

Electromagnetism using Geometric Algebra versus Components

1 Introduction

The task for today is to compare some more-sophisticated and less-sophisticated ways of expressing the laws of electromagnetism. In particular we compare Geometric Algebra, ordinary vectors, and vector components.

We do this in the spirit of the correspondence principle: whenever you learn a new formalism, you should check that it is consistent with what you already know.

Contents

1	Introduction	1
2	Preview	2
3	Vectors	3
4	Components	4
5	Electromagnetism using Geometric Algebra	4
6	Charge, Force, and Energy	6
6.1	Conservation of Charge	6
6.2	Lorentz Force Law	7
6.3	Lagrangian Density	9
6.4	Poynting Vector	9
6.5	Stress-Energy Tensor	10
7	Vector-ish Potential	11
7.1	The Basic Idea	11
7.2	$D = 3$ versus $D = 4$	11
7.3	Gauge Invariance	12
7.4	The Maxwell Equation in terms of the Vector-ish Potential	13

2	<i>PREVIEW</i>	2
8	Geometric Algebra – General Remarks	13
8.1	Overview	13
8.2	No Decorations	14
8.3	No Cross Product	14
8.4	No Chirality	14
8.5	Multiple Approaches	14
9	Pitfalls to Avoid	15
9.1	Definition of Dot Product	15
9.2	Unit Pseudoscalar	15
9.3	Exponents; Squared versus Norm Squared	16
9.4	Dot Product Not Necessarily Commutative	16
10	Additional Remarks	16
10.1	More about the Notation	16
10.2	Factors of c	17
11	An Application: Plane Waves	18
11.1	Running Waves	18
11.2	Running Wave Phase Relationships	20
11.3	Standing Wave Phase Relationships	21
11.4	Spacetime Picture	21
12	References	22

2 Preview

As we shall see in [section 5](#), Maxwell’s equations for the electromagnetic field can be written in the remarkably compact and elegant form:

$$\nabla F = \frac{1}{c\epsilon_0} J \tag{1}$$

where J a vector in spacetime, representing the charge and current, and F is a bivector, representing the electromagnetic field. It is worth learning the geometric algebra (aka Clifford algebra) formalism just to see this result.

It is also interesting to apply the correspondence principle, to see how this equation reproduces results that may be more familiar in other forms. Therefore let’s take a step back and review the prosaic non-spacetime non-geometric version of Maxwell’s equation.

3 Vectors

We start by writing the Maxwell equations in terms of vector fields in three dimensions, namely:

$$\begin{aligned}
 \nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} \\
 \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
 c^2 \nabla \times \mathbf{B} &= \frac{\partial \mathbf{E}}{\partial t} + \frac{\mathbf{j}}{\epsilon_0} \\
 \nabla \cdot \mathbf{B} &= 0
 \end{aligned}
 \tag{2}$$

These equations have several deep symmetries. We can make some of the symmetries more apparent by making a few superficial changes. The reasons for this will be explained in moment.

$$\begin{aligned}
 \nabla \cdot \mathbf{E} &= \frac{1}{c\epsilon_0} c\rho \\
 \nabla \times c\mathbf{B} - \frac{\partial}{\partial ct} \mathbf{E} &= \frac{1}{c\epsilon_0} \mathbf{j} \\
 \nabla \cdot c\mathbf{B} &= 0 \\
 \nabla \times \mathbf{E} + \frac{\partial}{\partial ct} c\mathbf{B} &= 0
 \end{aligned}
 \tag{3}$$

These equations are invariant with respect to rotations in three dimensions. They are *manifestly* invariant, because they have been written in vector notation. We have not yet specified a basis for three-dimensional space, so if Alice uses a reference frame that is rotated relative to Bob's reference frame, [equation 3](#) not only means the same thing to both of them, it looks the same, verbatim.

In contrast, these equations have a relativistic invariance that is *not* manifest. The t coordinate appears explicitly. If Alice uses a reference frame that is moving relative to Bob's reference frame, they won't be able to agree on the value of t . For that matter, they won't be able to agree on the values of the \mathbf{E} -field and \mathbf{B} -field.

Of course the non-agreement about the coordinates and the non-agreement about the fields cancel in the end, so Alice and Bob eventually agree about what the equations predict will happen physically.

Therefore [equation 3](#) represents an *intermediate* level of sophistication: manifest invariance with respect to rotations, but non-manifest invariance with respect to **boosts**.

In passing from [equation 2](#) to [equation 3](#), we added factors of c in strategic places. This helps make the equations more manifestly symmetric. Specifically:

1. In every place where t appears, we have arranged things so that ct appears, rather than t alone. The rationale is that ct has the same dimensions as x , y , and z . To say the same thing another way, in spacetime, the partner to x , y , and z is not t but rather ct .
2. Similarly, the partner to \mathbf{j} is not ρ but rather $c\rho$. In spacetime, $c\rho$ represents a certain amount of charge that sits at one spatial location and flows toward the future, whereas \mathbf{j} represents charge flowing from one spatial location to another.
3. Last but not least, the proper partner for \mathbf{E} is not \mathbf{B} but rather $c\mathbf{B}$. In every place where \mathbf{B} appears, we have arranged things so the combination $c\mathbf{B}$ appears, rather than \mathbf{B} alone. This is just an exercise in algebraic re-arrangement, and does not change the meaning of the equations. The rationale is that $c\mathbf{B}$ has the same dimensions as \mathbf{E} , and arranging things this way makes the equations more manifestly symmetric. (There have been proposals from Gauss and others to consider $c\mathbf{B}$ to be “the” magnetic field, but we decline to do so, since that would depart from the conventional meaning of the terms.)

Some tangential remarks:

- The Maxwell equations are not very useful without the Lorentz force law, as discussed in [section 6.2](#).

- As always, $\mu_0\epsilon_0c^2 = 1$.
- These equations are true in any system of units, including SI among others. (In *some* systems of units, it may be possible to formulate less-general but more-compact equations, perhaps by setting $c = 1$ and/or setting $4\pi\epsilon_0 = 1$. These compact formulations are consistent with [equation 3](#). They are corollaries of [equation 3](#) ... and not vice versa. We will stick with the more general formulation in this document. It is slightly less elegant but considerably more practical. Also, converting from the more-general form to some less-general form is vastly easier than vice versa. For details on units etc., see [reference 1](#) and [reference 2](#).)

4 Components

We can construct an even-less-sophisticated expression by choosing a basis and writing out the components:

$$\begin{aligned}
 \nabla_i E_i &= \frac{1}{\epsilon_0} \rho \\
 \epsilon_{ijk} \nabla_j cB_k - \frac{\partial}{\partial ct} E_i &= \frac{1}{c\epsilon_0} j_i \\
 \nabla_i cB_i &= 0 \\
 \epsilon_{ijk} \nabla_j E_k + \frac{\partial}{\partial ct} cB_i &= 0
 \end{aligned}
 \tag{4}$$

See [section 10.1](#) for information about the notation used here.

Expressing things in components like this is sometimes convenient for calculations, but it conceals rotation-invariance. If Alice uses a reference frame that is rotated relative to Bob's, they won't be able to agree on what x_i means or what E_i means. Of course the rotation-invariance is still there; it has just become non-manifest.

5 Electromagnetism using Geometric Algebra

Geometric Algebra (also known as Clifford Algebra) has many advantages, as discussed in [section 8](#). It turns out we can write the Maxwell equations in the easy-to-remember form

$$\nabla F = \frac{1}{c\epsilon_0} J
 \tag{5}$$

which contains the entire meaning of the less-sophisticated version, [equation 3](#), as we shall demonstrate in a moment.

This expression has the advantage of being manifestly Lorentz invariant (including boosts as well as rotations). Contrast this with [equation 3](#) in which the Lorentz invariance is not manifest.

Overall, the best approach would be to solve practical problems by direct appeal to [equation 1](#). Some examples can be found in [section 11](#) and [reference 3](#).

However, that's not the main purpose of this document. Instead, we want to derive the less-sophisticated Maxwell equations ([equation 3](#)) starting from [equation 1](#). This can be considered a test or an application of the correspondence principle.

For starters, we need to establish the correspondence between the 3-dimensional electric current \mathbf{j} and the corresponding four-vector current J . That is,

$$J = c\rho\gamma_0 + j^k\gamma_k
 \tag{6}$$

where we have chosen a reference frame in which $\gamma_0, \gamma_1, \gamma_2,$ and γ_3 are the orthonormal basis vectors. In particular, γ_0 is the timelike basis vector. We see that ρ has to do with *continuity of flow* of charge in the *time* direction, just as the ordinary three-dimensional current \mathbf{j} represents flow in the spacelike directions. See [reference 4](#) for more about the idea of conservation and continuity of flow.

