**Advances in
Applied Clifford Algebras**

CrossMark

Geometric Algebra as a Unifying Language for Physics and Engineering and Its Use in the Study of Gravity

Anthony N. Lasenby ⓘ *

Abstract. Geometric Algebra (GA) is a mathematical language that aids a unified approach and understanding in topics across mathematics, physics and engineering. In this contribution, we introduce the Space-time Algebra (STA), and discuss some of its applications in electromagnetism, quantum mechanics and acoustic physics. Then we examine a gauge theory approach to gravity that employs GA to provide a coordinate free formulation of General Relativity, and discuss what a suitable Lagrangian for gravity might look like in two dimensions. Finally the extension of the gauge theory approach to include scale invariance is briefly introduced, and attention drawn to the interesting properties with respect to the cosmological constant of the type of Lagrangians which are favoured in this approach. The intention throughout is to provide a survey accessible to anyone, equipped only with an introductory knowledge of GA, whether in maths, physics or engineering.

Mathematics Subject Classification. 00A79, 83C02 (Physics and General Relativity).

Keywords. Geometric algebra, Mathematical physics, Gravitation, Relativity, Electromagnetism, Acoustic physics.

1. Introduction

Geometric Algebra (GA) is an extremely useful approach to the mathematics of physics and engineering, that allows one to use a common language in a wide variety of contexts. For example, complex variables, vectors, quaternions, matrix theory, differential forms, tensor calculus, spinors and twistors, are all subsumed under a common approach. This therefore results in great efficiency, enabling one quickly to get into new areas. It also tends to suggest new geometrical, and therefore physically clear, and coordinate-independent

*Corresponding author.

ways of looking at things. Despite the title of this paper, we are not going to attempt a survey of each area (there are good talks by experts in each field already in this volume). Instead we want to look briefly at *why* GA is so useful, and then look at a couple of areas, Electromagnetism and Acoustic Physics, in a bit more detail to illustrate the usefulness of this approach. We then pass on to consider the use of the same tools of GA in the study of Gravity, showing that one can quickly achieve novel insights and understanding. Topics we consider, in a simple fashion made possible by GA, include how gauge theory in flat space can replace parts of the differential geometry needed for gravity, and issues about what the correct Lagrangian for gravity should be.

1.1. Some Main Features of the Usefulness of GA

One of the major aspects of the usefulness of GA, at least in my own experience, is that one can do virtually everything with just geometric objects in spacetime. For introductions to the GA of 3d space see the book [9] by David Hestenes and the article [14] by Alan Macdonald in this volume. Here we explain the elements of GA that we will work with in 4d spacetime, traditionally known as the STA (Spacetime Algebra).

The STA is the geometric algebra of spacetime [2,7], and is generated by four vectors $\{\gamma_\mu\}$ which satisfy

$$\gamma_\mu \cdot \gamma_\nu = \tfrac{1}{2}(\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu) = \eta_{\mu\nu} = \text{diag}(+---). \tag{1.1}$$

Throughout, Greek indices run from 0 to 3 and Latin indices run from 1 to 3. We use a signature in which $\gamma_0^2 = -\gamma_i^2 = 1$, and natural units $c = \hbar$ are assumed throughout. The full STA is spanned by

$$\begin{array}{ccccc} 1 & \{\gamma_\mu\} & \{\gamma_\mu \wedge \gamma_\nu\} & \{I\gamma_\mu\} & I = \gamma_0\gamma_1\gamma_2\gamma_3 \\ 1 \text{ scalar} & 4 \text{ vectors} & 6 \text{ bivectors} & 4 \text{ trivectors} & 1 \text{ pseudoscalar} \end{array} \tag{1.2}$$

The algebraic properties of the STA are those of the Dirac matrices, but there is never any need to introduce an explicit matrix representation in calculations.

Suppose now that we wish to study physics in the rest frame defined by the γ_0 vector, which we can think of as the velocity vector of a specific observer. We define

$$\sigma_k = \gamma_k\gamma_0, \tag{1.3}$$

so that

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij} \tag{1.4}$$

The set $\{\sigma_i\}$ therefore generate the geometric algebra of the three-dimensional space orthogonal to γ_0. We call the three σ_i and their linear combinations, *relative vectors* to indicate that their definition depends on the choice of the timelike vector γ_0. We see that

$$\sigma_1\sigma_2\sigma_3 = \gamma_1\gamma_0\gamma_2\gamma_0\gamma_3\gamma_0 = \gamma_0\gamma_1\gamma_2\gamma_3 = I, \tag{1.5}$$

so relative space and spacetime share the same pseudoscalar. The algebra of space is therefore the *even subalgebra* of the STA. This subalgebra contains the scalars and pseudoscalars, and six (spacetime) bivectors. These bivectors are split into timelike and spacelike bivectors by the chosen velocity vector
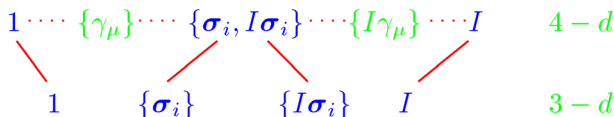
FIGURE 1. The even elements of the 4d spacetime algebra provide the full geometric algebra of 3d space

(γ_0 in this case)—see Fig. 1. This construction can be generalised to a general velocity v—see Section 5.2.3 in [2].

1.2. Lorentz Transformations

The geometry of spacetime is usually expressed in terms of Lorentz transformations, and these in turn are usually expressed as a *coordinate transformation*, e.g.

$$\begin{aligned} x' &= \gamma(x - \beta t), \quad t' = \gamma(t - \beta x) \\ x &= \gamma(x' + \beta t'), \quad t = \gamma(t' + \beta x') \end{aligned} \tag{1.6}$$

where $\gamma = (1 - \beta^2)^{-1/2}$ and β is the scalar velocity. Note we are using the symbol x for the x-coordinate to distinguish it from the 4d position vector x.

As an alternative to passive coordinate transformations, the GA approach encourages one to think about *frame axes* as being actively changed in a transformation. This articulates well with the very useful machinery of *reciprocal frames*, which we explain briefly here.

Suppose the position vector x is decomposed in two frames, $\{e_\mu\}$ and $\{e'_\mu\}$,

$$x = x^\mu e_\mu = x^{\mu'} e'_\mu \tag{1.7}$$

The relation of coordinates to these frames comes from the notion of reciprocal frame. Given the frame $\{e_\mu\}$ we define the reciprocal frame $\{e^\mu\}$ via

$$e^\mu \cdot e_\nu = \delta^\mu_\nu \tag{1.8}$$

where δ^μ_ν is the Kronecker delta. So e.g. $\gamma^i = -\gamma_i$, $(i = 1, 2, 3)$, and $\gamma^0 = \gamma_0$. With these definitions then

$$x^\mu = x \cdot e^\mu \tag{1.9}$$

and so for example

$$t = e^0 \cdot x, \qquad t' = e^{0'} \cdot x \tag{1.10}$$

This is very useful for working with curvilinear coordinates in particular, and articulates well with Geometric Calculus.

Concentrating on the 0, 1 components:

$$t e_0 + x e_1 = t' e'_0 + x' e'_1, \tag{1.11}$$

from which we derive the *vector* relations

$$e'_0 = \gamma(e_0 + \beta e_1), \quad e'_1 = \gamma(e_1 + \beta e_0) \tag{1.12}$$

which give the new frame in terms of the old. We now introduce the 'hyperbolic angle' or 'rapidity' α,

$$\tanh \alpha = \beta, \qquad (\beta < 1), \tag{1.13}$$

which gives

$$\gamma = (1 - \tanh^2\alpha)^{-1/2} = \cosh\alpha. \tag{1.14}$$

The vector e_0' is now

$$e_0' = \cosh\alpha\, e_0 + \sinh\alpha\, e_1 = (\cosh\alpha + \sinh\alpha\, e_1 e_0)e_0 \\ = e^{\alpha\, e_1 e_0} e_0 = e^{\alpha\, e_1 e_0/2} e_0 e^{-\alpha\, e_1 e_0/2} \tag{1.15}$$

Similarly, we have

$$e_1' = \cosh\alpha\, e_1 + \sinh\alpha\, e_0 = e^{\alpha\, e_1 e_0} e_1 = e^{\alpha\, e_1 e_0/2} e_1 e^{-\alpha\, e_1 e_0/2} \tag{1.16}$$

The other two frame vectors are unchanged, and we have written the transformation equations for e_0 and e_1 in the split 'half-angle' form so that exactly the same transformation applies to e_2 and e_3 as well. Defining

$$R = e^{\alpha\, e_1 e_0/2} \tag{1.17}$$

we see the relationship between the frames is

$$e_\mu' = R e_\mu \tilde{R}, \quad e^{\mu'} = R e^\mu \tilde{R}, \quad R\tilde{R} = 1. \tag{1.18}$$

Here $\tilde{R}$ denotes the *reverse* of R, which has the same elements of each grade as R, except each geometric product, e.g. $ab\ldots c$ is written in the reverse order $c\ldots ba$. Also $R\tilde{R} = 1$, for an even element R, is the definition of a *rotor*. These are all the same definitions as for rotations in 3d. Thus the same *rotor* prescription works for *boosts*, i.e. hyperbolic rotations between the timelike and spacelike axes, as for 3d rotations [14]. Spacetime is a unified entity now.

