<table>
<thead>
<tr>
<th>Title</th>
<th>Geometric Algebra for Mathematics and Physics: A Unified Language (Innovative Teaching of Mathematics with Geometric Algebra)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Hestenes, David</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 (2004), 1378: 7-32</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2004-05</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/25643">http://hdl.handle.net/2433/25643</a></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>

Kyoto University
Geometric Algebra for Mathematics and Physics
- a Unified Language

Department of Physics and Astronomy,
Arizona State University, Tempe, Arizona 85287-1504
David Hestenes

Abstract

Physics and other applications of mathematics employ a miscellaneous assortment of mathematical tools in ways that contribute to a fragmentation of knowledge. We can do better! Research on the design and use of mathematical systems provides a guide for designing a unified mathematical language for the whole of physics that facilitates learning and enhances insight. The result of developments over several decades is a comprehensive language called Geometric Algebra with wide applications to physics and engineering. This lecture is an introduction to Geometric Algebra with the goal of incorporating it into the math/physics curriculum.

I. Introduction

The main subject of my lecture is a constructive critique of the mathematical language used in physics with an introduction to a unified language that has been developed over the last forty years to replace it. The generic name for that language is Geometric Algebra (GA). The material is developed in sufficient detail to be useful in instruction and research and to provide an entrée to the published literature.

After explaining the utter simplicity of the GA grammar in Section III, I explicate the following unique features of the mathematical language:

(1) GA seamlessly integrates the properties of vectors and complex numbers to enable a completely coordinate-free treatment of 2D physics.

(2) GA articulates seamlessly with standard vector algebra to enable easy contact with standard literature and mathematical methods.

(3) GA reduces "grad, div, curl and all that" to a single vector derivative that, among other things, combines the standard set of four Maxwell equations into a single equation and provides new methods to solve it.

Moreover, the GA formulation of spinors facilitates the treatment of rotations and rotational dynamics in both classical and quantum mechanics without co-
ordinates or matrices. GA provides fresh insights into the geometric structure of quantum mechanics with implications for its physical interpretation.19–30 All of this generalizes smoothly to a completely coordinate-free language for spacetime physics and general relativity.1–3,36,37 The development of GA has been a central theme of my own research in theoretical physics and mathematics.

II. Mathematics for Modeling Physical Reality

Mathematics is taken for granted in the physics curriculum—a body of immutable truths to be assimilated and applied. The profound influence of mathematics on our conceptions of the physical world is never analyzed. The possibility that mathematical tools used today were invented to solve problems in the past and might not be well suited for current problems is never considered.

One does not have to go very deeply into the history of physics to discover the profound influence of mathematical invention. Two famous examples will suffice to make the point: The invention of analytic geometry and calculus was essential to Newton’s creation of classical mechanics.4 The invention of tensor analysis was essential to Einstein’s creation of the General Theory of Relativity.

Note my use of the terms “invention” and “creation” where others might have used the term “discovery.” This conforms to the epistemological stance of Modeling Theory4,6–8 and Einstein himself, who asserted that scientific theories “cannot be extracted from experience, but must be freely invented.”9

The point I wish to make by citing these two examples is that without essential mathematical concepts the two theories would have been literally inconceivable. The mathematical modeling tools we employ at once extend and limit our ability to conceive the world. Limitations of mathematics are evident in the fact that the analytic geometry that provides the foundation for classical mechanics is insufficient for General Relativity. This should alert one to the possibility of other conceptual limits in the applications of mathematics.

Since Newton’s day a variety of different symbolic systems have been invented to address problems in different contexts. Figure 1 lists nine such systems in use today. Few scientists are proficient with all of them, but each system has advantages over the others in some application domain. For example, for applications to rotations, quaternions are demonstrably more efficient than the vectorial and matrix methods taught in standard linear algebra courses. The difference hardly matters in the world of academic exercises, but in the aerospace industry, for instance, where rotations are bread and butter, engineers opt for quaternions.

Each of the mathematical systems in Fig. 1 incorporates some aspect of geometry. Taken together, they constitute a highly redundant system of multiple representations for geometric concepts that are essential in physics and other applications of mathematics. As mathematical language, this Babel of mathematical tongues has the following defects:

1. Limited access. Scientific ideas, methods and results are distributed broadly across these diverse mathematical systems. Since most scientists are
proficient with only a few of the systems, their access to knowledge formulated in other systems is limited or denied. Of course, this language barrier is even greater for students.

2. Wasteful redundancy. In many cases, the same information is represented in several different systems, but one of them is invariably better suited than the others for a given application. For example, Goldstein's textbook on mechanics gives three different ways to represent rotations: coordinate matrices, vectors and Pauli spin matrices. The costs in time and effort for translation between these representations are considerable.

3. Deficient integration. The collection of systems in Fig. 1 is not an integrated mathematical structure. This is especially awkward in problems that call for the special features of two or more systems. For example, vector algebra and matrices are often awkwardly combined in rigid body mechanics, while Pauli matrices are used to express equivalent relations in quantum mechanics.

4. Hidden structure. Relations among scientific concepts represented in different symbolic systems are difficult to recognize and exploit.