We also need to know how F is related to the old-fashioned fields \mathbf{E} and \mathbf{B} . In any particular frame,

$$F = (\mathbf{E} + ic\mathbf{B})\gamma_0 \quad (7)$$

where i is the unit pseudoscalar ([equation 45](#)). We can expand this as:

$$\begin{aligned} F &= \mathbf{E}\gamma_0 - c\mathbf{B}\gamma_1\gamma_2\gamma_3 \\ &= E^k\gamma_k\gamma_0 - cB^k\gamma_k\gamma_1\gamma_2\gamma_3 \end{aligned} \quad (8)$$

where E^k and B^k are the components of the usual electric field and magnetic field as measured in our chosen frame.

This equation has quite an interesting structure. It tells us we ought to view the electromagnetic field as a bivector. In any particular frame this bivector F has two contributions: one contribution is a bivector having one edge in the timelike direction, associated with \mathbf{E} , while the other contribution is a bivector having both edges in spacelike directions, associated with $c\mathbf{B}$.

We are making heavy use of the central feature of the Clifford Algebra, namely the ability to multiply vectors. This multiplication obeys the usual associative and distributive laws, but is not in general commutative.¹ In particular because our basis vectors γ_μ are orthogonal, each of them anticommutes with the others:

$$\gamma_\mu\gamma_\nu = -\gamma_\nu\gamma_\mu \quad \text{for all } \mu \neq \nu \quad (9)$$

and the normalization condition² in $D = 1 + 3$ requires a minus sign in the timelike component:

$$\gamma_0\gamma_0 = -1, \quad \gamma_1\gamma_1 = +1, \quad \gamma_2\gamma_2 = +1, \quad \gamma_3\gamma_3 = +1 \quad (10)$$

Now all we have to do is plug [equation 7](#) into [equation 1](#) and turn the crank.

There will be 12 terms involving E , because E has three components E^k and the derivative operator has four components ∇^μ . Similarly there will be 12 terms involving B .

$$\begin{aligned} \nabla F = & \quad +\nabla^0 E^1 \gamma_1 & +\nabla^1 E^1 \gamma_0 & -\nabla^2 E^1 \gamma_0 \gamma_1 \gamma_2 & +\nabla^3 E^1 \gamma_0 \gamma_3 \gamma_1 & \\ & +\nabla^0 E^2 \gamma_2 & +\nabla^1 E^2 \gamma_0 \gamma_1 \gamma_2 & +\nabla^2 E^2 \gamma_0 & -\nabla^3 E^2 \gamma_0 \gamma_2 \gamma_3 & \\ & +\nabla^0 E^3 \gamma_3 & -\nabla^1 E^3 \gamma_0 \gamma_3 \gamma_1 & +\nabla^2 E^3 \gamma_0 \gamma_2 \gamma_3 & +\nabla^3 E^3 \gamma_0 & \\ & -\nabla^0 cB^1 \gamma_0 \gamma_2 \gamma_3 & -\nabla^1 cB^1 \gamma_1 \gamma_2 \gamma_3 & -\nabla^2 cB^1 \gamma_3 & +\nabla^3 cB^1 \gamma_2 & \\ & -\nabla^0 cB^2 \gamma_0 \gamma_3 \gamma_1 & +\nabla^1 cB^2 \gamma_3 & -\nabla^2 cB^2 \gamma_1 \gamma_2 \gamma_3 & -\nabla^3 cB^2 \gamma_1 & \\ & -\nabla^0 cB^3 \gamma_0 \gamma_1 \gamma_2 & -\nabla^1 cB^3 \gamma_2 & +\nabla^2 cB^3 \gamma_1 & -\nabla^3 cB^3 \gamma_1 \gamma_2 \gamma_3 & \end{aligned} \quad (11)$$

Let's discuss what this means. We start with the nine terms highlighted in blue. The six terms involving cB are the components of $\nabla \times c\mathbf{B}$. Similarly, the three terms involving E are the components of $+\nabla^0 \mathbf{E}$, which is the same as $-(\partial/c\partial t)\mathbf{E}$. These terms each involve exactly one of the spacelike basis vectors ($\gamma_1, \gamma_2,$ and γ_3), so we are dealing with a plain old vector in $D = 3$ space. The RHS of the [equation 1](#) has a vector that matches

¹You may be familiar with matrix multiplication, which has many of the same axioms as Geometric Algebra, including the associative law, the distributive law, and non-commutative multiplication. But the analogy is not perfect: the product of two matrices is another matrix, whereas the geometric product of two vectors isn't another vector: it could be a scalar (force times distance = work) or a bivector (force times distance = torque) or perhaps a combination of the two – but it won't be a proper vector.

²Some other authors use the opposite convention, in which $\gamma_0\gamma_0 = +1$ and all the others are -1 . It doesn't make much difference; the physics works out the same using either convention. But the convention used here makes it slightly easier to see the correspondence with plain old $D = 3$ vectors.

this, namely the $D = 3$ current density. So the blue terms are telling us that $\nabla \times c\mathbf{B} - (\partial/c\partial t)\mathbf{E} = (1/c\epsilon_0)\mathbf{j}$, which agrees nicely with [equation 3](#).

Next, we consider the nine terms highlighted in red. The six terms involving E are the components of $\nabla \times \mathbf{E}$. Similarly, the three terms involving cB are the components of $-\nabla^0 c\mathbf{B}$, which is the same as $+(\partial/c\partial t)c\mathbf{B}$. These nine terms are all the trivectors with a projection in the timelike direction (γ_0). Since the RHS of [equation 1](#) doesn't have any trivector terms, we must conclude that these red terms add up to zero, that is, $\nabla \times \mathbf{E} + (\partial/c\partial t)c\mathbf{B} = 0$, which also agrees with [equation 3](#).

The three black terms involving E match up with the timelike piece of J and tell us that $\nabla \cdot \mathbf{E} = (1/\epsilon_0)\rho$. The three black terms involving cB tell us that $\nabla \cdot c\mathbf{B} = 0$.³

Let me say few words about how this was calculated. It really was quite mechanical, just following the formalism. Consider the term $+\nabla^2 cB^3 \gamma_1$ in the last row. We started from the expression ∇F which has two factors, so the term in question will have two factors, $\nabla^2 \gamma_2$ and $-cB^3 \gamma_3 \gamma_1 \gamma_2 \gamma_3$, which combine to make $-\nabla^2 \gamma_2 cB^3 \gamma_3 \gamma_1 \gamma_2 \gamma_3$. All we have to do is permute the γ vectors to get this into standard form. Pull the scalars to the front and permute the first two vectors using [equation 9](#) to get $+\nabla^2 cB^3 \gamma_3 \gamma_2 \gamma_1 \gamma_2 \gamma_3$. Permute again to get $-\nabla^2 cB^3 \gamma_3 \gamma_1 \gamma_2 \gamma_2 \gamma_3$ which reduces using [equation 10](#) to $-\nabla^2 cB^3 \gamma_3 \gamma_1 \gamma_3$. Then one more permutation and one more reduction and the job is done.

The only part that required making a decision was writing $\gamma_0 \gamma_3 \gamma_1$ in places where I could have written $-\gamma_0 \gamma_1 \gamma_3$. This is just cosmetic; it makes the signs fall into a nice pattern so it is easier to see the correspondence with the old-fashioned cross product. We can make this seem more elegant and less arbitrary if we say the rule is to write all pseudovectors using the basis $\{i\gamma_\mu \text{ for } \mu = 0, 1, 2, 3\}$, where i is the unit pseudoscalar ([equation 45](#)).

After the calculation was done, deciding how to color the terms took some judgment, but not much, because the terms naturally segregate as vectors and trivectors, spacelike and timelike.

6 Charge, Force, and Energy

6.1 Conservation of Charge

Preview: Our goal is to prove that charge is conserved, i.e. that $\nabla \cdot J = 0$. We are not going to assume conservation; we are going to prove that conservation is already guaranteed as a consequence of [equation 1](#), the Maxwell equation. We will do that by taking the divergence of both sides of the equation.