1.3. The Electromagnetic Field Strength

We wish to illustrate how useful this version of the Lorentz transformations is, by looking at the *electromagnetic field strength tensor*, or *Faraday* tensor.

The Faraday tensor is a rank-2 antisymmetric quantity $F^{\mu\nu}$. As a matrix, it has components

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \tag{1.19}$$

This is the standard definition, but it hides the natural complex structure.

In our version, F is a spacetime *bivector*. We generate it from the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$, which are relative vectors in the 3-space orthogonal to the time axis γ_0 (and therefore bivectors in the full STA), via the very simple expression

$$F = \boldsymbol{E} + I\boldsymbol{B} \tag{1.20}$$

which we can relate to the matrix elements via

$$F^{\mu\nu} = (\gamma^\nu {\wedge} \gamma^\mu){\cdot}F \tag{1.21}$$

(Note here the '$\cdot$' in $(\gamma^\nu{\wedge}\gamma^\mu){\cdot}F$ is the standard GA inner product, which for blades A_r and B_s, returns the lowest grade element in the geometric product $A_r B_s$. We mention this to distinguish this product from the left and right

contractions which are sometimes used in GA.) Since $\gamma_0 F \gamma_0 = (-\boldsymbol{E} + I\boldsymbol{B})$, we can recover $\boldsymbol{E}$ and $I\boldsymbol{B}$ individually from

$$
\begin{aligned}
\boldsymbol{E} &= \tfrac{1}{2}(F - \gamma_0 F \gamma_0) \\
I\boldsymbol{B} &= \tfrac{1}{2}(F + \gamma_0 F \gamma_0)
\end{aligned}
\tag{1.22}
$$

The split into $\boldsymbol{E}$ and $I\boldsymbol{B}$ depends on the *observer velocity* (γ_0). Different observers measure different fields.

A second observer, velocity $v = R\gamma_0\tilde{R}$, with comoving frame $\gamma'_\mu = R\gamma_\mu\tilde{R}$, measures the components of electric field

$$
E'_i = (\gamma'_i\gamma'_0)\cdot F = (R\sigma_i\tilde{R})\cdot F = \sigma_i\cdot(\tilde{R}FR)
\tag{1.23}
$$

This is the same transformation law as for vectors (here for the transformation of frames rather than the physical quantities themselves, hence the inverse transformation $\tilde{R}FR$ is found for F), and provides a very efficient machinery for transformation of all electromagnetic quantities, which we illustrate with a couple of examples.

- *Example 1* Stationary charges in the γ_0 frame set up a field

$$
F = \boldsymbol{E} = E_x\sigma_1 + E_y\sigma_2
\tag{1.24}
$$

with $E_z = 0$. What are the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields measured by a second observer, moving with velocity $\tanh\alpha$ in the γ_1 direction? For this the relevant rotor is

$$
R = e^{\alpha\sigma_1/2}
\tag{1.25}
$$

and the second observer measures the σ_i components of

$$
\tilde{R}FR = e^{-\alpha\sigma_1/2}Fe^{\alpha\sigma_1/2} = E_x\sigma_1 + E_y e^{-\alpha\sigma_1}\sigma_2
\tag{1.26}
$$

This gives

$$
E'_x = E_x, \quad E'_y = \cosh\alpha\, E_y, \quad B'_z = -\sinh\alpha\, E_y
\tag{1.27}
$$

The ease with which we can carry this out compares favourably with the same done with tensors.

- *Example 2* What are the *electromagnetic invariants* that can be constructed from the Faraday bivector? We take

$$
F^2 = \langle FF \rangle + \langle FF \rangle_4 = \alpha + I\beta
\tag{1.28}
$$

But

$$
(\tilde{R}FR)(\tilde{R}FR) = \tilde{R}(\alpha + I\beta)R = \alpha + I\beta
\tag{1.29}
$$

so both are *Lorentz invariant*, i.e. *independent* of the observer frame. In the γ_0 frame

$$
\begin{aligned}
\alpha &= \langle(\boldsymbol{E} + I\boldsymbol{B})(\boldsymbol{E} + I\boldsymbol{B})\rangle = \boldsymbol{E}^2 - \boldsymbol{B}^2 \\
\beta &= -\langle I(\boldsymbol{E} + I\boldsymbol{B})(\boldsymbol{E} + I\boldsymbol{B})\rangle = 2\boldsymbol{E}\cdot\boldsymbol{B}
\end{aligned}
\tag{1.30}
$$

The first is the *Lagrangian density*, while the second is less commonly encountered. The invariance of both together means that if $\boldsymbol{E}$ and $\boldsymbol{B}$ are of the same magnitude and perpendicular in one frame, then they are in all frames, which is relevant to radiation fields.

1.4. Composition of Boosts

As a further example of the utility of 4d rotors, but in a different area from electromagnetism, we will look at the same topic as treated by Macdonald [14] in his Section 4.1.4, namely the composition of pure boosts. We have already said a boost is a hyperbolic rotation between the timelike and space-like axes, and a spacetime rotor R (which remember is an even element of the 4d geometric algebra satisfying $R\tilde{R} = 1$), will in general represent a mixture of such rotations with purely spatial rotations. The definition of a pure boost depends on singling out a particular timelike direction—we will assume this is the γ_0 axis, though one can work more generally as well, and we shall give a 4d formula for the spatial rotation which is generated when two pure boosts in different directions are combined. This formula gives a neat explicit result, which aids understanding of how this spatial part, which is maybe somewhat unexpected, and of course is intimately linked with the phenomenon of Thomas precession, comes about.

A rotor L is a boost relative to γ_0 if

$$v = L\gamma_0\tilde{L} = L^2\gamma_0 \tag{1.31}$$

and we can write explicitly that

$$L = \frac{1 + v\gamma_0}{\sqrt{2\left(1 + v\cdot\gamma_0\right)}} = \exp\left(\frac{\alpha}{2}\frac{v\wedge\gamma_0}{|v\wedge\gamma_0|}\right) \tag{1.32}$$

where $\cosh(\alpha) = v\cdot\gamma_0$.

So we suppose that we have two boosts, L_1 and L_2 with corresponding velocities v_1 and v_2. We now define

$$R = L_2L_1, \quad \text{and} \quad v = R\gamma_0\tilde{R} \tag{1.33}$$

In general R will not be a pure boost, but will have a spatial rotation part U which commutes with γ_0. Writing $R = LU$, The 'L' part of R is given immediately by (1.32), and we can then form the U part via

$$U = \tilde{L}R \tag{1.34}$$

We can see this encodes the spatial-only part of the rotation since

$$U\gamma_0\tilde{U} = \tilde{L}R\gamma_0\tilde{R}L = \tilde{L}vL = \gamma_0 \tag{1.35}$$

We can get a more explicit formula for U as follows. We have that

$$\tilde{L} = \frac{1 + \gamma_0 v}{\sqrt{2\left(1 + v\cdot\gamma_0\right)}} = \frac{1 + \gamma_0 L_2 L_1 \gamma_0 \tilde{L}_1 \tilde{L}_2}{\sqrt{2\left(1 + v\cdot\gamma_0\right)}} \tag{1.36}$$

Thus the expression for U is

$$U = \frac{L_2L_1 + \gamma_0 L_2 L_1 \gamma_0}{\sqrt{2\left(1 + v\cdot\gamma_0\right)}} \tag{1.37}$$

and we can see that U is the projection of the composite rotor $R = L_2L_1$ into the part of 'rotor space' that commutes with γ_0. This provides an interesting alternative viewpoint on the (perhaps more explicit) $3 + 1$ formulae given in [14]. (See also [11] for a further $3 + 1$ derivation.)

2. Quantum Mechanics

The STA treatment of quantum mechanics is a large subject (see e.g. [2,3,8]), and we give only a brief treatment, highlighting how different this approach can be from the usual ones, and how it aids a unified understanding of several different 'wave equations'.