5. Reduced information density. The density of information about nature is reduced by distributing it over several different symbolic systems.

Evidently elimination of these defects will make physics (and other scientific disciplines) easier to learn and apply. A clue as to how that might be done lies in recognizing that the various symbolic systems derive geometric interpretations from a common coherent core of geometric concepts. This suggests that one can create a unified mathematical language for physics (and thus for a large portion of mathematics and its applications) by designing it to incorporate an optimal representation of geometric concepts. In fact, Hermann Grassmann recognized this possibility and took it a long way more than 150 years ago. However, his program to unify mathematics was forgotten and his mathematical ideas...
were dispersed, though many of them reappeared in the several systems of Fig. 1. A century later the program was reborn, with the harvest of a century of mathematics and physics to enrich it. This has been the central focus of my own scientific research.

Creating a unified geometric language for physics and mathematics is a problem in the design of mathematical systems. Here are some general criteria that I have applied to the design of Geometric Algebra as a solution to that problem:

1. **Optimal algebraic encoding** of the basic geometric concepts: magnitude, direction, sense (or orientation) and dimension.

2. **Coordinate-free methods** to formulate and solve basic equations.

3. **Optimal uniformity** of method across various domains (like as classical, quantum and relativistic theories in physics) to make common structures as explicit as possible.

4. **Smooth articulation** with widely used alternative systems (Fig. 1) to facilitate access and transfer of information.

5. **Optimal computational efficiency.** The unified system must be at least as efficient as any alternative system in every application.

Obviously, these design criteria ensure built-in benefits of the unified language. In implementing the criteria I deliberately sought out the best available mathematical ideas and conventions. I found that it was frequently necessary to modify the mathematics to simplify and clarify the physics.

In the development of any scientific theory, a major task for theorists is to construct a mathematical language that optimizes expression of the key ideas and consequences of the theory. Although existing mathematics should be consulted in this endeavor, it should not be incorporated without critically evaluating its suitability. I might add that the process also works in reverse. Modification of mathematics for the purposes of other sciences serves as a stimulus for further development of mathematics. There are many examples of this effect in the history of physics.

Perhaps the most convincing evidence for validity of a new scientific theory is successful prediction of a surprising new phenomenon. Similarly, the most impressive benefits of Geometric Algebra arise from surprising new insights into the structure of physics and other sciences. 38–43

The following Sections survey the elements of Geometric Algebra and its application. Many details and derivations are omitted, as they are available elsewhere. The emphasis is on highlighting the unique advantages of Geometric Algebra as a unified mathematical language.

### III. Understanding Vectors

A recent study on the use of vectors by introductory physics students summarized the conclusions in two words: "**vector avoidance!**" 11,15 I maintain that the origin of this serious problem lies not so much in pedagogy as in the mathematics. The fundamental geometric concept of a vector as a directed magnitude is not adequately represented in standard mathematics. The basic definitions of
vector addition and scalar multiplication are essential to the vector concept but not sufficient. To complete the vector concept we need multiplication rules that enable us to compare directions and magnitudes of different vectors.

A. The Geometric Product

I take the standard concept of a real vector space for granted and define the geometric product \( \mathbf{a} \mathbf{b} \) for vectors \( \mathbf{a}, \mathbf{b}, \mathbf{c} \) by the following rules:

\[
\begin{align*}
(\mathbf{a} \mathbf{b}) \mathbf{c} &= \mathbf{a} (\mathbf{b} \mathbf{c}), \\
\mathbf{a} (\mathbf{b} + \mathbf{c}) &= \mathbf{a} \mathbf{b} + \mathbf{a} \mathbf{c}, \\
(\mathbf{b} + \mathbf{c}) \mathbf{a} &= \mathbf{b} \mathbf{a} + \mathbf{c} \mathbf{a}, \\
\mathbf{a}^2 &= |\mathbf{a}|^2.
\end{align*}
\]

where \(|\mathbf{a}|\) is a positive scalar called the magnitude of \(\mathbf{a}\), and \(|\mathbf{a}| = 0\) implies that \(\mathbf{a} = 0\).

All of these rules should be familiar from ordinary scalar algebra. The main difference is absence of a commutative rule. Consequently, left and right distributive rules must be postulated separately. The contraction rule (4) is peculiar to geometric algebra and distinguishes it from all other associative algebras. But even this is familiar from ordinary scalar algebra as the relation of a signed number to its magnitude.

From the geometric product \(\mathbf{a} \mathbf{b}\) we can define two new products, a symmetric inner product

\[
\mathbf{a} \cdot \mathbf{b} = \frac{1}{2}(\mathbf{a} \mathbf{b} + \mathbf{b} \mathbf{a}) = \mathbf{b} \cdot \mathbf{a},
\]

and an antisymmetric outer product

\[
\mathbf{a} \wedge \mathbf{b} = \frac{1}{2}(\mathbf{a} \mathbf{b} - \mathbf{b} \mathbf{a}) = - \mathbf{b} \wedge \mathbf{a}.
\]

Therefore, the geometric product has the canonical decomposition

\[
\mathbf{a} \mathbf{b} = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b}.
\]

From the contraction rule (4) it is easy to prove that \(\mathbf{a} \cdot \mathbf{b}\) is scalar-valued, so it can be identified with the standard Euclidean inner product.