Background: We are going to need a mathematical lemma that says the divergence of the divergence of a bivector is always zero. To derive this, consider an arbitrary bivector W . We temporarily assume W is a simple blade, i.e. $W = a\gamma_5 \gamma_6$. Then the divergence is

$$\begin{aligned} \nabla \cdot W &= \nabla \cdot a\gamma_5 \gamma_6 \\ &= \langle \nabla a\gamma_5 \gamma_6 \rangle_1 \\ &= \nabla^5 \gamma_5 a\gamma_5 \gamma_6 + \nabla^6 \gamma_6 a\gamma_5 \gamma_6 \\ &= \nabla^5 a\gamma_6 - \nabla^6 a\gamma_5 \end{aligned} \tag{12}$$

where on the second line we have used the general rule that the dot product is the low-grade piece of the full geometric product. On the last line we have temporarily assumed that γ_5 and γ_6 are spacelike, but we shall see that this assumption is unnecessary.

³Magnetic monopoles would be described by a trivector (i.e. pseudovector) term on the RHS of [equation 1](#).

Let us now take the divergence of the divergence.

$$\begin{aligned}
\nabla \cdot (\nabla \cdot W) &= \nabla \cdot (\nabla^5 a \gamma_6 - \nabla^6 a \gamma_5) \\
&= \langle \nabla (\nabla^5 a \gamma_6 - \nabla^6 a \gamma_5) \rangle_0 \\
&= \nabla^6 \gamma_6 \nabla^5 a \gamma_6 - \nabla^5 \gamma_5 \nabla^6 a \gamma_5 \\
&= \nabla^6 \nabla^5 a - \nabla^5 \nabla^6 a \\
&= 0
\end{aligned} \tag{13}$$

On the last line we have used the fact that the various components of the gradient operator commute with each other.

We now lift the assumption that our basis vectors are timelike. You should verify that it doesn't really matter whether γ_5 and γ_6 are spacelike or timelike. Hint: a fuller calculation would give us:

$$\begin{aligned}
\nabla \cdot (\nabla \cdot W) &= \nabla^6 \nabla^5 a \gamma_5^2 \gamma_6^2 - \nabla^5 \nabla^6 a \gamma_5^2 \gamma_6^2 \\
&= 0
\end{aligned} \tag{14}$$

We now lift the assumption that W is a blade. By the distributive law, if $\nabla \cdot (\nabla \cdot W)$ is zero for any grade=2 blade, it is zero for any sum of such blades, i.e. for any bivector whatsoever. We conclude in all generality:

$$\nabla \cdot (\nabla \cdot W) = 0 \quad (\text{for any bivector } W) \tag{15}$$

As another lemma, for any bivector we can always write

$$\begin{aligned}
\nabla W &= \langle \nabla W \rangle_1 + \langle \nabla W \rangle_3 \\
&= \nabla \cdot W + \nabla \wedge W
\end{aligned} \tag{16}$$

This allows us to pick apart ∇F as follows:

$$\nabla F = \nabla \cdot F + \nabla \wedge F \tag{17a}$$

$$\nabla \cdot F = \frac{1}{c\epsilon_0} J \tag{17b}$$

$$\nabla \wedge F = \text{something} \quad (\text{in all generality}) \tag{17c}$$

$$= 0 \quad (\text{assuming no monopoles}) \tag{17d}$$

For the purposes of this section, all we need is [equation 17b](#). That is the grade=1 piece of the Maxwell equation. We do not need to assume the non-existence of monopoles. We do not need to know anything about the trivector piece of the Maxwell equation. We do not need [equation 17d](#) or even [equation 17c](#).

Using our lemma ([equation 15](#)), we can write

$$\begin{aligned}
\nabla \cdot J &= c\epsilon_0 \nabla \cdot (\nabla \cdot F) \\
&= 0
\end{aligned} \tag{18}$$

We are of course using the four-dimensional divergence. Zero divergence expresses the continuity of world-lines in spacetime. For an explanation of why this is the right way to express the idea of conservation in terms of continuity of flow, see [reference 4](#).

6.2 Lorentz Force Law

As remarked above, our theory of electromagnetism would be incomplete without the Lorentz force law.

The old-fashioned way of writing the Lorentz force law is:

$$\frac{\partial}{\partial t} \mathbf{p} = q(\mathbf{E} + \frac{\mathbf{v}}{c} \times c\mathbf{B}) \tag{19}$$

where \mathbf{p} is the momentum, q is the charge, and \mathbf{v} is the ordinary 3-dimensional velocity.

As with practically any equation involving cross products, [equation 19](#) can be improved by rewriting it using Geometric Algebra instead:

$$\frac{\partial}{\partial \tau} p = qu \cdot F \tag{20}$$

where τ is the proper time, $u = dx/d\tau$ is the 4-dimensional proper velocity,⁴ $p = mu$ is the momentum, and m is the invariant mass. Here p and u are vectors in $D = 1 + 3$ spacetime. This is the relativistically-correct generalization of equation 19.

Equation 20, unlike previous equations, involves a dot product. In particular, it involves the dot product of a vector with a bivector. Such things are not quite as easy to compute as the dot product between two vectors, but they are still reasonably easy to compute in terms of the geometric product. In general, the dot product is the lowest-grade part of the full geometric product, as discussed in reference 5. In the case of a vector dotted with a bivector, we have:

$$\begin{aligned} A \cdot (B \wedge C) &= \langle A(B \wedge C) \rangle_1 \\ &= \frac{1}{2} \langle ABC - ACB \rangle_1 \end{aligned} \quad (21)$$

That means we just form the geometric product and throw away everything but the grade=1 part. Another way of dealing with “vector dot bivector” is:

$$A \cdot (B \wedge C) = (A \cdot B)C - (A \cdot C)B \quad (22)$$

which can be considered a sort of “distributive law” for distributing the dot-operator over the wedge-operator. Equation 22 tells us that the product $A \cdot (B \wedge C)$ is a vector that lies in the plane spanned by B and C .

The following examples are useful for checking the validity of the foregoing equations:

$$\begin{aligned} \gamma_1 \cdot (\gamma_1 \wedge \gamma_1) &= 0 \\ \gamma_1 \cdot (\gamma_1 \wedge \gamma_2) &= \gamma_2 \\ \gamma_2 \cdot (\gamma_1 \wedge \gamma_2) &= -\gamma_1 \\ \gamma_3 \cdot (\gamma_1 \wedge \gamma_2) &= 0 \end{aligned} \quad (23)$$

To say the same thing in geometric (rather than algebraic) terms, you can visualize the product of a vector with a bivector as follows:

- Throw away the component of the vector perpendicular to the plane of the bivector. Keep the projection in the plane.
- The result (the dot product) will be in the plane and perpendicular to the projection. Its length will be the magnitude of the bivector times the magnitude of the projection.

An example of the Lorentz law in action is shown in figure 1, for the case of an electromagnetic field bivector (F) that is uniform in space, oriented purely in the plane of the paper. The cyclotron orbit shown in the figure corresponds to the motion of a positive test charge with some initial velocity, free of forces other than the indicated electromagnetic field.

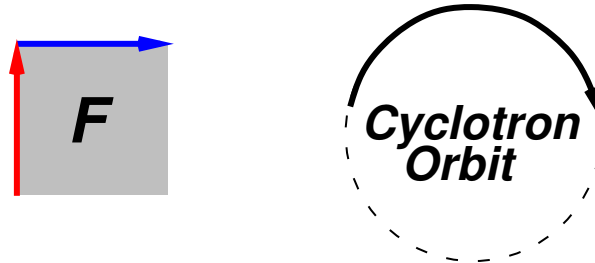


Figure 1: Lorentz Law

It is straightforward to understand this result. If the particle is moving in the direction of the red vector, it will experience a force in the blue direction. If the particle is moving in the blue direction, it will experience a force opposite to the red direction.

⁴The proper velocity $u = dx/d\tau$ is not to be confused with the coordinate velocity $v = dx/dt$. They're the same when they're small, but in general they differ by a factor of gamma. You could rewrite equation 20 in terms of t -derivatives rather than τ -derivatives, but it would be less elegant and less useful.

To summarize: The magnetic part of the Lorentz force law is super-easy to remember:

A field bivector in the plane of the paper
leads to a cyclotron orbit in the plane of the paper.

Motion perpendicular to the field bivector is unaffected by the field.

The foregoing applies if the field F is already expressed in modern terms, as a bivector. Now, in the spirit of this document, we re-examine the situation to exhibit the correspondence between the bivector idea and old-fashioned ideas such as the electric field vector and the magnetic field pseudovector.

The bivector shown in [figure 1](#) is purely spatial, so it must correspond to a magnetic field, with no electric field in our frame of reference. The magnetic field pseudovector is perpendicular to the paper, directed *out of* the paper. You can check using the right-hand force rule that the cyclotron orbit shown in [figure 1](#) is correct for a positive test charge moving in such a magnetic field.