First let us consider the Pauli algebra. Equation (1.4) above showed how the σ_i vectors provide a representation-free version of the Pauli matrices $\hat{\sigma}_i$, which are defined as

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.1}$$

But what about what they operate on? Conventionally the $\hat{\sigma}_i$ act on 2-component *Pauli spinors*

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{2.2}$$

with ψ_1 and ψ_2 complex. In the GA approach, something rather remarkable happens, we can replace both objects (operators and spinors), by elements of the *same* algebra, via the following equivalence (conventional on the left, STA on the right):

$$|\psi\rangle = \begin{pmatrix} a^0 + ia^3 \\ -a^2 + ia^1 \end{pmatrix} \quad \leftrightarrow \quad \psi = a^0 + a^k I \sigma_k \tag{2.3}$$

Thus for example, the spin-up $|+\rangle$, and spin-down $|-\rangle$ states are translated by

$$|+\rangle \leftrightarrow 1 \qquad |-\rangle \leftrightarrow -I\sigma_2 \tag{2.4}$$

The action of the quantum operators $\{\hat{\sigma}_k\}$ on states $|\psi\rangle$ has an analogous operation on the multivector ψ:

$$\hat{\sigma}_k|\psi\rangle \leftrightarrow \sigma_k\psi\sigma_3 \quad (k = 1, 2, 3) \tag{2.5}$$

which can be verified by an explicit computation. We can understand the STA object ψ as follows. As an even element of the 3d geometric algebra of space, it satisfies

$$\psi\tilde{\psi} = (a^0)^2 + (a^1)^2 + (a^2)^2 + (a^3)^2 = \rho^2 \tag{2.6}$$

where ρ is a positive scalar quantity. (We note in passing that formulae such as this are familiar from the theory of quaternions whose algebra is isomorphic to the even sub-algebra of the 3d GA, the specific correspondence being

$$1, i, j, k \leftrightarrow 1, I\sigma_1, -I\sigma_2, I\sigma_3.) \tag{2.7}$$

We can then define

$$R = \rho^{-1/2}\psi, \quad \text{so that} \quad R\tilde{R} = 1 \tag{2.8}$$

meaning that the 'non-scale' part of ψ is a 3d rotor. This is very interesting in meaning that Pauli spinors are (up to scale) instructions for rotating objects in 3d space.

This view offers a number of insights. For example, the spin-vector $\boldsymbol{s}$ defined by

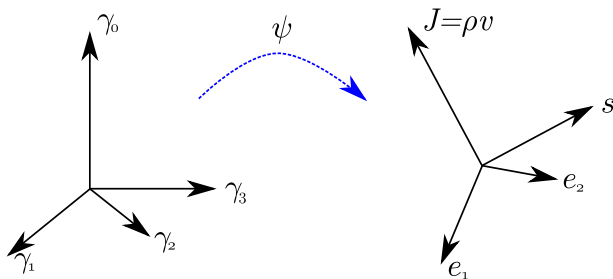$$\langle\psi|\hat{\sigma}_k|\psi\rangle = \sigma_k\cdot\boldsymbol{s}. \tag{2.9}$$

FIGURE 2. The Lorentz spinor ψ pictured as an instruction to rotate and dilate the fixed frame axes $\{\gamma_\mu\}$ to a new 'body frame' attached to the particle

can now be written as

$$s = \rho R \sigma_3 \tilde{R}. \tag{2.10}$$

The double-sided construction of the expectation value contains an instruction to rotate the fixed σ_3 axis into the spin direction and dilate it.

Also, suppose that the vector s is to be rotated to a new vector $R_0 s \tilde{R}_0$. The rotor group combination law tells us that R transforms to $R_0 R$. This induces the spinor transformation law

$$\psi \mapsto R_0 \psi. \tag{2.11}$$

This explains the 'spin-1/2' nature of spinor wave functions.

Similar things happen in the relativistic case. Instead of the wavefunction being a weighted spatial rotor, it is now a full Lorentz spinor of the type discussed in Sect. 1.4, which we can decompose into

$$\psi = \rho^{1/2} e^{I\beta/2} R \tag{2.12}$$

with the addition of a slightly mysterious β term related to antiparticle states (see e.g. [5] for further discussion of this decomposition, and the role of β). There are five observables in all, including the current, $J = \psi\gamma_0\psi = \rho R\gamma_0\tilde{R}$, and the spin vector $s = \psi\gamma_3\psi = \rho R\gamma_3\tilde{R}$, and a pictorial representation of the observables corresponding to the 4 frame vectors is given in Fig. 2.

This picture provides a very interesting link with the GA treatment of rigid body mechanics, which has motivated much of David Hestenes' work on the electron [10].

Another very significant unification is from the nature of the derivative operator

$$\nabla \equiv \gamma^\mu \partial_\mu \tag{2.13}$$

Using this, and the Faraday bivector $F = \boldsymbol{E} + I\boldsymbol{B}$ discussed above, the entire set of Maxwell equations can be written as

$$\nabla F = J \tag{2.14}$$

as we discuss further below. The point we wish to make here is that someone used to the STA version of these equations, perhaps in an engineering application, can immediately proceed to understanding the wave equation for the

neutrino
$$\nabla\psi = 0 \tag{2.15}$$
In fact their only problem is they won't know whether neutrinos are *Majorana*, in which case
$$\psi = \phi\frac{1}{2}(1 + \sigma_2) \tag{2.16}$$
(here ϕ is a Pauli spinor, and the idempotent $\frac{1}{2}(1 + \sigma_2)$ removes 4 d.o.f.), or a full Dirac spinor ψ, since currently no one knows this! They could then proceed to the Dirac equation
$$\nabla\psi I\sigma_3 = m\psi\gamma_0 \tag{2.17}$$
where the $I\sigma_3$ at the right of ψ reveals a geometrical origin for the unit imaginary of quantum mechanics.

Such a person might then wonder about generalising this choice, and allowing spatial rotations at the right of ψ to transform between $I\sigma_1$, $I\sigma_2$ and $I\sigma_3$. This would then be the $SU(2)$ part of electroweak theory! So our hypothetical investigator, starting in engineering applications, would have succeeded in getting quite a long way into High Energy physics with exactly the same tools as needed for e.g. electromagnetism and rigid body mechanics.

3. Electromagnetism

We now return to electromagnetism to examine two further applications in which the power of the STA is evident, and the second of which also provides a prelude for what we shall look at in acoustic physics.

We have already said that defining the Lorentz-covariant field strength $F = \boldsymbol{E} + I\boldsymbol{B}$ and additionally the current $J = (\rho + \boldsymbol{J})\gamma_0$, we obtain the single, covariant equation
$$\nabla F = J \tag{3.1}$$

The advantage here is not merely notational—just as the geometric product is invertible, unlike the separate dot and wedge product, the geometric product with the vector derivative is invertible (via multivector Green's functions) where the separate divergence and curl operators are not. This leads to the development of a new method for calculating the EM response of a system to incoming plane radiation, which is partially described in Sections 7.5 and 7.6 of [2]. For a problem with harmonic time dependence it proves convenient to use a complex representation (with uninterpreted imaginary i) for the Faraday bivector, and we write
$$F(x) = F(\boldsymbol{r})e^{-i\omega t} \tag{3.2}$$

The appropriate Green's function for this case is then
$$G(\boldsymbol{r}) = \frac{e^{i\omega r}}{4\pi}\left(\frac{i\omega}{r}(1 - \boldsymbol{\sigma}_r) + \frac{\boldsymbol{r}}{r^3}\right) \tag{3.3}$$
where $\boldsymbol{\sigma}_r = \boldsymbol{r}/r$ is the unit (relative) vector in the direction $\boldsymbol{r}$. (See also the equivalent expression in [16].) The advantages of using this multivector Green's function are (a) it provides a first order formulation—we just need
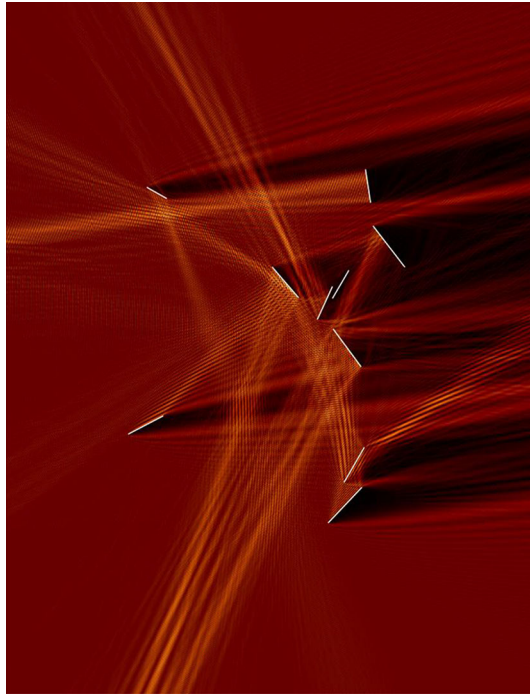
FIGURE 3. Image showing the response of a set of conducting plates (mirrors) to a plane electromagnetic wave coming in from the *left*, using a technique based on the Green's function (3.3).

a single integral over a surface to compute the fields at all spatial positions, thus conforming closely to what we would expect from a Huygen's principle formulation, and (b) provides the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields simultaneously, along with all necessary polarisation and obliquity factors, which is not achieved in such a simple way in conventional approaches (see e.g. Chapter 12 of [1] for a good exposition of the standard techniques).