The geometric significance of the outer product \(\mathbf{a} \wedge \mathbf{b}\) should also be familiar from the standard vector cross product \(\mathbf{a} \times \mathbf{b}\). The quantity \(\mathbf{a} \wedge \mathbf{b}\) is called a bivector, and it can be interpreted geometrically as an oriented plane segment, as shown in Fig. 2. It differs from \(\mathbf{a} \times \mathbf{b}\) in being intrinsic to the plane containing \(\mathbf{a}\) and \(\mathbf{b}\), independent of the dimension of any vector space in which the plane lies.

From the geometric interpretations of the inner and outer products, we can infer an interpretation of the geometric product from extreme cases. For orthogonal vectors, we have from (5)

\[
\mathbf{a} \cdot \mathbf{b} = 0 \quad \iff \quad \mathbf{a} \mathbf{b} = - \mathbf{b} \mathbf{a}.
\]
Fig. 2. Bivectors $a \wedge b$ and $b \wedge a$ represent plane segments of opposite orientation as specified by a “parallelogram rule” for drawing the segments.

On the other hand, collinear vectors determine a parallelogram with vanishing area (Fig. 2), so from (6) we have

$$a \wedge b = 0 \iff ab = ba.$$  \hspace{1cm} (9)

Thus, the geometric product $ab$ provides a measure of the relative direction of the vectors. Commutativity means that the vectors are collinear. Anticommutativity means that they are orthogonal. Multiplication can be reduced to these extreme cases by introducing an orthonormal basis.

**B. Basis and Bivectors**

For an orthonormal set of vectors $\{\sigma_1, \sigma_2, \ldots\}$, the multiplicative properties can be summarized by putting (5) in the form

$$\sigma_i \cdot \sigma_j = \frac{1}{2}(\sigma_i \sigma_j + \sigma_j \sigma_i) = \delta_{ij}$$  \hspace{1cm} (10)

where $\delta_{ij}$ is the usual Kroenecker delta. This relation applies to a Euclidean vector of any dimension, though for the moment we focus on the 2D case.

A unit bivector $i$ for the plane containing vectors $\sigma_1$ and $\sigma_2$ is determined by the product

$$i = \sigma_1 \sigma_2 = \sigma_1 \wedge \sigma_2 = -\sigma_2 \sigma_1$$  \hspace{1cm} (11)

The suggestive symbol $i$ has been chosen because by squaring (11) we find that

$$i^2 = -1$$  \hspace{1cm} (12)

Thus, $i$ is a truly geometric $\sqrt{-1}$. We shall see that there are others.

From (11) we also find that

$$\sigma_2 = \sigma_1 i = -i \sigma_1 \quad \text{and} \quad \sigma_1 = i \sigma_2.$$  \hspace{1cm} (13)

In words, multiplication by $i$ rotates the vectors through a right angle. It follows that $i$ rotates every vector in the plane in the same way. More generally, it follows that every unit bivector $i$ satisfies (12) and determines a unique plane in Euclidean space. Each $i$ has two complementary geometric interpretations: It represents a unique oriented area for the plane, and, as an operator, it represents an oriented right angle rotation in the plane.
C. Vectors and Complex Numbers

Assigning a geometric interpretation to the geometric product is more subtle than interpreting inner and outer products — so subtle, in fact, that the appropriate assignment has been generally overlooked until recently. The product of any pair of unit vectors \( \mathbf{a} \), \( \mathbf{b} \) generates a new kind of entity \( U \) called a rotor, as expressed by the equation

\[
U = \mathbf{a} \mathbf{b}. \tag{14}
\]

The relative direction of the two vectors is completely characterized by the directed arc that relates them (Fig. 3), so we can interpret \( U \) as representing that arc. The name "rotor" is justified by the fact that \( U \) rotates \( \mathbf{a} \) and \( \mathbf{b} \) into each other, as shown by multiplying (14) by vectors to get

\[
\mathbf{b} = \mathbf{a} U \quad \text{and} \quad \mathbf{a} = U \mathbf{b}. \tag{15}
\]

Further insight is obtained by noting that

\[
\mathbf{a} \cdot \mathbf{b} = \cos \theta \quad \text{and} \quad \mathbf{a} \wedge \mathbf{b} = \mathbf{i} \sin \theta, \tag{16}
\]

where \( \theta \) is the angle from \( \mathbf{a} \) to \( \mathbf{b} \). Accordingly, with the angle dependence made explicit, the decomposition (7) enables us to write (14) in the form

\[
U_\theta = \cos \theta + \mathbf{i} \sin \theta = e^{i \theta}. \tag{17}
\]

It follows that multiplication by \( U_\theta \), as in (15), will rotate any vector in the \( \mathbf{i} \)-plane through the angle \( \theta \). This tells us that we should interpret \( U_\theta \) as a directed arc of fixed length that can be rotated at will on the unit circle, just as we interpret a vector \( \mathbf{a} \) as a directed line segment that can be translated at will without changing its length or direction (Fig. 4).