It is amusing to check the general case, for any F that is known in terms of the old-fashioned electric field vector and magnetic field pseudovector, as in [equation 7](#) or [equation 8](#). As suggested by [equation 20](#), we should take the dot product of u with both sides of our expression for F . The correspondence principle suggests we should recover the old-fashioned 3-vector version of the force law, i.e. [equation 19](#). To carry out the dot product, we could just turn the crank ... but in fact we hardly need to do any work at all. The dot product in $u \cdot F$ uses a subset of the full geometric product uF , namely the plain vector (grade=1) terms. See [equation 18](#) in [reference 6](#). We can avoid some work, because uF has the same structure as ∇F – it’s just the geometric product of some vector with F – so we can just re-use [equation 11](#), replacing ∇ by u everywhere. Then we throw away all the trivector terms, and what remains is the dot product.

In the nonrelativistic limit, the timelike component of the velocity equals unity, plus negligible higher-order terms. So the blue terms in [equation 11](#) give us the usual Lorentz equation for the spacelike components of the momentum-change: $\mathbf{1E} + \mathbf{v} \times \mathbf{B}$.

The black terms involving E give us a bonus: They tell us the power (i.e. the rate of work, i.e. the time-derivative of the kinetic energy), namely $\mathbf{v} \cdot \mathbf{E}$.

6.3 Lagrangian Density

Let us consider the gorm of the electromagnetic field, namely $\text{gorm}(F) \equiv \langle FF^\sim \rangle_0$. You can readily verify that:

$$\langle FF^\sim \rangle_0 = (cB)^2 - E^2 \tag{24}$$

This is a scalar, a Lorentz-invariant scalar. It is useful in a number of ways, not least of which is the fact that $-\epsilon_0((cB)^2 - E^2)$ is the Lagrangian density for the electromagnetic field.

6.4 Poynting Vector

Let’s continue looking for energy-related expressions involving F . [Section 6.3](#) gives us hint as to where to look; the Lagrangian density is not “the” energy density, but it at least has dimensions of energy density.

We know from old-fashioned electromagnetism that there should be an energy density that goes like the square of the field strength. This tells us the amount of energy per unit volume. In old-fashioned terms, the energy density is $1/2 \epsilon_0 (E^2 + c^2 B^2)$.

There is also a Poynting vector, which tells us the amount of energy flowing per unit surface area, per unit time. In old-fashioned terms, it is $c \epsilon_0 \mathbf{E} \times c\mathbf{B}$.

So, without further motivation, we use 20/20 hindsight to assert that $F \gamma_0 F$ will be interesting. Following the spirit of this document, let's check that assertion by working out $F \gamma_0 F$ in terms of the old-fashioned E and B fields, and seeing what we get. We substitute for F using [equation 7](#) and turn the crank:

$$\begin{aligned}
F \gamma_0 F &= (E^k \gamma_k \gamma_0 - c B^j \gamma_j \gamma_1 \gamma_2 \gamma_3) \gamma_0 (E^k \gamma_k \gamma_0 - c B^j \gamma_j \gamma_1 \gamma_2 \gamma_3) \\
&= E^k \gamma_k \gamma_0 \gamma_0 E^j \gamma_j \gamma_0 \\
&\quad - E^k \gamma_k \gamma_0 \gamma_0 c B^j \gamma_j \gamma_1 \gamma_2 \gamma_3 - c B^k \gamma_k \gamma_1 \gamma_2 \gamma_3 \gamma_0 E^j \gamma_j \gamma_0 \\
&\quad + c B^k \gamma_k \gamma_1 \gamma_2 \gamma_3 \gamma_0 c B^j \gamma_j \gamma_1 \gamma_2 \gamma_3 \\
&= -E^k \gamma_k E^j \gamma_j \gamma_0 \\
&\quad + E^k \gamma_k c B^j \gamma_j \gamma_1 \gamma_2 \gamma_3 - c B^k \gamma_k \gamma_1 \gamma_2 \gamma_3 E^j \gamma_j \\
&\quad + c B^k \gamma_k \gamma_1 \gamma_2 \gamma_3 c B^j \gamma_j \gamma_1 \gamma_2 \gamma_3 \gamma_0 \\
&= -E^k \gamma_k E^j \gamma_j \gamma_0 \\
&\quad + E^k \gamma_k c B^j \gamma_j \gamma_1 \gamma_2 \gamma_3 - c B^j \gamma_j \gamma_1 \gamma_2 \gamma_3 E^k \gamma_k \\
&\quad - c B^k \gamma_k c B^j \gamma_j \gamma_0 \\
&= -E \cdot E \gamma_0 \\
&\quad + E^k c B^j (\gamma_k \gamma_j - \gamma_j \gamma_k) \gamma_1 \gamma_2 \gamma_3 \\
&\quad - c^2 B \cdot B \gamma_0 \\
&= -(E \cdot E + c^2 B \cdot B) \gamma_0 \\
&\quad - 2(E \times cB)^k \gamma_k
\end{aligned} \tag{25}$$

In going from the second line to the third line, we used the fact that $(\gamma_0)^2 = -1$. We also used the fact that $\gamma_0 \gamma_k = -\gamma_0 \gamma_k$ for all $k \in \{1, 2, 3\}$. On the other hand, $(\gamma_1 \gamma_2 \gamma_3) \gamma_k = +\gamma_k (\gamma_1 \gamma_2 \gamma_3)$. That is, when we commute γ_k across the three factors in $(\gamma_1 \gamma_2 \gamma_3)$, we pick up only two factors of -1 , not three, since for one of the factors the subscript on that factor will match the subscript k , and γ_k obviously commutes with itself.

In the next step, we used the fact that $(\gamma_1 \gamma_2 \gamma_3)^2 = -1$. We also changed some dummy indices.

So we see that we should be particularly interested in the quantity

$$\begin{aligned}
T(\gamma_0) &:= -^{1/2} \epsilon_0 F \gamma_0 F \\
&= \epsilon_0 \begin{bmatrix} (E^2 + c^2 B^2)/2 \\ (E \times cB)_1 \\ (E \times cB)_2 \\ (E \times cB)_3 \end{bmatrix}
\end{aligned} \tag{26}$$

The spacelike part of $T(\gamma_0)$ is the old-fashioned three-dimensional Poynting vector (apart from a missing factor of c), while the timelike component represents the corresponding energy density.

Although this $T(\gamma_0)$ -vector has four components, it is not a well-behaved Lorentz-covariant four-vector. It is actually just one column of a 4×4 object, namely the stress-energy tensor, T . Writing $T(\gamma_0)$ in terms of E and B (as in the second line of [equation 26](#)) only makes sense in the particular frame where E and B are defined. Also, if you want to connect $T(\gamma_0)$ to the Poynting vector in a given frame, γ_0 cannot be just any basis vector, but must be the 4-velocity of the frame itself, i.e. the unit vector in the time direction in the given frame.

More generally, the quantity

$$T(a) := -^{1/2} \epsilon_0 F a F \tag{27}$$

represents the flow of [energy, momentum] across the hypersurface perpendicular to the vector a . A more general way of looking at this is presented in [section 6.5](#).

6.5 Stress-Energy Tensor

The stress-energy tensor T for the electromagnetic field (in a vacuum) has the following matrix elements:

$$T_{\mu\nu} = F \gamma_\mu F \gamma_\nu \tag{28}$$

for any set of basis vectors $\{\gamma_\mu\}$. Equation 26 and equation 27 can be understood as special cases of equation 28.

7 Vector-ish Potential

7.1 The Basic Idea

In four dimensions, the electromagnetic field bivector F can always be written as the exterior derivative of a quasi-vector-ish potential A . Conversely, we can integrate the electromagnetic field to find the potential difference between point P and point Q .

$$\begin{aligned} F &= dA \\ &= \nabla \wedge A \end{aligned} \quad (29)$$

This implicitly defines what we mean by A . However, A is not uniquely defined, as discussed in section 7.3. Furthermore, even though A looks like it might be a four-vector, it's not.

7.2 $D = 3$ versus $D = 4$

- In electrostatics, in any chosen reference frame, the electric field can always be written as the derivative of some potential φ . This φ is called the scalar potential or the electrostatic potential.

Conversely, you can always integrate the electrostatic field to find the potential difference between any two points.