Methods of this kind were employed to provide a fast route to calculating the response of a set of conducting plates or facets to incoming plane waves, with a typical result as shown in Fig. 3. The technique was fast enough that it was possible to change the illumination arbitrarily in real time and see the effects this caused, whilst working on just a single GPU. It should be stressed that the calculations involved were *exact*, i.e. all diffractive and other effects are included.

3.1. Radiation from a Moving Charge

For a more detailed example, and as preparation for the consideration of acoustic physics, we will now consider radiation from a moving charge. David Hestenes and Gull et al. [4] introduced the relevant STA techniques on this
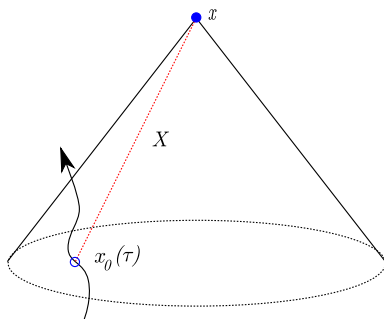
FIGURE 4. An observer at spacetime position x receives electromagnetic influences from the point $x_0(\tau)$ on a particle's world line, and is joined to this point by the null vector $X = x - x_0(\tau)$

some years ago, but they are still not widely known, and also we want to highlight a particular numerical implementation.

So we have a charge q moving arbitrarily, and wish to calculate the radiation fields this produces. From the fact that $\nabla \wedge F = 0$, we know we can introduce a vector potential A such that $F = \nabla \wedge A$. If we impose the Lorenz gauge $\nabla \cdot A = 0$, so that $F = \nabla A$, then A obeys the wave equation

$$\nabla F = \nabla^2 A = J \tag{3.4}$$

Since radiation does not travel backwards in time, the electromagnetic influence propagates along the future light-cone of the charge, and it is where this future cone intersects the observer position x that we need the fields. Conversely, taking a backward light-cone from x, we can ask where it intersects the path of the particle in spacetime. We call the path of the particle $x_0(\tau)$, which is parameterised in terms of the particle proper time τ. This situation is shown in Fig. 4. Clearly the separation vector down the light-cone $X = x - x_0$ is null, and this quantity is key to what follows (both here, and in acoustic physics).

Now suppose that the charge is in fact at rest, so its velocity vector is γ_0, and that its path through spacetime is $x_0(\tau) = \tau \gamma_0$ relative to the origin in the γ_0 frame.

The key observation is now the following. If the spacetime observer (assumed also at rest) is located at a spatial distance r from the charge, then for an observer time t we know that the τ at the intersection point of the backward light-cone from the observer with the path of the particle, satisfies

$$r = t - \tau \tag{3.5}$$

This is because the travel time available to get from the particle to the observer is $t - \tau$ and the distance needed to be traversed is r (we are taking $c = 1$).

The null vector X is therefore given by

$$X = (t\gamma_0 + re_r) - \tau\gamma_0 = r(\gamma_0 + e_r) \tag{3.6}$$

where e_r is a unit 4d vector in the r direction (so will satisfy $e_r^2 = -1$). We note this means $X \cdot \gamma_0 = r$.

Meanwhile, for this simple static case, we can immediately write down the potential A, since it is just the usual static Coulomb potential:

$$A = \frac{q}{4\pi r} \gamma_0 \tag{3.7}$$

We can turn this into a covariant expression by replacing the r by $X \cdot \gamma_0$, and then in turn replacing each occurrence of γ_0 by a general velocity v for the particle. This yields

$$A = \frac{q}{4\pi} \frac{v}{X \cdot v} \tag{3.8}$$

This expression is in fact the fully covariant result for the A field generated by an arbitrarily moving particle, and is the STA version of the Liénard–Wiechert potential. What 'fully covariant' means here, can be understood as follows. If we wish to transform to a new situation with a Lorentz rotated velocity $v' = R v \tilde{R}$, then the new A-field is given by

$$A'(x) = R A(\tilde{R} x R) \tilde{R} \tag{3.9}$$

i.e. the rotated A evaluated at the back-rotated spacetime position. We can illustrate this simply in the case where the velocity v is constant. Then the only position dependence in A lies in the $X(x)$ part of the $X \cdot v$ denominator. We can therefore predict that the new A should be

$$A' = \frac{q}{4\pi} \frac{v'}{X(\tilde{R} x R) \cdot v} \tag{3.10}$$

For a constant v we have that $x_0(\tau) = \tau v$ and so

$$X(\tilde{R} x R) = \tilde{R} x R - \tau v = \tilde{R} \left(x - \tau R v \tilde{R} \right) R = \tilde{R} \left(x - \tau v' \right) R \tag{3.11}$$

meaning that

$$A' = \frac{q}{4\pi} \frac{v'}{(x - \tau v') \cdot v'} \tag{3.12}$$

where the τ is chosen to make $(x - \tau v')$ null. This is precisely the original A:

$$A = \frac{q}{4\pi} \frac{v}{(x - \tau v) \cdot v} \tag{3.13}$$

with all vs replaced by v's, and illustrating the covariant nature of our construction in this special case. More generally, (3.9) will always work, and is effectively the definition of what we mean by covariance in the general case.

Now we want to find $F = \nabla A$. One needs a few differential identities of which the following is perhaps the most interesting. Since $X^2 = 0$, and using a small overcircle to denote the scope of a differential operator, we have

$$\begin{aligned}
0 = \mathring{\nabla}(\mathring{X} \cdot X) &= \mathring{\nabla}(\mathring{x} \cdot X) - \mathring{\nabla}(x_0\mathring{}(\tau) \cdot X) \\
&= X - \gamma^\mu(X \cdot \partial_\mu x_0(\tau)) \\
&= X - \gamma^\mu(X \cdot (\partial_\mu \tau)\partial_\tau x_0) \\
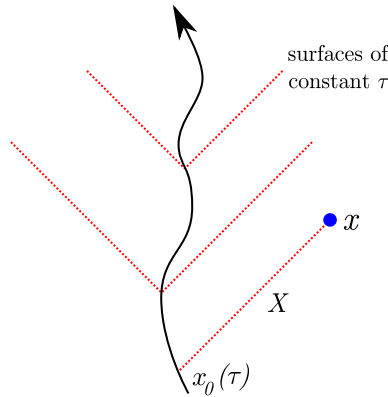&= X - (\nabla \tau)(X \cdot v)
\end{aligned} \tag{3.14}$$

FIGURE 5. The proper time τ at which the backwards light-cone from an observer at x intersects a particle's path, can be viewed as a scalar field filling spacetime, via extension along forward light-cones from the particle

where we used the chain rule in the third step. This result implies

$$\nabla\tau = \frac{X}{X \cdot v} \tag{3.15}$$

This is a very interesting result, since it shows us that the particle proper time at the past light-cone intersection can be treated as a spacetime field. Such a field, which does not carry any physical attributes (such as energy or angular momentum), is called an *adjunct field*, and we can think of it arising via the value of the field at $x_0(\tau)$ being extended over the charge's forward light-cone, as indicated in Fig. 5.

Proceeding using this result, and defining

$$\Omega_v = \dot{v}{\wedge}v \tag{3.16}$$

which is the *acceleration bivector*, then the result for F itself can be found relatively quickly. One finds

$$F = \frac{q}{4\pi} \frac{X \wedge v + \frac{1}{2}X\Omega_v X}{(X \cdot v)^3} \tag{3.17}$$

This equation displays a clean split into a Coulomb field in the rest frame of the charge, and a radiation term

$$F_{\text{rad}} = \frac{q}{4\pi} \frac{\frac{1}{2}X\Omega_v X}{(X \cdot v)^3} \tag{3.18}$$

proportional to rest-frame acceleration projected down the null vector X.