![Fig. 3](https://via.placeholder.com/150)

**Fig. 3.** A pair of unit vectors \( \mathbf{a}, \mathbf{b} \) determine a directed arc on the unit circle that represents their product \( U = \mathbf{a} \mathbf{b} \). The length of the arc is (radian measure of) the angle \( \theta \) between the vectors.

With rotors, the composition of 2D rotations is expressed by the rotor product

\[
U_\theta U_\varphi = U_{\theta + \varphi}; \tag{18}
\]

and depicted geometrically in Fig. 5 as addition of directed arcs.
Fig. 4. All directed arcs with equivalent angles are represented by a single rotor $U_\theta$, just as line segments with the same length and direction are represented by a single vector $a$.

Fig. 5. The composition of 2D rotations is represented algebraically by the product of rotors and depicted geometrically by addition of directed arcs.

The generalization of all this should be obvious. We can always interpret the product $ab$ algebraically as a complex number

$$z = \lambda U = \lambda e^{i\theta} = ab,$$

(19)

with modulus $|z| = \lambda = \|a\|\|b\|$. And we can interpret $z$ geometrically as a directed arc on a circle of radius $|z|$ (Fig. 6). It might be surprising that this geometric interpretation never appears in standard books on complex variables. Be that as it may, the value of the interpretation is greatly enhanced by its use in geometric algebra.

Fig. 6. A complex number $z = \lambda U$ with modulus $\lambda$ and angle $\theta$ can be interpreted as a directed arc on a circle of radius $\lambda$. Its conjugate $z^\dagger = \lambda U^\dagger$ represents an arc with opposite orientation.
The connection to vectors via (19) removes a lot of the mystery from complex numbers and facilitates their application to physics. For example, comparison of (19) to (7) shows at once that the real and imaginary parts of a complex number are equivalent to inner and outer products of vectors. The complex conjugate of (19) is

$$z^\dagger = \lambda U^\dagger = \lambda e^{-i\theta} = ba,$$  \hspace{1cm} (20)

which shows that it is equivalent to reversing order in the geometric product. This can be used to compute the modulus of $z$ in the usual way:

$$|z|^2 = zz^\dagger = \lambda^2 = baab = a^2b^2 = |a|^2|b|^2$$ \hspace{1cm} (21)

Anyone who has worked with complex numbers in applications knows that it is usually best to avoid decomposing them into real and imaginary parts. Likewise, in GA applications it is usually best practice to work directly with the geometric product instead of separating it into inner and outer products.

GA gives complex numbers new powers to operate directly on vectors. For example, from (19) and (20) we get

$$b = a^{-1}z = z^\dagger a^{-1},$$ \hspace{1cm} (22)

where the multiplicative inverse of vector $a$ is given by

$$a^{-1} = \frac{1}{a} = \frac{a}{a^2} = \frac{a}{|a|^2}.$$ \hspace{1cm} (23)

Thus, $z$ rotates and rescales $a$ to get $b$. This makes it possible to construct and manipulate vectorial transformations and functions without introducing a basis or matrices.

This is a good point to pause and note some instructive implications of what we have established so far. Complex numbers, especially equations (17) and (18), are ideal for dealing with plane trigonometry and 2D rotations. However, students in introductory science and engineering courses are often denied access to this powerful tool, evidently because it has a reputation for being conceptually difficult, and class time would be lost by introducing it. GA removes these barriers to use of complex numbers by linking them to vectors and giving them a clear geometric meaning.

GA also makes it possible to formulate and solve 2D physics problems in terms of vectors without introducing coordinates. Conventional vector algebra cannot do this, in part because the vector cross product is defined only in 3D. That is the main reason why coordinate methods dominate introductory physics, computer science, engineering, etc. The available math tools are too weak to do otherwise. GA changes all that!

Most of the mechanics problems in introductory physics are 2D problems. Coordinate-free GA solutions for the standard problems are worked out in my mechanics book.\textsuperscript{12} Although the treatment there is for a more advanced course,
it can easily be adapted to the introductory level. The essential GA concepts for that level have already been presented in this section.

Will comprehensive use of GA significantly enhance student learning? We have noted theoretical reasons for believing that it will. However, mathematical reform at the introductory level makes little sense unless it is extended to the whole curriculum. Taking physics as example, the following sections provide strong justification for doing just that. We shall see how simplifications at the introductory level get amplified to greater simplifications and surprising insights at the advanced level.

IV. Classical Physics with Geometric Algebra

This Section surveys the fundamentals of GA as a mathematical framework for classical physics and demonstrates some of its unique advantages. Detailed applications can be found in the references.

A. Geometric Algebra for Physical Space

The arena for classical physics is a 3D Euclidean vector space $\mathcal{P}^3$, which serves as a model for “Physical Space.” By multiplication and addition the vectors generate a geometric algebra $\mathcal{G}_3 = \mathcal{G}(\mathcal{P}^3)$. In particular, a basis for the whole algebra can be generated from a standard frame $\{\sigma_1, \sigma_2, \sigma_3\}$, a righthanded set of orthonormal vectors.