- In non-static situations in three dimensions, the electric field cannot in general be written as the derivative of any scalar potential. In particular, consider the electric field inside a betatron, or inside an ordinary transformer, or a ground loop, or a radio wave. The curl of E is nonzero:

$$\begin{aligned} \nabla \times E &\equiv *(\nabla \wedge E) && \text{[in D=3]} \\ &\neq 0 && \text{[when the magnetic field is changing]} \end{aligned} \quad (30)$$

That suffices to prove that:

$$E \neq \nabla\varphi \quad \text{for any } \varphi \quad \text{[when the magnetic field is changing]} \quad (31)$$

This tells us that any attempt to integrate E to find the scalar potential difference between point P and point Q will fail; the integral will depend on the path from P to Q , not just on the endpoints.

- Also, for what it's worth:

$$E \neq \nabla \wedge \varphi \quad \text{for any } \varphi \quad \text{[in D=3, when charge is present]} \quad (32)$$

- The electrostatic potential is somewhat related to the vector-ish potential as follows: It is superficially attractive to write:

$$\stackrel{\otimes}{\times} A = [\varphi/c, A_x, A_y, A_z] \quad \text{(the four-vector-ish potential)} \quad (33a)$$

$$\stackrel{\otimes}{\times} \varphi = cA_0 \quad (33b)$$

However, beware that equation 33a is a swindle, because it defines an object that is *not* a four-vector. It has four components, but that is not sufficient to make it a well-behaved 4-vector. It does not behave properly with respect to Lorentz transformations.

This is not tragic, because the potentials are not directly observable. The only thing that matters is the *difference* between two potentials, and that turns out to be well behaved, for the following reason:

Loosely speaking, if you start out with a vector-ish potential in a certain gauge and then change to a different reference frame, you get a vector-ish potential with the same physical meaning *in some other*

screwy gauge. If you try to calculate A by evaluating it in one frame and then boosting it into another frame, you will almost certainly get the wrong value for A . However, when you compute any physical observable, the gauge drops out, so you might end up with the right physics.

In particular, the key [equation 29](#) is OK. The electromagnetic field F is a well-behaved bivector. The exterior derivative on the RHS annihilates any and all gauge fields.

In any case, if you choose a particular reference frame and a particular gauge, then you can think of φ/c as being the timelike component of A .

At this point you should be asking yourself, how can $\nabla \wedge$ (field) be nonzero in three dimensions but zero in four dimensions? How does that not violate the correspondence principle? How does that not contradict the claim made in [reference 7](#) that Minkowski spacetime is very very similar to Euclidean space?

The answer is that when we switch from three dimensions to four, we redefine mean by “the” field, “the” potential, and “the” wedge product. In four dimensions, the exterior derivative of a vector has more terms. Invoking the correspondence principle, we can explain this in terms of the old-style E and B fields as follows: when we compute $\nabla \wedge F$, the time derivative of the B -component cancels the spatial derivatives of the E -component.

This is a trap for the unwary. Don’t let your experience with $D = 3$ poison your intuition about $D = 4$. Consider the contrast:

In $D = 3$ it is important to remember that “the field” (E) is not generally the derivative of any potential.	In $D = 4$, we can always write “the field” (F) as $F = dA$.
---	--

For some problems, there is a natural reference frame that has immense practical significance.	For some problems, the frame-independent space-time approach is simple, convenient, powerful, and elegant.
--	--

For example, if you are dealing with transformers or ground loops, you care a lot about the electric field in the frame of the device. The fact that this field cannot be written as the gradient of any potential is important. See [reference 8](#) for suggestions on how to visualize what’s going on.

7.3 Gauge Invariance

The vector-ish potential is implicitly defined by [equation 29](#). However, for any given field F , you don’t know whether the vector-ish potential is A or $A + \lambda'$, since we can write:

$$F = \nabla \wedge (A + \lambda') \tag{34}$$

for any vector field λ' such that

$$\nabla \wedge \lambda' = 0 \tag{35}$$

In particular, we can use the gradient of any scalar field λ :

$$\lambda' = \nabla \lambda \tag{36}$$

which is guaranteed to work since $\nabla \wedge \nabla$ (anything) is automatically zero. Beware of inconsistent terminology: Sometimes λ is called «the» gauge field, and sometimes λ' is called «the» gauge field.

7.4 The Maxwell Equation in terms of the Vector-ish Potential

The fact that we can write the electromagnetic field bivector as the derivative of a vector field is related to the fact that there are no trivector terms on the RHS of the Maxwell equation (equation 1). In particular, because ∇ is a vector, we can always write:

$$\nabla F = \begin{array}{ll} \nabla \cdot F & \text{(vector piece)} \\ + \nabla \wedge F & \text{(trivector piece)} \end{array} \quad (37)$$

Equation 37 is a mathematical identity, valid for any F you can think of. Applying it to the electromagnetic field in particular and plugging in equation 29 we obtain:

$$\nabla F = \begin{array}{ll} \nabla \cdot \nabla \wedge A & \text{(vector piece)} \\ + \nabla \wedge \nabla \wedge A & \text{(trivector piece)} \end{array} \quad (38)$$

So we could not write $F = \nabla \wedge A$ unless we already knew that $\nabla \wedge F$ was zero, since $\nabla \wedge \nabla \wedge A$ is automatically zero. Indeed $\nabla \wedge \nabla \wedge$ (anything) is automatically zero; see equation 17.

Combining these ideas, we see that another way of writing the Maxwell equation is:

$$\nabla \cdot \nabla \wedge A = \frac{1}{c\epsilon_0} J \quad (39)$$

or equivalently:

$$\nabla^2 A = \frac{1}{c\epsilon_0} J \quad (40)$$

where ∇^2 is called the d'Alembertian, or (equivalently) the four-dimensional Laplacian. It's the dot product of the derivative operator with itself.

Some references express the same idea using a different symbol:

$$\square^2 A = \frac{1}{c\epsilon_0} J \quad (41)$$

Beware that yet other references use plain unsquared \square to represent the d'Alembertian. The idea is that they reserve ∇^2 to represent the three-dimensional Laplacian, and use \square^2 to represent the four-dimensional generalization. However, in this document, we assume that all vectors are four-dimensional unless otherwise specified; for example, p is the four-momentum, A is the four-vector-ish potential, ∇ is the four-dimensional gradient, et cetera.

8 Geometric Algebra – General Remarks

8.1 Overview

Geometric Algebra has some tremendous advantages. It provides a unified view of inner products, outer products, $D = 2$ flatland, $D = 3$ space, $D = 1+3$ spacetime, vectors, tensors, complex numbers, quaternions, spinors, rotations, reflections, boosts, and more. This may sound too good to be true, but it actually works.

If you need an introduction to Geometric Algebra, please see reference 9, reference 10, and other references in section 12. Just as I did not include an introductory discussion of the divergence and curl operators in equation 3, I will not include an introductory discussion of Geometric Algebra here. There's no point in duplicating what's in the references. In particular, reference 10 discusses electromagnetism using $D = 3$ Clifford Algebra, which is easier to follow than the $D = 4$ discussion here, but the results are not as simple and elegant as equation 1. The calculation here, while not particularly difficult, does not pretend to be entirely elementary.

8.2 No Decorations

In Geometric Algebra, it traditional to not distinguish vectors using boldface or other decorations. This is appropriate, since the Clifford Algebra operates on *multivectors* and treats all multivectors on pretty much the same footing. Multivectors can be scalars, vectors, bivectors, pseudovectors, pseudoscalars — or linear combinations of the above.

8.3 No Cross Product

Observe that there is no cross-product operator in [equation 1](#) or [equation 20](#). That is good. Cross products are trouble. They don't exist in two dimensions, they are worse than useless in four dimensions, and aren't even 100% trustworthy in three dimensions. For example, consider a rotating object and its angular-momentum vector $\mathbf{r} \times \mathbf{p}$. If you look at the object in a mirror, the angular-momentum vector is reversed. You can't draw a picture of the rotating object and its angular-momentum vector and expect the picture to be invariant under reflections.

As far as I can tell, every physics formula involving a cross product can be improved by rewriting it using a wedge product instead.

For a rotating object, the cross product $\mathbf{r} \times \mathbf{p}$ is a vector oriented according to the axis of rotation, while the wedge product $r \wedge p$ is an area oriented according to the plane of rotation. The concept of “axis of rotation” is not portable to $D = 2$ or $D = 4$, but the concept of “plane of rotation” works fine in all dimensions.

If you think cross products are trouble, wait till you see Euler angles. They are only defined with respect to a particular basis. It's pathetic to represent rotations in a way that is not rotationally invariant. Geometric Algebra fixes this.