That this is a radiation term we can establish from its properties with distance. We can see from Eq. (3.6), that in the rest frame of the charge X scales with distance r. This means that the two Xs at the top and three at the bottom of (3.18) yield an overall dependence of 1/distance.
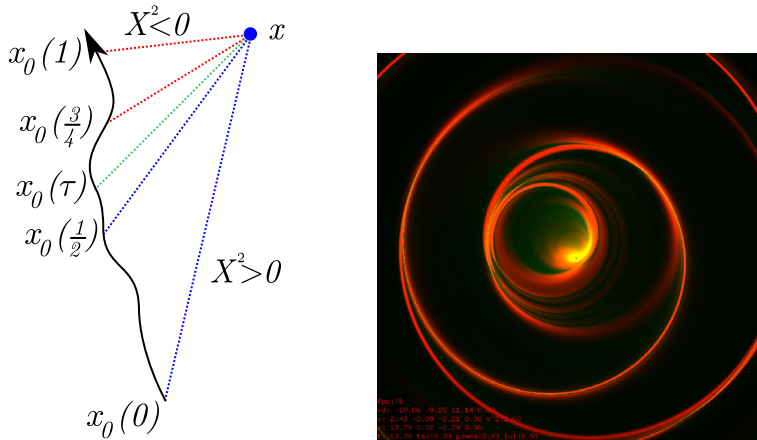
FIGURE 6. Calculations for the field produced by a moving charge. *Left* shows a bisection search for identifying the point where the backward light-cone from the observer's position intersects the particle worldline. *Right* shows some captured screen output for the magnitude of the Faraday bivector produced by a charge being moved in a circle

The *stress-energy tensor* for the electromagnetic field, which has quite a complicated expression when written in tensor form, is very simple in the STA. We find that the flux of energy and momentum in the direction of a vector a is given by

$$T(a) = -\frac{1}{2}FaF \qquad (3.19)$$

i.e. we simply *reflect a* in the bivector F to find the energy-momentum flow. From what we have just ascertained about the properties of the radiation part of F with distance, it is clear that the $T(a)$ due to this drops off as $1/\text{distance}^2$. Thus the surface integral of T does not vanish at infinity—energy-momentum is being carried away from the charge by radiation.

The expression for the electromagnetic field from a moving charge given in (3.17) is unique to the STA, and is probably the most compact in the literature, as well as (in the opinion of the author), being the most revealing of the underlying physics.

It also lends itself well to numerical solution techniques. In Fig. 6 we show two plots arising from an algorithm which was used to produce a plot of the field radiated from a charged particle, the movements of which are controlled in real time by the user. The left illustrates a bisection search which can be used to find the end point $x_0(\tau)$ of the vector X. Schematically the overall algorithm is:

- Store particle's history (position, velocity, acceleration)
- To calculate the fields at x, find the null vector X by bisection search (or similar)

- Retrieve the particle velocity and acceleration at the corresponding τ—the above formulae give us A and F

The image on the right of Fig. 6 is captured screen output in which the intensity is proportional to the magnitude of the Faraday bivector. Here we see the beginnings of synchrotron radiation as the particle is moved approximately in a circle. This is just a single moment from the continuous output produced as the particle is moved around (via mouse inputs) by the user, and the computations are carried out in real time on a single PC GPU. We give these details since programs of this kind may be useful in teaching applications, where students could get a real feel for EM radiation by being able to control its production themselves in this way.

Finally in this section, we note two further formulae which may be deduced from those we already have, and which are very striking.

$$A = -\frac{q}{8\pi}\nabla^2 X, \qquad F = -\frac{q}{8\pi}\nabla^3 X \qquad (3.20)$$

These make clear the primacy of the vector X, which we will now also encounter in the next topic.

4. Acoustic Physics

We now look at a perhaps surprising application of these techniques. This is the topic of the wave equation for linearised perturbations in a stationary fluid

$$\frac{1}{c_0^2}\frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial^2 \phi}{\partial y^2} - \frac{\partial^2 \phi}{\partial z^2} = 0 \qquad (4.1)$$

where c_0 is the speed of sound in the fluid. Oscillatory solutions for a moving source might look typically like the function imaged in Fig. 7
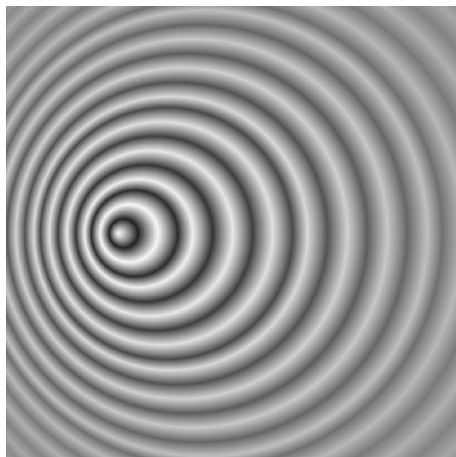

FIGURE 7. A typical solution of the wave equation (4.1) for an oscillatory moving source within the fluid

To make things look as simple as possible, and emphasise the tie-in with special relativity, we will henceforth use units of length such that $c_0 = 1$. (So for propagation in air, the unit of length is about 330 m.) The wave equation is then identical to

$$\nabla^2 \phi = 0 \tag{4.2}$$

where ∇^2 is the usual relativistic Laplacian, and we are using a Special Relativistic (SR) metric of the form $ds^2 = dt^2 - dx^2 - dy^2 - dz^2$.

We have written the wave equation without a source, but we want solutions corresponding to a δ-function source which follows a given path. It is known already that the solution for ϕ for this case corresponds to the electrostatic part of the electromagnetic Liénard–Wiechert potential (see e.g. [15]).

For a fluid source with variable strength $Q(t)$ then using the notation from [15]

$$\phi(t, x, y, z) = \frac{1}{4\pi} \frac{Q_s}{R_s \left(1 - M_s \cos \theta_s\right)} \tag{4.3}$$

Subscript s means that the corresponding quantity is evaluated at the *retarded* position. As in EM this is where backwards null cone from the observer's position $((t, x, y, z))$ intersects the world line of the source, and *null cone* is defined in terms of the above SR metric. M is the Mach number of the moving source, i.e. the ratio of its speed to the speed of sound in the fluid (assumed subsonic). θ is the angle between the source velocity vector and the observer's position, seen from the source, and R is the distance between the source position and the observer's position.

Can we tie this in what we have just looked at for EM? An immediate aspect we need to deal with, is that of course, in this Newtonian case, there is no concept of particle *proper time.* Instead, the time of the particle is the same as the time recorded by any observer, in whatever state of motion. This is essentially the defining characteristic of *Newtonian time,* which is universal, and flows equably and imperturbably, unaffected by anything else.

We can still define a *retarded time τ,* however, by exactly the same construction as above. This is the *Newtonian time at the point where the backward null-cone from the observer's position intersects the worldline of the particle.*

The key to what follows is to note that we can define a covariant Newtonian 4-velocity in the following way. Consider the particle moving in Newtonian time. We use the 'projective split' in which relative vectors in a 3d frame orthogonal to γ_0 are bivectors in the overall spacetime. Suppose the 3d track of the particle as a function of Newtonian time τ is $\boldsymbol{x}_0(\tau)$. The particle position in 4d Newtonian spacetime is then given by

$$x_{0N} = (\tau + \boldsymbol{x}_0(\tau)) \gamma_0 \tag{4.4}$$

so that we can define its Newtonian 4-velocity as

$$v_N = \frac{dx_{0N}}{d\tau} = (1 + \boldsymbol{M}) \gamma_0 \tag{4.5}$$

where $\boldsymbol{M}$ is the (relative) ordinary velocity divided by the sound speed.

The two key observations, on which the entire equivalence rests, are that (a), for a given particle path, although it will in general not have the same length, v_N is *in the same direction* as the relativistic 4-velocity v, and (b), the velocity v appears 'projectively' in the formula for the EM 4-potential (3.8). 'Projectively' here means that it appears only linearly, and any scale associated with it will cancel out between numerator and denominator.

Putting these observations together, means that an equally good expression for A in EM is

$$A = \frac{q}{4\pi} \frac{v_N}{X \cdot v_N} \tag{4.6}$$

in which just the Newtonian 4-velocity appears. We can then use the time part of this to provide the solution for the potential due to a moving source in a fluid. To illustrate this approach, we will now look at a well known phenomena for waves in a fluid, and derive an interesting 4d covariant form for the Doppler effect, which also turns out to have implications in electromagnetism.

4.1. The Doppler Effect

Starting with the fluid case, suppose we have a wave with modulation function $f(t, x, y, z)$, and a moving observer, with Newtonian 4-velocity

$$V_N = (1 + \mathbf{N}) \gamma_0 \tag{4.7}$$

say. ($\mathbf{N}$ here is being used to indicate the observer's ordinary velocity, divided by the sound speed.) The (4d) gradient in the V_N direction is

$$V_N \cdot \nabla = \frac{\partial}{\partial t} + \mathbf{N} \cdot \boldsymbol{\nabla} \tag{4.8}$$

which we recognise as the 'convective derivative' for the given observer. We then claim that $i V_N \cdot \nabla \ln f$ provides a covariant definition of the 'effective frequency' observed. As an example, if the modulation is purely harmonic:

$$f(t, x, y, z) = \exp\left(i\left(\mathbf{k} \cdot \mathbf{x} - \omega t\right)\right) \tag{4.9}$$

with $|\mathbf{k}| = \omega$, then we get

$$i V_N \cdot \nabla \ln f = \omega \left(1 - \hat{\mathbf{k}} \cdot \mathbf{N}\right) \tag{4.10}$$

as expected. (Note that as in the equivalent EM case, the i here is an uninterpreted imaginary, being used for convenience when we have modulation in time.)