With multiplication specified by (10), the standard frame generates a unique trivector (3-vector) or pseudoscalar

$$i = \sigma_1 \sigma_2 \sigma_3,$$  \hspace{1cm} (24)

and a bivector (2-vector) basis

$$\sigma_1 \sigma_2 = i \sigma_3, \quad \sigma_2 \sigma_3 = i \sigma_1, \quad \sigma_3 \sigma_1 = i \sigma_2.$$  \hspace{1cm} (25)

Geometric interpretations for the pseudoscalar and bivector basis elements are depicted in Figs. 7 and 8.

The pseudoscalar $i$ has special properties that facilitate applications as well as articulation with standard vector algebra. It follows from (24) that

$$i^2 = -1,$$  \hspace{1cm} (26)

and it follows from (25) that every bivector $\mathbf{B}$ in $\mathcal{G}_3$ is the dual of a vector $\mathbf{b}$ as expressed by

$$\mathbf{B} = i \mathbf{b} = \mathbf{b} i.$$  \hspace{1cm} (27)

Thus, the geometric duality operation is simply expressed as multiplication by the pseudoscalar $i$. This enables us to write the outer product defined by (6) in the form

$$\mathbf{a} \wedge \mathbf{b} = i \mathbf{a} \times \mathbf{b}.$$

\hspace{1cm} (28)
Fig. 7. Unit pseudoscalar \( i \) represents an oriented unit volume. The volume is said to be righthanded, because \( i \) can be generated from a righthanded vector basis by the ordered product \( \sigma_1 \sigma_2 \sigma_3 = i \).

Fig. 8. Unit bivectors representing a basis of directed areas in planes with orthogonal intersections.
Thus, the conventional vector cross product \( \mathbf{a} \times \mathbf{b} \) is implicitly defined as the dual of the outer product. Consequently, the fundamental decomposition of the geometric product (7) can be put in the form

\[
\mathbf{a} \mathbf{b} = \mathbf{a} \cdot \mathbf{b} + i \mathbf{a} \times \mathbf{b}.
\]  

(29)

This is the definitive relation among vector products that we need for smooth articulation between geometric algebra and standard vector algebra, as is demonstrated with many examples in my mechanics book.\(^\text{12}\)

The elements in any geometric algebra are called multivectors. The special properties of \( i \) enable us to write any multivector \( M \) in \( G_3 \) in the expanded form

\[
M = \alpha + \mathbf{a} + i \mathbf{b} + i \beta,
\]  

(30)

where \( \alpha \) and \( \beta \) are scalars and \( \mathbf{a} \) and \( \mathbf{b} \) are vectors. The main value of this form is that it reduces multiplication of multivectors in \( G_3 \) to multiplication of vectors given by (29).

The expansion (30) has the formal algebraic structure of a "complex scalar" \( \alpha + i \beta \) added to a "complex vector" \( \mathbf{a} + i \mathbf{b} \), but any physical interpretation attributed to this structure hinges on the geometric meaning of \( i \). A very important example is the expression of the electromagnetic field \( \mathbf{F} \) in terms of an electric vector field \( \mathbf{E} \) and a magnetic vector field \( \mathbf{B} \):

\[
\mathbf{F} = \mathbf{E} + i \mathbf{B}.
\]  

(31)

Geometrically, this is a decomposition of \( \mathbf{F} \) into vector and bivector parts. In standard vector algebra \( \mathbf{E} \) is said to be a polar vector while \( \mathbf{B} \) is an axial vector, the two kinds of vector being distinguished by a difference in sign under space inversion. GA reveals that an axial vector is just a bivector represented by its dual, so the magnetic field in (31) is fully represented by the complete bivector \( i \mathbf{B} \), rather than \( \mathbf{B} \) alone. Thus GA makes the awkward distinction between polar and axial vectors unnecessary. The vectors \( \mathbf{E} \) and \( \mathbf{B} \) in (31) have the same behavior under space inversion, but an additional sign change comes from space inversion of the pseudoscalar.

To facilitate algebraic manipulations, it is convenient to introduce a special symbol for the operation (called reversion) of reversing the order of multiplication. The reverse of the geometric product is defined by

\[
(\mathbf{a} \mathbf{b})^\dagger = \mathbf{b} \mathbf{a}.
\]  

(32)

We noted in (20) that this is equivalent to complex conjugation in 2D. From (24) we find that the reverse of the pseudoscalar is

\[
i^\dagger = -i.
\]  

(33)

Hence the reverse of an arbitrary multivector in the expanded form (30) is

\[
M^\dagger = \alpha + \mathbf{a} - i \mathbf{b} - i \beta,
\]  

(34)
The convenience of this operation is illustrated by applying it to the electromagnetic field $F$ in (31) and using (29) to get

$$\frac{1}{2}FF^\dagger = \frac{1}{2}(E + iB)(E - iB) = \frac{1}{2}(E^2 + B^2) + E \times B,$$

(35)

which is recognized as an expression for the energy and momentum density of the field.

You have probably noticed that the expanded multivector form (30) violates one of the basic math strutures that is drilled into our students, namely, that "it is meaningless to add scalars to vectors," not to mention bivectors and pseudoscalars. On the contrary, GA tells us that such addition is not only geometrically meaningful, it is essential to simplify and unify the mathematical language of physics and other applications, as can be seen in many examples that follow.