8.4 No Chirality

Note that Clifford Algebra does not require any right-hand rule. In [equation 10](#), the timelike vector is distinguished from the spacelike vector, but otherwise that equation and [equation 9](#) treat all the basis vectors on an equal footing; renaming or re-ordering them doesn't matter.

In $D = 3$ or $D = 1 + 3$ the unit pseudoscalar ([equation 45](#)) is chiral; that is, constructing it requires the right-hand rule. The axioms of Clifford Algebra sometimes permit but never require the construction of such a critter. The laws of electromagnetism are completely left/right symmetric. The magnetic term in [equation 7](#) contains B , which is chiral because it was defined via the old-fashioned cross product ... but the same term contains a factor of i which makes the overall expression left/right symmetric. It would be better to write the magnetic field as a bivector to begin with (as in [reference 3](#)), so the equations would make manifest the intrinsic left/right symmetry of the physical laws.

8.5 Multiple Approaches

There are at three different approaches to defining an F -like quantity as part of a geometric-algebra formulation of electromagnetism.

- The bivector + bivector approach, as used here, for example in [equation 7](#).
- The vector + axial vector approach, as used e.g. in [reference 10](#).
- The scalar + vector approach, as used in e.g. [reference 11](#).

Each approach is self-consistent, and most of the equations, such as [equation 1](#), are the same across all systems.

The advantage of the bivector + bivector approach is that it is “at home in spacetime” i.e. it treats x and t on the same footing, and treats B and E on the same footing (to the extent possible). It makes it easy and intuitive to draw bivector diagrams of the sort used in [reference 3](#).

9 Pitfalls to Avoid

9.1 Definition of Dot Product

You may be accustomed to expanding the dot product as

$$\mathbf{A} \cdot \mathbf{B} \text{ ??? } A_1B_1 + A_2B_2 + A_3B_3 \quad (42)$$

as if that were the definition of dot product ... but that is not the definition, and you’ll get the wrong answer if you try the corresponding thing in a non-Euclidean space, such as spacetime. So what you should do instead is to expand

$$\mathbf{A} = A^\mu \gamma_\mu = A^0 \gamma_0 + A^1 \gamma_1 + A^2 \gamma_2 + A^3 \gamma_3 \quad (43)$$

where the γ_μ are the basis vectors. Such an expansion is always legal. That is what defines the components A^μ . The superscripts on A label the components of \mathbf{A} ; they are not exponents. The subscripts on γ do not indicate components; they simply label which of the basis vectors we are talking about. It is possible but not particularly helpful to think of γ_0 as the zeroth component of some “vector of vectors”; in any case remember that γ_0 is a vector unto itself.

When you take the dot product $\mathbf{A} \cdot \mathbf{B}$, the expansion [equation 43](#) (and a similar expansion for \mathbf{B}) gives you sixteen terms, since the dot product distributes over addition in the usual way. The twelve off-diagonal terms vanish, since they involve things like $\gamma_1 \cdot \gamma_2$ and the basis vectors are mutually orthogonal. So we are left with

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} &= A^0 B^0 \gamma_0 \cdot \gamma_0 + A^1 B^1 \gamma_1 \cdot \gamma_1 + A^2 B^2 \gamma_2 \cdot \gamma_2 + A^3 B^3 \gamma_3 \cdot \gamma_3 \\ &= -A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3 \end{aligned} \quad (44)$$

where the term $A_0 B_0$ has picked up a minus sign, because γ_0^2 is -1 .

9.2 Unit Pseudoscalar

Another thing to watch out for when reading the Geometric Algebra literature concerns the use of the symbol i for the unit pseudoscalar:

$$i := \gamma_0 \gamma_1 \gamma_2 \gamma_3 \quad (45)$$

It’s nice to have a symbol for the unit pseudoscalar, and choosing i has some intriguing properties stemming from the fact that $i^2 = -1$, but there’s a pitfall: you may be tempted to treat i as a scalar, but it’s not. Scalars commute with everything, whereas this i anticommutes with vectors (and all odd-grade multivectors). This is insidious because in $D = 3$ the unit pseudoscalar commutes with everything. For these reasons we have mostly avoided using i in the main part of this note.

9.3 Exponents; Squared versus Norm Squared

Logical consistency requires that when using superscripts as exponents, they should denote simple powers:

$$\begin{aligned} M^2 &:= MM \\ M^3 &:= MMM \\ \text{etc.} \end{aligned} \tag{46}$$

for any multivector M . However, there is an unfortunate tendency for some authors to write M^2 when they mean MM^\sim where M^\sim is the *reverse* of M , formed by writing in reverse order all the vectors that make up M ; for example the reverse of [equation 7](#) tells us that $F^\sim = \gamma_0(\mathbf{E} + c\mathbf{B}i)$.

This is insidious because for scalars and vectors $MM^\sim = MM$; the distinction is only important for grade-2 objects and higher.

I recommend writing out MM^\sim whenever you mean MM^\sim . Many authors are tempted to come up with a shorthand for this – perhaps M^2 , $|M|^2$, or $\|M\|^2$ – but in my experience such things are much more trouble than they are worth. You need to be especially careful in the case where there are timelike vectors involved, since MM^\sim might well be negative. In such a case, any notation that suggests that MM^\sim is the square of anything is just asking for trouble.

A related and very important idea is the *gorm* of an object M , defined to be the scalar part of MM^\sim , i.e. $\langle MM^\sim \rangle_0$. (We saw a good physical example, namely the gorm of the electromagnetic field, in [section 6.3](#).)

9.4 Dot Product Not Necessarily Commutative

The dot product of a vector with a bivector is anticommutative, so be careful how you write the Lorentz force law:

$$u \cdot F = -F \cdot u \tag{47}$$

This is insidious because the dot product is commutative when acting on two vectors, or on “almost” any combination of multivectors. It is anticommutative only in cases where one of them has odd grade, and the other has a *larger* even grade. That is, in general,

$$A \cdot B = (-1)^{\min(r,s)|r-s|} B \cdot A \tag{48}$$

where r is the grade of A and s is the grade of B . This result may seem somewhat counterintuitive, but it is easy to prove; compare [equation 22](#) in [reference 6](#).

10 Additional Remarks

10.1 More about the Notation

1. You can think of x_1 as the X -direction, x_2 as the Y -direction, and x_3 as the Z -direction, but for most purposes we prefer the $\{x_1x_2x_3\}$ notation to the $\{XYZ\}$ notation. We will use x^0 and t almost interchangeably. There are unsettled issues about t versus ct , as discussed in [section 10.2](#).
2. We are using the Einstein summation convention, which calls for implied summation over repeated dummy indices, so that e.g.

$$\nabla_k E_k := \nabla_1 E_1 + \nabla_2 E_2 + \nabla_3 E_3 \tag{49}$$

Roman-letter indices run over the values 1,2,3 while Greek-letter indices run over the values 0,1,2,3.

3. By definition of what we mean by component, we can expand the ∇ operator in terms of its components

$$\nabla = \nabla^\mu \gamma_\mu \quad (50)$$

Naturally $\nabla^1 = (\partial/\partial x^1)$ and similarly for x^2 and x^3 , but you have to be careful of the minus sign in

$$\nabla^0 = -(\partial/\partial x^0) \quad (51)$$

Note that [equation 50](#) expresses a vector in terms of components times basis vectors, in contrast to [equation 51](#) which expresses only one component.

Here's how I like to remember where the minus sign goes. Imagine a scalar field $f(x)$, that is, some dimensionless scalar as a function of position. Positions are measured in inches. The length of the gradient vector ∇f is *not* measured in the same units as the length of position vectors. In fact it will have dimensions of *reciprocal* inches. So in this spirit we can write

$$\nabla = \frac{1}{\gamma_0} \frac{\partial}{\partial x^0} + \frac{1}{\gamma_1} \frac{\partial}{\partial x^1} + \frac{1}{\gamma_2} \frac{\partial}{\partial x^2} + \frac{1}{\gamma_3} \frac{\partial}{\partial x^3} \quad (52)$$

We can easily evaluate the reciprocals of the γ_μ vectors according to [equation 10](#), resulting in:

$$\nabla = -\gamma_0 \frac{\partial}{\partial x^0} + \gamma_1 \frac{\partial}{\partial x^1} + \gamma_2 \frac{\partial}{\partial x^2} + \gamma_3 \frac{\partial}{\partial x^3} \quad (53)$$

which has the crucial minus sign in front of the first term, and has the basis vectors in the numerators where they normally belong.