We now apply this to our function $f(\tau)$ of Newtonian retarded time. This gives

$$i V_N \cdot \nabla \ln f = \frac{i\left(V_N \cdot \nabla \tau\right) df/d\tau}{f(\tau)} \tag{4.11}$$

where we have as usual used the chain rule in evaluating the ∇ applied to τ. But $i(df/d\tau)/f(\tau)$ is the effective frequency as observed at the emitter. Also we can use Eq. (3.15) with v replaced by v_N for evaluating $\nabla \tau$, since nothing in its derivation depended on $v^2 = 1$. So

$$\nabla \tau = \frac{X}{X \cdot v_N} \tag{4.12}$$

Putting these two facts together, we can deduce

$$\frac{\text{effective frequency measured by observer}}{\text{effective frequency at transmitter}} = \frac{V_N \cdot X}{v_n \cdot X} \qquad (4.13)$$

So we have a nice compact, covariant, expression for the Doppler effect, given solely in terms of 4d geometric quantities. We could of course have worked with purely harmonically varying quantities at a single frequency, but we wanted to illustrate that the essence of it rests with the action of $V_N \cdot \nabla$, and could in principle be applied to any time-varying quantity, to give an effective 'stretching' effect. Also, we can now go back to EM case, and can recover an interesting result there.

4.2. Return to Electromagnetism

We now need to work with the relativistic 4-velocities v and V instead of the Newtonian 4-velocities v_N and V_N, and the retarded proper time instead of retarded Newtonian time. The main difference arises in the 'convective derivative', which acquires a factor $\cosh\alpha$, where $\tanh\alpha = |\boldsymbol{N}|$, compared to the Newtonian case. However, this is exactly what is required to convert the 'laboratory frame' time t to the proper time of the observer. Thus we now get the result for EM

$$\frac{\text{effective frequency measured by observer}}{\text{effective frequency at transmitter}} = \frac{V \cdot X}{v \cdot X} \qquad (4.14)$$

This is an interesting expression for the redshift in special relativity, which I have not seen before! The usual expression, and one which works in general relativity (GR) as well, is derived by working with a photon with 4-momentum p. There one finds

$$\frac{\text{photon energy measured by observer}}{\text{photon energy at emission}} = \frac{V \cdot p}{v \cdot p} \qquad (4.15)$$

(In the GR version, the p in the numerator may be different from the p in the denominator, despite referring to the same photon, due to gravitational redshift.) We see that in the current approach, the role of the null-momentum p is taken over by the retarded null vector X, which is an interesting equivalence. Equation (4.14) is more general than (4.15), since it refers to the stretching or compression in time of any information flow from source to receiver.

5. Gauge Theory Gravity

The rest of this paper will concentrate on gravity, and in particular the Gauge Theory approach to gravity. There is not enough space to explain this properly, but continuing the theme of seeking to show how *Spacetime Algebra* is able to reach deep into modern physics using just the same tools (and entities!) as useful in classical physics and engineering applications, we will illustrate it in action in a very simple setting where one can see all the details. Hopefully this will convince the reader that the techniques involved are worth getting to grips with, and that as argued above for High Energy Physics,

advanced topics in gravitational theory are well within the reach of anyone equipped with the toolbox of geometric algebra.

The setting will be *Gravity in 2 dimensions*. We want to show how one can recover the needed results of differential geometry via a *gauge theory* approach. Then following this, we will consider an extension (in 4d) to a larger class of gauge symmetries I am excited about.

What is Gauge Theory Gravity? This is a version of gravity that aims to be as much as possible like our best descriptions of the other 3 forces of nature:

- the *strong force* (nuclei forces);
- the *weak force* (e.g. radioactivity etc.);
- *electromagnetism.*

These are all described in terms of *Yang-Mills type gauge theories* (unified in quantum chromodynamics) in a flat spacetime background. In the same way, *Gauge Theory Gravity* (*GTG*) is expressed in a flat spacetime. It has two gauge fields:

$\overline{h}(a)$: this allows an arbitrary remapping of position to take place in which $x \mapsto f(x)$ (position gauge change). The vector derivative ∇ is altered by this remapping and defining the transformation law for $\overline{h}(a)$ so as to undo this change, then $\overline{h}(\nabla)$ can become position gauge covariant. $\overline{h}(a)$ is a vector function of vectors and therefore has 16 d.o.f.

$\Omega(a)$: this allows Lorentz rotations to be gauged locally (rotation gauge change). Again, the introduction of this gauge field is necessary to make derivatives covariant, and it is a bivector function of vectors with 24 d.o.f.

Standard General Relativity (GR) cannot even see changes of the latter type, since the metric corresponds to $g_{\mu\nu} = \underline{h}^{-1}(e_\mu) \cdot \underline{h}^{-1}(e_\nu)$ and is invariant under local Lorentz rotations. Note here, and in writing $\overline{h}(a)$, we have introduced a useful notation in which the *adjoint* to a linear function $\underline{h}(a)$ (with the underbar or overbar denoting the linearity of the function), is given by $\overline{h}(a)$, and is defined by

$$\overline{h}(a) \cdot b = a \cdot \underline{h}(b) \tag{5.1}$$

for all vectors a and b.

The covariant derivative in the a direction is

$$\mathcal{D}_a \equiv a \cdot \nabla + \Omega(a) \times \tag{5.2}$$

The $\times$ operator is defined by $A \times B = \frac{1}{2}(AB - BA)$. If A is a bivector, then $A \times$ preserves the grade of the object being acted upon. Moreover, it is easy to show $A \times (BC) = (A \times B)C + B(A \times C)$, so that $A \times$ behaves similarly to a derivation. Thus, despite appearances, $\mathcal{D}_a$ is actually a scalar operator, i.e. preserves grades, and can be commuted through the objects it acts upon! We get the full vector covariant derivative via

$$\mathcal{D} \equiv \overline{h}(\partial_a)\mathcal{D}_a \tag{5.3}$$

∂_a here is part of the other new GA linear algebra machinery we need to introduce, and is the *multivector derivative* w.r.t. a, $\partial_a \equiv e^\mu \frac{\partial}{\partial a^\mu}$. As two

examples to give the flavour of this derivative, $\partial_a a = 4$ and $\nabla \equiv \partial_x$. The field strength tensor is got by commuting covariant derivatives, and one finds

$$[\mathcal{D}_a, \mathcal{D}_b]M = R(a \wedge b) \times M \quad \text{where } M \text{ is any multivector field} \tag{5.4}$$

This leads to the *Riemann tensor*

$$R(a \wedge b) = \partial_a \Omega(b) - \partial_b \Omega(a) + \Omega(a) \times \Omega(b) \tag{5.5}$$

Note this is a mapping of *bivectors* to *bivectors*. The Ricci scalar (rotation gauge and position gauge invariant) is

$$\mathcal{R} = \left[\overline{h}(\partial_b) \wedge \overline{h}(\partial_a) \right] \cdot R(a \wedge b) \tag{5.6}$$

The gravitational action is then $\mathcal{L}_{\text{grav}} = \det h^{-1} \mathcal{R}$, with the $\det h^{-1}$ inserted to make the action integral $\int d^n x \mathcal{L}_{\text{grav}}$ position gauge covariant. The dynamical variables are $\overline{h}(a)$ and $\Omega(a)$ and the field equations correspond to taking the *multivector functional derivatives* $\partial_{h(a)}$ and $\partial_{\Omega(a)}$. How these work is explained in Appendix B of the paper by Lasenby et al. [12], which can be consulted for a full exposition and many other details of Gauge Theory Gravity.

To help make this approach concrete in the current context, and also throw up some interesting issues, we will now illustrate Gauge Theory Gravity by considering the Gauge Theory approach to differential geometry in the simplest non-trivial context.

5.1. 2d Differential Geometry

The simple illustration of the Gauge Theory approach to differential geometry we want to give is the application to 2 Euclidean dimensions—there is not really a gravity theory available here (for interesting reasons we discuss), but this case is nevertheless instructive.