Shall we say that this stricture against addition of scalars to vectors is a misconception or even a conceptual virus? At least it is a design flaw in standard vector algebra that has been almost universally overlooked. As we have just seen, elimination of the flaw enables us to combine electric and magnetic fields into a single electromagnetic field. And we shall see below how it enables us to construct spinors from vectors (contrary to the received wisdom that spinors are more basic than vectors)!

**B. Reflections and Rotations**

Rotations play an essential role in the conceptual foundations of physics as well as in many applications, so our mathematics should be designed to handle them as efficiently as possible. We have noted that conventional treatments employ an awkward mixture of vector, matrix and spinor or quaternion methods. My purpose here is to show how GA provides a unified, coordinate-free treatment of rotations and reflections that leaves nothing to be desired.

The main result is that any orthogonal transformation $U$ can be expressed in the canonical form

$$U \mathbf{x} = \pm U \mathbf{x} U^\dagger,$$

(36)

where $U$ is a unimodular multivector called a versor, and the sign is the parity of $U$, positive for a rotation or negative for a reflection. The condition

$$U^\dagger U = 1,$$

(37)

defines unimodularity. The underbar notation serves to distinguish the linear operator $\underline{U}$ from the versor $U$ that generates it. The great advantage of (36) is that it reduces the study of linear operators to algebraic properties of their versors. This is best understood from specific examples.

The simplest example is reflection in a plane with unit normal $\mathbf{a}$ (Fig. 9),

$$\mathbf{x}' = -\mathbf{a} \mathbf{a} \mathbf{x} = -\mathbf{a}(\mathbf{x}_\perp + \mathbf{x}_\parallel)\mathbf{a} = \mathbf{x}_\perp - \mathbf{x}_\parallel$$

(38)
To show how this function works, the vector $\mathbf{x}$ has been decomposed on the right into a parallel component $\mathbf{x}_\parallel = (\mathbf{x} \cdot \mathbf{a})\mathbf{a}$ that commutes with $\mathbf{a}$ and an orthogonal component $\mathbf{x}_\perp = (\mathbf{x} \wedge \mathbf{a})\mathbf{a}$ that anticommutes with $\mathbf{a}$. As can be seen below, it is seldom necessary or even advisable to make this decomposition in applications. The essential point is that the normal vector defining the direction of a plane also represents a reflection in the plane when interpreted as a versor. A simpler representation for reflections is inconceivable, so it must be the optimal representation for reflections in every application, as shown in some important applications below. Incidentally, the term versor was coined in the 19th century for an operator that can re-verse a direction. Likewise, the term is used here to indicate a geometric operational interpretation for a multivector.

![Fig. 9. Reflection in a plane.](image)

The reflection (38) is not only the simplest example of an orthogonal transformation, but all orthogonal transformations can be generated by reflections of this kind. The main result is expressed by the following theorem: The product of two reflections is a rotation through twice the angle between the normals of the reflecting planes. This important theorem seldom appears in standard textbooks, primarily, I presume, because its expression in conventional formalism is so awkward as to render it impractical. However, it is an easy consequence of a second reflection applied to (38). Thus, for a plane with unit normal $\mathbf{b}$, we have

$$x'' = -\mathbf{b} x' \mathbf{b} = \mathbf{b} \mathbf{a} \mathbf{b} \mathbf{a} = U x U^\dagger,$$

where a new symbol has been introduced for the versor product $U = \mathbf{b} \mathbf{a}$. The theorem is obvious from the geometric construction in Fig. 10. For an algebraic proof that the result does not depend on the reflecting planes, we use (17) to write

$$U = \mathbf{b} \mathbf{a} = \cos \frac{1}{2} \theta + i \sin \frac{1}{2} \theta = e^{\frac{1}{2}i\theta},$$

where, anticipating the result from Fig. 9, we denote the angle between $\mathbf{a}$ and $\mathbf{b}$ by $\frac{1}{2} \theta$ and the unit bivector for the $\mathbf{a} \wedge \mathbf{b}$-plane by $i$. Next, we decompose $\mathbf{x}$ into a component $\mathbf{x}_\perp$ orthogonal to the $i$-plane and a component $\mathbf{x}_\parallel$ in the plane. Note that, respectively, the two components commute (anticommute) with $i$, so

$$\mathbf{x}_\perp U^\dagger = U^\dagger \mathbf{x}_\perp, \quad \mathbf{x}_\parallel U^\dagger = U \mathbf{x}_\parallel.$$
Inserting this into (39) with \( x = x_\parallel + x_\perp \), we obtain

\[
x'' = UxU^t = x_\perp + U^2x_\parallel.
\] (42)

These equations show how the two-sided multiplication by the versor \( U \) picks out the component of \( x \) to be rotated, so we see that one-sided multiplication works only in 2D. As we learned from our discussion of 2D rotations, the versor \( U^2 = e^{i\theta} \) rotates \( x_\perp \) through angle \( \theta \), in agreement with the half-angle choice in (40).

