism.

The application to cosmology has revealed the existence of an alternative gauge choice in which a number of calculations can be performed more efficiently. This new gauge suggests a different interpretation for redshifts, which reinforces the importance of only discussing intrinsic relations between fields, and not the properties of matter relative to the spacetime background. One can only do the latter if it has already been arranged that properties of the background spacetime match intrinsic features of the gravitational or matter fields. A final surprise is obtained when studying the Dirac equation in a cosmological background. The spinor field probes the gauge structure at a deeper level than classical fields and reveals that, for Dirac particles, the only cosmological models consistent with homogeneity are spatially flat.

A fuller presentation of this work is in preparation [9]. This includes the derivation of all the field equations (matter and gravitational) from action integrals. In addition, we treat quantum mechanics in a black-hole background and give an intrinsic treatment of a rotating perfect fluid. It is also argued in [9] that splitting gravitational effects into the $\overline{h}$- and ω-fields suggests a novel route to a multiparticle theory of gravitationally-interacting particles.

References

[1] R. Utiyama. Invariant theoretical interpretation of interaction. *Phys. Rev.*, 101(5):1597, 1956.

[2] T.W.B. Kibble. Lorentz invariance and the gravitational field. *J. Math. Phys.*, 2(3):212, 1961.

[3] D. Ivanenko and G. Sardanashvily. The gauge treatment of gravity. *Phys. Rep.*, 94(1):1, 1983.

[4] M. Gockeler and T. Schucker. *Differential Geometry, Gauge Theories, and Gravity.* Cambridge University Press, 1987.

[5] F.W. Hehl, P. von der Heyde, G.D. Kerlick, and J.M. Nester. General relativity with spin and torsion: Foundations and prospects. *Rev. Mod. Phys.*, 48:393, 1976.

[6] D. Hestenes and G. Sobczyk. *Clifford Algebra to Geometric Calculus.* Reidel, Dordrecht, 1984.

[7] D. Hestenes. *Space–Time Algebra.* Gordon and Breach, New York, 1966.

[8] S.F. Gull, A.N. Lasenby, and C.J.L. Doran. Electron paths, tunnelling and diffraction in the spacetime algebra. *Found. Phys.*, 23(10):1329, 1993.

[9] A.N. Lasenby, C.J.L. Doran, and S.F. Gull. Gravity, gauge theories and geometric algebra. *Phil. Trans. R. Soc. Lond. A*, 356:487–582, 1998.

[10] S.F. Gull, A.N. Lasenby, and C.J.L. Doran. Imaginary numbers are not real — the geometric algebra of spacetime. *Found. Phys.*, 23(9):1175, 1993.

[11] C.J.L. Doran, A.N. Lasenby, and S.F. Gull. States and operators in the spacetime algebra. *Found. Phys.*, 23(9):1239, 1993.

[12] A.N. Lasenby, C.J.L. Doran, and S.F. Gull. A multivector derivative approach to Lagrangian field theory. *Found. Phys.*, 23(10):1295, 1993.

[13] D. Hestenes. *New Foundations for Classical Mechanics (Second Edition).* Kluwer Academic Publishers, Dordrecht, 1999.

[14] D. Hestenes. A unified language for mathematics and physics. In J.S.R. Chisholm and A.K. Common, editors, *Clifford Algebras and their Applications in Mathematical Physics (1985)*, page 1. Reidel, Dordrecht, 1986.

[15] D. Hestenes. The design of linear algebra and geometry. *Acta Appl. Math.*, 23:65, 1991.

[16] T.G. Vold. An introduction to geometric algebra with an application to rigid body mechanics. *Am. J. Phys.*, 61(6):491, 1993.

[17] T.G. Vold. An introduction to geometric calculus and its application to electrodynamics. *Am. J. Phys.*, 61(6):505, 1993.

[18] C.J.L. Doran, A.N. Lasenby, and S.F. Gull. Gravity as a gauge theory in the spacetime algebra. In F. Brackx, R. Delanghe, and H. Serras, editors, *Clifford Algebras and their Applications in Mathematical Physics (1993)*, page 375. Kluwer Academic, Dordrecht, 1993.

[19] A.N. Lasenby, C.J.L. Doran, and S.F. Gull. Cosmological consequences of a flat-space theory of gravity. In F. Brackx, R. Delanghe, and H. Serras, editors, *Clifford Algebras and their Applications in Mathematical Physics (1993)*, page 387. Kluwer Academic, Dordrecht, 1993.

[20] D. Hestenes. Proper particle mechanics. *J. Math. Phys.*, 15(10):1768, 1974.

[21] C.J.L. Doran, D. Hestenes, F. Sommen, and N. van Acker. Lie groups as spin groups. *J. Math. Phys.*, 34(8):3642, 1993.

[22] D. Hestenes. Observables, operators, and complex numbers in the Dirac theory. *J. Math. Phys.*, 16(3):556, 1975.

[23] S.W. Hawking and G.F.R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge University Press, 1973.

[24] C.W. Misner, K.S. Thorne, and J.A. Wheeler. *Gravitation*. W.H. Freeman and Company, San Francisco, 1973.

[25] I.D. Novikov and V.P. Frolov. *Physics of Black Holes*. Kluwer Academic, Dordrecht, 1989.

[26] E.T. Copson. On electrostatics in a gravitational field. *Proc. R. Soc. A*, 118:184, 1928.

[27] B. Linet. Electrostatics and magnetostatics in the Schwarzschild metric. *J. Phys. A*, 9(7):1081, 1976.

[28] K.S. Thorne, R.H. Price, and D.A. Macdonald. *Black Holes: The Membrane Paradigm*. Yale University Press, 1986.

[29] R. S. Hanni and R. Ruffini. Lines of force of a point charge near a Schwarzschild black hole. *Phys. Rev. D.*, 8(10):3259, 1973.

[30] A.G. Smith and C.M. Will. Force on a static charge outside a Schwarzschild black hole. *Phys. Rev. D*, 22(6):1276, 1980.

[31] J.V. Narlikar. *Introduction to Cosmology.* Cambridge University Press, 1993.

[32] C.J. Isham and J.E. Nelson. Quantization of a coupled Fermi field and Robertson-Walker metric. *Phys. Rev. D*, 10(10):3226, 1974.