We start with a general h-function, which we write as

$$\begin{aligned} \overline{h}(e^1) &= f_1(x, y)e^1 + f_2(x, y)e^2 \\ \overline{h}(e^2) &= g_2(x, y)e^1 + g_1(x, y)e^2 \end{aligned} \tag{5.7}$$

where we have used the frame $\{e^1, e^2\}$, which is reciprocal to the basis frame $\{e_1, e_2\}$, in the definition. We also define a general Ω function

$$\begin{aligned} \Omega(e_1) &= A_1(x, y)I \\ \Omega(e_2) &= A_2(x, y)I \end{aligned} \tag{5.8}$$

Here $I = e_1 e_2$ is the pseudoscalar of the 2d space, and the functions f_i, g_i and A_i are all scalar functions of position in the 2d space.

Note, defining a vector field $A = A_i e^i$ we have

$$\Omega(a) = (a \cdot A)I \tag{5.9}$$

We can now find the Riemann tensor and Ricci scalar via Eqs. (5.5) and (5.6). The latter works out to something nice-looking in terms of the vector field A:

$$\mathcal{R} = 2 \left(\det h \right) \left(\nabla \wedge A \right) I \tag{5.10}$$

and we are guaranteed that this is position-gauge and rotation-gauge covariant. However, it turns out to be *disastrous* as regards being a suitable Lagrangian for gravity!

First, we know that we should multiply by $\det h^{-1}$ to form the position gauge covariant Lagrangian, so the full action is

$$S = \int d^2x \, \det h^{-1} \det h \, (\nabla \wedge A) \, I = \int d^2x \, (\nabla \wedge A) \, I \qquad (5.11)$$

Thus this doesn't even depend on h! Moreover, things are even worse. We can write

$$S = \int d^2x \, (\nabla \wedge A) \, I = \int d^2x \, \nabla \cdot A' = \int dl \, n \cdot A' \qquad (5.12)$$

where $A' = AI$ is the vector dual to A. The final integral is around the 'boundary' in 2d space and n is a vector normal to the boundary.

We see from this that we will not get any equations of motion—the action consists of just a *'topological'* boundary term. So *Einstein-Hilbert* gravity, based on just the first power of the Ricci scalar does not work in 2d.

So what should one do instead to get 2d gravity? (This is a subject of current research in *Quantum Gravity*—2d can provide a test bed for more complicated theories—see for example the interesting review talk by Daniel Grumiller [6].)

What if instead of $\mathcal{R}$ we used $\mathcal{R}^2$? Suddenly everything looks much more sensible. We get

$$S = \int d^2x \, \det h^{-1} \mathcal{R}^2 = -\int d^2x \, \det h \, F \cdot F \qquad (5.13)$$

where we have written $F = \nabla \wedge A$. In fact one can go further. Let us define $\mathcal{F} = \bar{h}(\nabla \wedge A)$, which is the 'covariant' version of F. Then

$$S = -\int d^2x \, \det h^{-1} \, \mathcal{F} \cdot \mathcal{F} \qquad (5.14)$$

So this is exactly the $\mathcal{F} \cdot \mathcal{F}$ Lagrangian of electromagnetism within Gauge Theory gravity! Thus we can tell that doing gravity in 2d is going to be a lot like doing electromagnetism. (Note that original $\mathcal{R}$ Lagrangian now looks very strange—it is equal to

$$\int d^2x \, \det h^{-1} \mathcal{F} I \qquad (5.15)$$

which would be an odd way of doing electromagnetism.)

We will not pursue the full setup here, but suffice to say that to determine the equations of motion for both h and Ω, we need to bring in the *torsion* defined by

$$\mathcal{Q}(\bar{h}(a)) \equiv \mathcal{D} \wedge \bar{h}(a) \qquad (5.16)$$

and then this provides another term (specifically $\mathcal{Q}(\partial_a) \cdot \mathcal{Q}(a))$ we can put in the Lagrangian, and which 'stiffens up' the equations for h.

This is effectively at the boundary of what people are working on for 2d gravity, but emerges very quickly and naturally here in this GTG approach.

Here to proceed further in terms of illustration, we will go back to the differential geometry aspects, and will assume for the rest of this development that torsion $= 0$, i.e.

$$\mathcal{D}\wedge\bar{\mathsf{h}}(a) = 0 \tag{5.17}$$

Using

$$\mathcal{D}\wedge\bar{\mathsf{h}}(a) = \bar{\mathsf{h}}(\nabla)\wedge\bar{\mathsf{h}}(a) + \bar{\mathsf{h}}(\partial_b)\wedge\left(\Omega(b)\cdot\bar{\mathsf{h}}(a)\right) \tag{5.18}$$

this gives a relation between $\bar{\mathsf{h}}$ and Ω that we can solve for Ω (i.e. A in this 2d case) in terms of $\bar{\mathsf{h}}$.

The details are not very instructive, so we will just jump straight to the answer we get for $\mathcal{R} = 2\det h\,(\nabla\wedge A)\,I$. Also, we will further specialise to where h, and therefore the implied metric $g_{\mu\nu}$ is diagonal, i.e. $g_{11} = 1/f_1^2$, $g_{22} = 1/g_1^2$.

We obtain

$$\mathcal{R} = \frac{1}{g_{11}g_{22}}\left\{-\frac{\partial^2 g_{11}}{\partial x_2^2} - \frac{\partial^2 g_{22}}{\partial x_1^2} + \frac{1}{2g_{11}}\left[\frac{\partial g_{11}}{\partial x_1}\frac{\partial g_{22}}{\partial x_1} + \left(\frac{\partial g_{11}}{\partial x_2}\right)^2\right]\right.$$
$$\left. + \frac{1}{2g_{22}}\left[\frac{\partial g_{11}}{\partial x_2}\frac{\partial g_{22}}{\partial x_2} + \left(\frac{\partial g_{22}}{\partial x_1}\right)^2\right]\right\}$$

Those used to differential geometry, will recognise this as the quantity which appears in (the diagonal version of) **Gauss' Theorema Egregium**, which we can state as: *No matter what coordinate transformations we carry out (thereby changing the $g_{\mu\nu}$ of course), then in two dimensions the quantity that we have just found is invariant, and its value is twice the Gaussian curvature, K.*

This is a very useful result in General Relativity, since we can work out many problems in 4d in terms of two-dimensional hypersurfaces, e.g. the (r, ϕ) plane for spherically symmetric systems, or the (t, r) plane for cosmology—2d is often all we need. Also since the metric tensor is symmetric, we can always diagonalise it, so we have derived the essential formula. Note in terms of the covariant $\mathcal{F}$, we have quite generally

$$K = \mathcal{F}I \tag{5.19}$$

which gives an interesting view of the Gaussian curvature

A special case worth looking at, where the value of A becomes transparent, is for a *conformal* metric, i.e. where

$$\bar{\mathsf{h}}(a) = f(x, y)\,a \tag{5.20}$$

for some scalar function f. We quickly find that

$$A = (\nabla \ln f)\,I \quad \text{which note means} \quad \nabla\cdot A = 0 \tag{5.21}$$

and thus A is already in Lorenz gauge.

For a constant curvature space (two-dimensional version of de Sitter space, or for the spatial sections of any Friedmann–Robertson–Walker metric) we find a possible solution is

$$f = \frac{1}{2}\left(1 + Kr^2\right) \tag{5.22}$$

where $r^2 = x^2 + y^2$, so we have here recovered the spatial part of the line element for constant curvature universes.

Note A has a form analogous for what we would expect for a 'constant magnetic field', but modified by the conformal factor

$$A = \frac{2K}{1 + Kr^2} (-y, x) \tag{5.23}$$

Overall it is hoped this has given the reader a feel for Gauge Theory Gravity, and how despite working in a flat space and without tensor calculus, it can recover standard differential geometry results in a simple fashion.

5.2. Progressing to Scale Invariance

Now we want to add an additional symmetry to those of position gauge and rotation gauge covariance. This is *scale invariance*, where we want to be able to rescale the h-function by an arbitrary function of position

$$\overline{h}(a) \mapsto e^{\alpha(x)}\overline{h}(a) \tag{5.24}$$

where $\alpha(x)$ is a scalar. Then the 'metric' obeys

$$g_{\mu\nu} = \underline{h}^{-1}(e_\mu) \cdot \underline{h}^{-1}(e_\nu) \mapsto g'_{\mu\nu} = \Omega(x)g_{\mu\nu} \quad \text{with} \quad \Omega(x) = e^{-2\alpha(x)} \tag{5.25}$$

and we want the physical quantities to respond *covariantly* under these changes. Note that the change where we remap x to an arbitrary function of x ($x \mapsto f(x)$), is already included in the position-gauge freedom, so we are not talking about $x \mapsto e^\alpha x$. Instead we mean a change in the standard of length at each point (the original Weyl idea [17]).