![Fig. 10. Rotation as double reflection, depicted in the plane containing unit normals \( a, b \) of the reflecting planes.](image)

The great advantage of the canonical form (36) for an orthogonal transformation is that it reduces the composition of orthogonal transformations to versor multiplication. Thus, composition expressed by the operator equation

\[
U_2 U_1 = U_3
\] (43)

is reduced to the product of corresponding versors

\[
U_2 U_1 = U_3.
\] (44)

The orthogonal transformations form a mathematical group with (43) as the group composition law. The trouble with (43) is that abstract operator algebra does not provide a way to compute \( U_3 \) from given \( U_1 \) and \( U_2 \). The usual solution to this problem is to represent the operators by matrices and compute by matrix multiplication. A much simpler solution is to represent the operators by versors and compute with the geometric product. We have already seen how the product of reflections represented by \( U_1 = a \) and \( U_2 = b \) produces a rotation \( U_3 = ba \). Matrix algebra does not provide such a transparent result.

As is well known, the \textit{rotation group} is a subgroup of the \textit{orthogonal group}. This is expressed by the fact that rotations are represented by unimodular versors of even parity, for which the term \textit{rotor} was introduced earlier. The composition of 2D rotations is described by the rotor equation (18) and depicted in Fig. 5. Its generalization to composition of 3D rotations in different planes...
is described algebraically by (44) and depicted geometrically in Fig. 11. This deserves some explanation.

In 3D a rotor is depicted as a directed arc confined to a great circle on the unit sphere. The product of rotors $U_1$ and $U_2$ is depicted in Fig. 11 by connecting the corresponding arcs at a point $c$ where the two great circles intersect. This determines points $a = cU_1$ and $b = U_2c$, so the rotors can be expressed as products with a common factor,

$$U_1 = ca, \quad U_2 = bc.$$  \hfill (45)

Hence (43) gives us

$$U_3 = U_2U_1 = (bc)(ca) = ba,$$  \hfill (46)

with the corresponding arc for $U_3$ depicted in Fig. 11. It should not be forgotten that the arcs in Fig. 11 depict half-angles of the rotations. The non-commutativity of rotations is illustrated in Fig. 12, which depicts the construction of arcs for both $U_1U_2$ and $U_2U_1$.

Those of you who are familiar with quaternions will have recognized that they are algebraically equivalent to rotors, so we might as well regard the two as one and the same. Advantages of the quaternion theory of rotations have been known for the better part of two centuries, but to this day only a small number of specialists have been able to exploit them. Geometric algebra makes them available to everyone by embedding quaternions in a more comprehensive mathematical system. More than that, GA makes a number of significant improvements in quaternion theory — the most important being the integration of reflections with rotations described above. To make this point more emphatic, I
describe two important practical applications where the generation of rotations by reflections is essential.

**Multiple reflections.** Consider a light wave (or ray) initially propagating with direction $\mathbf{k}$ and reflecting off a sequence of plane surfaces with unit normals $\mathbf{a}_1, \mathbf{a}_3, \ldots , \mathbf{a}_n$ (Fig. 12). By multiple applications of (38) we find that it emerges with direction

$$\mathbf{k'} = (-1)^n \mathbf{a}_n \ldots \mathbf{a}_2 \mathbf{a}_1 \mathbf{k} \mathbf{a}_1 \mathbf{a}_2 \ldots \mathbf{a}_n \quad (47)$$

The net reflection is completely characterized by a single unimodular multivector $U = \mathbf{a}_n \ldots \mathbf{a}_2 \mathbf{a}_1$, which, according to (30), can be reduced to the form $U = \mathbf{a} + i\beta$.
if $n$ is an odd integer, or $U = \alpha + ib$ if $n$ is even. This is one way that GA facilitates modeling of the interaction of light with optical devices.

![Diagram of benzene molecule with symmetry vectors](image1)

**Fig. 14.** Symmetry vectors for the benzene molecule.

![Diagram of methane molecule with symmetry vectors](image2)

**Fig. 15.** Symmetry vectors for the methane molecule.

**Point Symmetry Groups.** Molecules and crystals can be classified by their symmetries under reflections and rotations in planes through a fixed point. All such symmetries are generated by some combination of three unit vectors $a, b, c$ satisfying the versor conditions

$$(ab)^p = (bc)^q = (ca)^r = -1,$$  \hspace{1cm} (48)$$

where \{p, q, r\} is a set of three integers that characterize the symmetry group. These conditions describe $p$-fold, $q$-fold and $r$-fold rotation symmetries. For
example, the set is \{6, 2, 2\} for the planar benzene molecule (Fig. 14), and \((ab)^6 = -1\) represents a sixfold rotation that brings all atoms back to their original positions.

The methane molecule (Fig. 15) has tetrahedral symmetry characterized by \{3, 3, 3\}, which specifies the 3-fold symmetry of the tetrahedron faces. This particular symmetry cannot be extended to a space-filling crystal, for which it can be shown that at least one of the symmetries must be 2-fold.

There are precisely 32 crystallographic point groups distinguished by a small set of allowed values of \(p\) and \(q\) in \{\(p, q, 2\)\}. For example, generating vectors for the case \(\{4, 3, 2\}\) of crystals with cubic symmetry are shown in Fig. 16. A complete analysis of all the point groups is given elsewhere\(^\text{14}\) along with an extension of GA techniques to handle the 230 space groups.