4. We make use of ϵ_{ijk} which is the Levi-Civita completely antisymmetric symbol: it equals +1 when ijk is a cyclic permutation of 123, equals -1 when ijk is an odd permutation of 123, and zero otherwise (namely when one of the indices is equal to one of the others).

10.2 Factors of c

In the field of electromagnetism, when we move beyond the introductory level to the intermediate level or the professional level, it is traditional to measure time in units of length, so that the speed of light is $c = 1$ in the chosen units.

This is a reasonable choice. However, it should remain a choice, not an obligation. We should be allowed to choose old-fashioned units of time if we wish. There are sometimes non-perverse reasons for choosing $c \neq 1$ – such as when checking the correspondence principle, as we do in this document.

This causes difficulties, because in the literature, some of the key formulas blithely assume $c = 1$, and if you want to go back and generalize the formulas so that they work even when $c \neq 1$, it is not always obvious how to do it. It's “usually” obvious, but not always.

In particular, consider the gorm of a vector (i.e. 4-vector) R that specifies position in spacetime. For any grade=1 vector R , the gorm is equal to the dot product, $R \cdot R$. For a position vector, we can write the gorm in terms of components, namely $-c^2 t^2 + x^2 + y^2 + z^2$. Leaving out the factor of c^2 would make this expression incorrect, indeed dimensionally unsound ... unless $c = 1$. Working backwards from the usual definition of dot product, that tells us that the position vector is $R = [ct, x, y, z]$ not simply $[t, x, y, z]$.

A similar argument tells us that the [energy, momentum] 4-vector is $[E, cp_x, cp_y, cp_z]$ not simply $[E, p_x, p_y, p_z]$.

The terminology in this area is trap for the unwary. You need to be careful to distinguish between “the time” (namely t) and “the timelike component of the position vector” (namely ct).

It is sometimes suggested that the dot product (i.e. the metric) be redefined to include explicit factors of c , which would permit the position vector be written as simply $[t, x, y, z]$. I do not recommend this, because although it is helpful for position 4-vectors, it is quite unhelpful for [energy, momentum] 4-vectors.

11 An Application: Plane Waves

11.1 Running Waves

As a modest application of [equation 1](#), let's try to find some solutions for it. In keeping with the spirit of this document, we will emphasize simplicity rather than elegance. We will formulate the problem in modern 4-dimensional terms, but in a way that maintains contact with old-style 3-dimensional frame-dependent concepts such as E and B . Also we will restrict attention to plane waves in free space.

In free space, there is no charge or current, so [equation 1](#) simplifies to:

$$\nabla F = 0 \tag{54}$$

We will write down a simple Ansatz ([equation 55](#)), and then show that it does in fact solve [equation 54](#).

$$\begin{aligned} F &= E(ky - vt)\gamma_1\gamma_0 + D(ky - vt)\gamma_2\gamma_0 - cB(ky - vt)\gamma_1\gamma_2 \\ &= E(\Phi)\gamma_1\gamma_0 + D(\Phi)\gamma_2\gamma_0 - cB(\Phi)\gamma_1\gamma_2 \end{aligned} \tag{55}$$

where F is the electromagnetic field bivector, E , D , and B are simple scalar functions of one scalar argument with as-yet undetermined physical significance, and Φ is the scalar phase:

$$\begin{aligned} \Phi &:= ky - vt \\ k &= +1 && \text{for propagation in the } +y \text{ direction} \\ k &= -1 && \text{for propagation in the } -y \text{ direction} \end{aligned} \tag{56}$$

Here is some motivation that may make this Ansatz less mysterious:

- When $k = +1$, writing the phase as a function of $y - vt$ is a standard way of creating something that keeps its shape while traveling in the $+y$ direction at velocity v . You can verify that whatever is happening at $[t, y] = [0, 0]$ is also happening at $[t, y] = [t_1, vt_1]$.
- By the same token, when $k = -1$, writing the phase as a function of $-y - vt$ creates something that moves in the $-y$ direction at velocity v .
- The RHS of [equation 55](#) is the most general bivector that can be written in 1+3 dimensional spacetime without mentioning γ_3 .

If we take a snapshot at any given time, we find that every plane parallel to the xz plane is a wavefront. That is to say, every such plane is a contour of constant phase. That's because it is, by construction, a contour of constant t and constant y . The phase depends on t and y , but not on x or z . This is what we would expect for a plane wave traveling in the y direction.

Using the chain rule we have:

$$\begin{aligned} \frac{\partial E}{\partial ct} &= \frac{dE}{d\Phi} \frac{\partial \Phi}{\partial t} \\ &= (-v/c) E' \\ \frac{\partial E}{\partial y} &= \frac{dE}{d\Phi} \frac{\partial \Phi}{\partial y} \\ &= kE' \end{aligned} \tag{57}$$

Corresponding statements can be made about B and D ... just apply the chain rule in the corresponding way. Here E' is pronounced “ E prime” and denotes the total derivative of E with respect to the scalar phase Φ .

Since there are three terms in [equation 55](#), taking the derivative gives us six terms; three for the timelike part of the gradient and three for the spacelike part. Plugging in and simplifying a bit gives us:

$$\begin{aligned}
\nabla F &= (-\partial/\partial ct)E\gamma_0\gamma_1\gamma_0 + (-\partial/\partial ct)D\gamma_0\gamma_2\gamma_0 - (-\partial/\partial ct)cB\gamma_0\gamma_1\gamma_2 \\
&\quad + (\partial/\partial y)E\gamma_2\gamma_1\gamma_0 + (\partial/\partial y)D\gamma_2\gamma_2\gamma_0 - (\partial/\partial y)cB\gamma_2\gamma_1\gamma_2 \\
&= (v/c)E'\gamma_0\gamma_1\gamma_0 + (v/c)D'\gamma_0\gamma_2\gamma_0 - (v/c)cB'\gamma_0\gamma_1\gamma_2 \\
&\quad + kE'\gamma_2\gamma_1\gamma_0 + kD'\gamma_2\gamma_2\gamma_0 - kcB'\gamma_2\gamma_1\gamma_2 \\
&= (v/c)E'\gamma_1 + (v/c)D'\gamma_2 - (v/c)cB'\gamma_0\gamma_1\gamma_2 \\
&\quad - kE'\gamma_0\gamma_1\gamma_2 - kD'\gamma_0 + kcB'\gamma_1
\end{aligned} \tag{58}$$

By [equation 54](#) we know this must equal zero. Each vector component must separately equal zero. Therefore:

$$\begin{aligned}
E' &= \frac{-v}{kc}cB' && \text{from the trivector part} \\
D' &= 0 && \text{from the } \gamma_0 \text{ (timelike) part} \\
cB' &= \frac{-v}{kc}E' && \text{from the } \gamma_1 \text{ (spacelike) part}
\end{aligned} \tag{59}$$

For additional follow-up on these results, see [section 11.2](#). For now, let's combine these results so as to obtain a consistency requirement for E' :

$$E' = (v^2/c^2)E' \tag{60}$$

where we have used the fact that $k^2 = 1$.

The first thing that we learn from the [equation 60](#) is that the electromagnetic plane wave in free space must propagate at speed $|v| = c$. This is an unavoidable consequence of the Maxwell equation in free space, [equation 54](#).

The second thing that we learn is that for any wave propagating at the required speed, the wavefunction can have any shape whatsoever, so long as it is differentiable function of its argument, i.e. a differentiable function of the phase Φ . It must be emphasized that we have not assumed that E is sinusoidal or even periodic. Any function $E(\Phi)$ you can think of, so long as it is differentiable, is an acceptable wavefunction for a plane wave in free space. Even an isolated blip, such as shown in [figure 2](#), can be a solution to [equation 54](#). The blip is moving left-to-right at the speed of light; the figure shows only a snapshot taken at time $t = 0$.

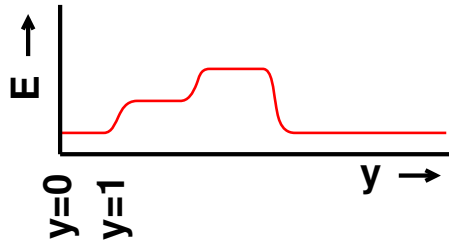


Figure 2: Snapshot of an Isolated Blip

The third thing we learn from [equation 60](#) in conjunction with [equation 59](#) is that once we have chosen E , then cB is constrained by [equation 59](#). That is, at every point in spacetime, $E = -kcB + g$, where g is some constant of integration. This g is not very interesting. It is constant across all of space and time, and represents some uniform, non-propagating background field. It has no effect on the propagating wave; the wave just propagates past it.

This completes the task of finding *some* solution.