There are a variety of ways of going about this. I have been working (in the background!) on a novel approach to this for the last 8 years, and gave a preliminary account in the Brazil ICCA meeting in 2008, but a lot has changed since then. (I did not manage to write up the presentation, but see http://www.ime.unicamp.br/icca8/videos.html for a video of the talk if interested.)

With a colleague (Mike Hobson) we have now finished writing up the theoretical foundations of the work [13]. (This first paper is unfortunately (for current purposes) not in GA notation, and indeed conversion of some parts of this to conventional notation was one of the hardest tasks associated with the work!) I will not give details here, but want to give a flavour of it by considering a subset of the full theory, which ties into the discussion we have just had of 2d.

We can immediately get a version of scale invariance in 4d, by using as Lagrangian, not the Ricci scalar $\mathcal{R}$, but the 'square' of the Riemann tensor, which in GA form we can write

$$\mathcal{S} = \int d^4x \det h^{-1} \beta \mathcal{R} \left(\partial_b \wedge \partial_a\right) \cdot \mathcal{R}\left(a \wedge b\right) \tag{5.26}$$

where β is a coupling factor. The point about this, is that if we think of h transforming as $\exp(\alpha)$, then the Riemann tensor transforms as

$\exp(2\alpha)\mathcal{R}(B)$. Thus the overall integrand is of right 'conformal weight' to be scale invariant. The (standard) Ricci scalar version

$$\mathcal{S} = \int d^4x \det h^{-1} \frac{\mathcal{R}}{2\kappa}, \quad \text{where} \quad \kappa = 8\pi G \qquad (5.27)$$

fails this test, and so cannot lead to a scale-invariant theory. Also the Riemann^2 version has the right weight in terms of dimensions for the β coupling factor to be dimensionless (again the Ricci scalar version fails this test).

Very importantly as well, the Riemann^2 Lagrangian is exactly what we'd expect if we were to model gravity as a gauge theory just like the electroweak and strong forces! This is because the Riemann tensor is the gravitational version of the 'field strength tensor' of the other theories, which is always found by commuting covariant derivatives (as here). In *electroweak* and *QCD*, we then form an invariant Lagrangian, by contracting the field-strength tensor with itself—again as here. So this directly parallels e.g. the Maxwell structure $\mathcal{F}\cdot\mathcal{F}$, which we saw above emerging as a viable candidate for 2d gravity. A very interesting feature of this approach, is that *torsion*, i.e.

$$\mathcal{D}\wedge\bar{\mathsf{h}}(a) \neq 0 \qquad (5.28)$$

becomes inevitable in general, and that quantum spin becomes a source not just for torsion (as happens in standard Einstein–Cartan type theories), but for the Riemann tensor itself.

In this connection, a beautiful feature is that the gravitational field equations then become 'Maxwell-like' in form, e.g. the Ω equation is (schematically)

$$\dot{\mathcal{D}}\dot{\mathcal{R}}(B) = \frac{1}{\beta}\overline{\mathcal{Q}}(B) \qquad (5.29)$$

where $\overline{\mathcal{Q}}(B)$ is the adjoint of the 'spin source' tensor. So quantum spin can feed through directly to give gravitational effects.

However, what is perhaps most interesting to me as a cosmologist, is that this approach gives unique insights into the 'cosmological constant' problem. We now know that on the largest scales in the universe we see not extra attraction, but 'repulsion'. The universe is *accelerating*, as measured by the brightness of distant supernovae—see Fig. 8. Is this the cosmological constant Λ? There are big problems with the physics of this if we interpret it as due to vacuum energy, and use vacuum energy as a source term in the Einstein equations—as is well known, particle physics predictions in this scenario are too big by about 10^{120} compared to the Λ we observe!

However, one finds something remarkable happens with a Riemann squared Lagrangian. I have been able to prove

(a) that all vacuum solutions of GR with Λ are solutions of Riemann^2 without a Λ;

(b) that all cosmological solutions (technically those with vanishing Weyl tensor) of GR with Λ and a certain type of 'matter', are solutions of Riemann^2 without a Λ.
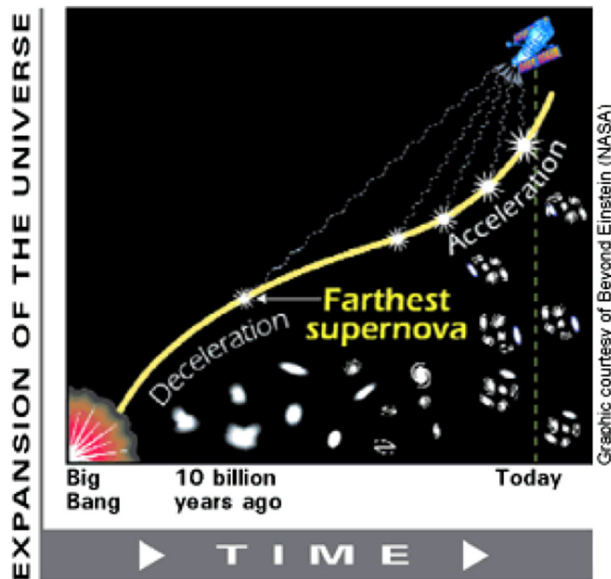
FIGURE 8. The evolution of the scale factor of the universe as traced by the brightness of distant supernovae

The effective Λ which is simulated in each case, is given by

$$\Lambda_{\text{eff}} = -\frac{3}{8\beta G} \qquad (5.30)$$

where β is the coupling constant mentioned above. So Λ can arise naturally from our modified gravity theory, and does not necessarily have to do with vacuum fluctuations as a source.

This is all very good. *However*, the big catch is the type of matter this works with. This also has to be *scale-invariant*—e.g. in cosmology one can use radiation, and have a radiation-filled universe (which was like our real universe for the first tens of thousands of years), but one cannot use ordinary baryonic matter, or indeed massive dark matter particles, such as dominate the universe today.

It was to get around this problem that I began exploring a more general scale-invariant theory, that *can* incorporate ordinary matter. The foundational aspects are now clear (see [13]). But coherently knitting together the applications so that one can determine whether it can evade all the current constraints on departures from GR, whilst doing useful things for Dark Energy and Λ is still not quite clear. Hopefully it is on the right lines, however!

Acknowledgments

References

[1] Born, M., Wolf, E.: Principles of optics: electromagnetic theory of propagation, interference and diffraction of light. CUP Archive (2000)

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

[3] Doran, C., Lasenby, A., Gull, S., Somaroo, S., Challinor, A.: Spacetime algebra and electron physics. Adv. Imaging Electron Phys. **95**, 271 (1996)

[4] Gull, S., Lasenby, A., Doran, C.: Imaginary numbers are not real—the geometric algebra of spacetime. Found. Phys. **23**, 1175 (1993)

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

[6] Grumiller, D.: http://quark.itp.tuwien.ac.at/~grumil/pdf/belgium (2011)

[7] Hestenes, D.: Space-Time Algebra. Birkhauser/Springer, Basel/Berlin (1966/2015)

[8] Hestenes, D.: Local observables in the Dirac theory. J. Math. Phys. **14**, 893 (1973)

[9] Hestenes, D.: New Foundations for Classical Mechanics. Kluwer, New York (2002)

[10] Hestenes, D.: Zitterbewegung in quantum mechanics. Found. Phys. **40**, 1 (2010)

[11] King, B.T.: An Einstein addition law for nonparallel boosts using the geometric algebra of space-time. Found. Phys. **25**, 1741 (1995)

[12] Lasenby, A., Doran, C., Gull, S.: Gravity, gauge theories and geometric algebra. Proc. R. Soc. Lond. Philos. Trans. Ser. A **356**, 487 (1998). arXiv:gr-qc/0405033

[13] Lasenby, A., Hobson, M.: Scale-invariant gauge theories of gravity: theoretical foundations (2015, arXiv e-prints). arXiv:1510.06699

[14] Macdonald, A.: A Survey of Geometric Algebra and Geometric Calculus, This Proceedings

[15] Rienstra, S., Hirschberg, A.: An Introduction to Acoustics. Eindhoven University of Technology (2013, online). http://www.win.tue.nl/~sjoerdr/papers/boek. Accessed 18 July 2016

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

[17] Weyl, H.: Gravitation und Elektrizität. Sitzungsber. Preuss. Akad. Wiss., Berlin, p. 465 (1918)

Anthony N. Lasenby
Kavli Institute for Cosmology
c/o Institute of Astronomy
Madingley Road
Cambridge CB3 0HA, UK

and

Astrophysics Group
Cavendish Laboratory
JJ Thomson Avenue
Cambridge CB3 0HE, UK
e-mail: `a.n.lasenby@mrao.cam.ac.uk`