Let's see if we can find a few more solutions.

First of all, we know the Maxwell equations are invariant under spacelike rotations, so we know there must exist plane waves propagating in any direction, not just the y direction. Any rotated version of our solution is another solution.

Secondly, you can easily verify that the factor of γ_1 in [equation 55](#) did not play any important role in the calculation; mostly it just went along for the ride. We could easily replace it with γ_3 and thereby obtain another solution, propagating in the same direction as the previous solution, but linearly independent of it. This phenomenon is called *polarization*. The Ansatz in [equation 55](#) is polarized in the γ_1 direction. You can verify that the polarization vector must be transverse to the direction of propagation; otherwise [equation 55](#) does not work as a solution to [equation 54](#).

We won't prove it, but we assert that we now have all the ingredients needed to construct the most general solution for plane waves in free space: first, pick a direction of propagation. Then choose a basis for the polarization vector, i.e. two unit vectors in the plane perpendicular to the direction of propagation. Then think of two arbitrary differentiable functions of phase, one for each component of the polarization vector. Finally, take arbitrary superpositions of all the above.

Tangential remark: Even though the Ansatz in [equation 55](#) contains three terms, the fact that $E = kcB$ and $D = 0$ means it can be written as a single blade, i.e. a bivector that is simply the product of two vectors. Specifically:

$$F = E(\Phi) \gamma_1 (\gamma_0 - k\gamma_2) \quad (61)$$

The structure here, and for any running plane wave, is simple. There are three factors: a scalar function $E(\Phi)$ that specifies the shape of the wave, times a spacelike vector that represents the polarization, times a null vector that represents the direction of propagation.

The general electromagnetic plane wave is not a single blade, but it can be written as a sum of blades of this form. Even more generally, there are lots of waves that are not plane waves.

11.2 Running Wave Phase Relationships

As noted in [section 11.1](#), there is a strict correspondence between the electric part and the magnetic part in an electromagnetic running plane wave. For a blip (or anything else) running left to right

$$E' = -cB' \quad \text{pointwise everywhere in space and time} \quad (62)$$

This is sometimes expressed by saying the E field and the cB field are “in phase.” (Such an expression makes more sense for sinusoidal waves than for blips.)

Meanwhile, for a blip (or anything else) running right to left,

$$E' = +cB' \quad \text{pointwise everywhere in space and time} \quad (63)$$

That is, once again there is a strict relationship between E and cB ... but the relationship in [equation 63](#) is diametrically opposite to the relationship in [equation 62](#). One of them is 180 degrees out of phase with the other.

If you consider the superposition of a left-running blip and a right-running blip, the whole notion of “phase relationship” goes out the window. You can have places where E is zero but cB is not, or vice versa, or anything you like, and the local relationship between E and cB will be wildly changing as a function of space and time. A particular type of superposition is considered in [section 11.3](#).

11.3 Standing Wave Phase Relationships

A standing wave can be viewed as the superposition of equal-and-opposite running waves. In particular, let's start with the sinusoidal waves

$$\begin{aligned}
 E_1 &= \cos(ct - y) \\
 E_2 &= \cos(ct + y) \\
 E &= E_1 + E_2 \\
 cB_1 &= -E_1 \\
 cB_2 &= +E_2 \\
 cB &= cB_1 + cB_2 \\
 &= -E_1 + E_2
 \end{aligned} \tag{64}$$

At any particular location y , the wave is a sinusoidal function of time. Choosing a different location just changes the phase. Let's apply the trigonometric sum-of-angles identity:

$$\begin{aligned}
 E_1 &= \cos(ct) \cos(y) + \sin(ct) \sin(y) \\
 E_2 &= \cos(ct) \cos(y) - \sin(ct) \sin(y) \\
 E &= E_1 + E_2 \\
 E &= 2 \cos(ct) \cos(y) \\
 cB &= -E_1 + E_2 \\
 &= -2 \sin(ct) \sin(y)
 \end{aligned} \tag{65}$$

So, as advertised above, we see that at most locations – i.e. any location where $\cos(y)$ and $\sin(y)$ are both nonzero – the E -field and the B -field are 90 degrees out of phase for a standing wave. (They are in phase for a running wave, as discussed in [section 11.2](#).)

11.4 Spacetime Picture

This section is restricted to the case where $k = +1$; that is, the wave is propagating in the $+y$ direction. Also we assume the constant of integration g is zero. Therefore $E = cB$ everywhere.

The blip we saw in [figure 2](#) is portrayed again in [figure 3](#). The former portrayed two variables, namely E versus y (at constant t). The latter portrays three variables, namely t , y , and E . The value of E is represented by the closeness of the flux lines. You can see that in the front half of the blip (larger y values) the E field is twice as large as in the back half of the blip.

The fact that $E = cB$ corresponds to the fact that, at each and every point in spacetime, the number of flux lines per unit distance in the timelike direction is equal to the number of flux lines per unit distance in the spacelike direction. An example of this is portrayed by the two small blue arrows in the figure. Not only does each arrow cross the same number of flux lines, it crosses the same flux lines.

You can see that this is a direct consequence of the geometry of spacetime, and the fact that the wave is propagating with velocity $v = c$.

As shown by the purple lines, contours of constant phase run from southwest to northeast. Phase increases toward the south and east. Phase increasing to the south corresponds to temporal period, and phase increasing to the east corresponds to spatial period i.e. wavelength. Note that any attempt to measure period or wavelength is utterly frame-dependent. Some properties of the wave (such as the total number of cycles) are frame-independent, but other properties (such as period, frequency, wavelength, and wavenumber) are necessarily frame-dependent.

In [figure 3](#), the x and z directions are not visible. If we made a more complicated diagram, from a different perspective, the electromagnetic field bivector F would be represented by *tubes*. The magnitude of F corresponds to the number of tubes per unit area.

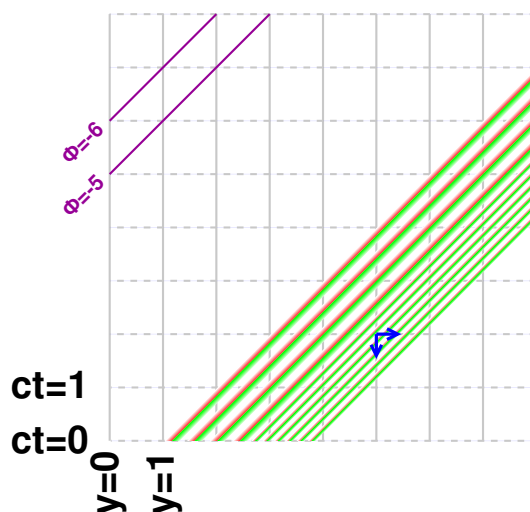


Figure 3: Radiation : Flux Lines in Spacetime

12 References

1. Wikipedia article: “Centimeter gram second system of units”
<http://en.wikipedia.org/wiki/CGS>
2. W. E. Baylis and G. W. F. Drake,
 “Units and Constants”
 in **Atomic Molecular & Optical Physics Handbook** AIP 1996
<http://www.atomwave.org/rmparticle/ao%20refs/aifm%20refs%20sorted%20by%20topic/other%20atom%20optics%20>
3. John Denker,
 “The Magnetic Field Bivector of a Long Straight Wire”
www.av8n.com/physics/straight-wire.htm
4. John Denker
 “Conservation as related to Continuity and Constancy”
www.av8n.com/physics/conservation-continuity.htm
5. John Denker,
 “Introduction to Clifford Algebra”
www.av8n.com/physics/clifford-intro.htm
6. Richard E. Harke,
 “An Introduction to the Mathematics of the Space-Time Algebra”
<http://www.harke.org/ps/intro.ps.gz>
7. John Denker,
 “Welcome to Spacetime”
www.av8n.com/physics/spacetime-welcome.htm
8. John Denker,
 “Visualizing A Field that is Not the Gradient of Any Potential”
www.av8n.com/physics/non-grady.htm

9. Stephen Gull, Anthony Lasenby, and Chris Doran,
“The Geometric Algebra of Spacetime”
<http://www.mrao.cam.ac.uk/~clifford/introduction/intro/intro.html>
10. David Hestenes,
“Oersted Medal Lecture 2002: Reforming the Mathematical Language of Physics”
Abstract: <http://geocalc.clas.asu.edu/html/Overview.html> Full paper:
<http://geocalc.clas.asu.edu/pdf/OerstedMedalLecture.pdf>
11. W. E. Baylis,
Electrodynamics, A Modern Geometric Approach
Birkhäuser, Boston (1999).