The point symmetry groups of molecules and crystals are increasing in importance as we enter the age of nanoscience and molecular biology. Yet the topic remains relegated to specialized courses, no doubt because the standard treatment is so specialized. However, we have just seen that the GA approach to reflections and rotations brings with it an easy treatment of the point groups at no extra cost.

This is a good place to summarize with a list of the advantages of the GA approach to rotations, including some to be explained in subsequent Sections:

1. Coordinate-free formulation and computation.
2. Simple algebraic composition.
3. Geometric depiction of rotors as directed arcs.
4. Rotor products depicted as addition of directed arcs.
5. Integration of rotations and reflections in a single method.
6. Efficient parameterizations (see ref.\(^\text{12}\) for details).
7. Smooth articulation with matrix methods.
8. Rotational kinematics without matrices.
Moreover, the approach generalizes to Lorentz transformations.²

C. Frames and Rotation Kinematics
Any orthonormal righthanded frame \{e_k, k = 1, 2, 3\} can be obtained from our standard frame \{σ_k\} by a rotation in the canonical form
\[ e_k = Uσ_k U^\dagger. \] (49)

Alternatively, the frames can be related by a rotation matrix
\[ e_k = α_{kj}σ_j. \] (50)

These two sets of equations can be solved for the matrix elements as a function of \(U\), with the result
\[ α_{kj} = e_k \cdot σ_j = <Uσ_k U^\dagger σ_j>, \] (51)
where \(<...>\) means scalar part. Alternatively, they can be solved for the rotor as a function of the frames or the matrix.¹² One simply forms the quaternion
\[ ψ = 1 + e_kσ_k = 1 + α_{kj}σ_jσ_k \] (52)
and normalizes to get
\[ U = \frac{ψ}{(ψψ^\dagger)^{1/2}}. \] (53)

This makes it easy to move back and forth between matrix and rotor representations of a rotation. We have already seen that the rotor is much to be preferred for both algebraic computation and geometric interpretation.

Let the frame \{e_k\} represent a set of directions fixed in a rigid body,⁴⁴ perhaps aligned with the principal axes of the inertia tensor. For a moving body the \(e_k = e_k(t)\) are functions of time, and (49) reduces the description of the rotational motion to a time dependent rotor \(U = U(t)\). By differentiating the constraint \(UU^\dagger = 1\), it is easy to show that the derivative of \(U\) can be put in the form
\[ \frac{dU}{dt} = \frac{1}{2}ΩU, \] (54)
where
\[ Ω = -iω \] (55)
is the rotational (angular) velocity bivector. By differentiating (49) and using (54), (55), we derive the familiar equations
\[ \frac{de_k}{dt} = \omega \times e_k \] (56)
employed in the standard vectorial treatment of rigid body kinematics.

The point of all this is that GA reduces the set of three vectorial equations (56) to the single rotor equation (54), which is easier to solve and analyze for given \( \Omega = \Omega(t) \). Specific solutions for problems in rigid body mechanics are discussed elsewhere.\(^{12}\) However, the main reason for introducing the classical rotor equation of motion in this lecture is its equivalence to equations in quantum mechanics.\(^{1, 2, 19–30}\)

**D. Maxwell’s Equation**

We have seen how electric and magnetic field vectors can be combined into a single multivector field

\[
F(x, t) = E(x, t) + iB(x, t)
\]

representing the complete electromagnetic field.\(^{18, 31–33}\) Standard vector algebra forces one to consider electric and magnetic parts separately, and it requires four field equations to describe their coordinated action. GA enables us to put Humpty Dumpty together and describe the complete electromagnetic field by a single equation. But first we need to learn how to differentiate with respect to the position vector \( x \).

We can define the derivative \( \nabla = \partial_x \) with respect to the vector \( x \) most quickly by appealing to your familiarity with the standard concepts of divergence and curl. Then, since \( \nabla \) must be a vector operator, we can use (29) to define the vector derivative by

\[
\nabla E = \nabla \cdot E + i \nabla \times E = \nabla \cdot E + \nabla \wedge E.
\]

This shows the divergence and curl as components of a single vector derivative. Both components are needed to determine the field. For example, for the field due to a static charge density \( \rho = \rho(x) \), the field equation is

\[
\nabla E = \rho.
\]

The advantage of this form over the usual separate equations for divergence and curl is that \( \nabla \) can be inverted to solve for

\[
E = \nabla^{-1} \rho.
\]

Of course \( \nabla^{-1} \) is an integral operator determined by a Green’s function, but GA provides new insight into such operators. For example, for a source \( \rho \) with 2D symmetry in a localized 2D region \( \mathcal{R} \) with boundary \( \partial \mathcal{R} \), the \( E \) field is planar and \( \nabla^{-1} \) can be given the explicit form\(^{16, 17}\)

\[
E(x) = \frac{1}{2\pi} \int_{\mathcal{R}} |d^2x| \frac{1}{\mathbf{x}' - \mathbf{x}} \rho(x') + \frac{1}{2\pi i} \oint_{\partial R} \frac{1}{\mathbf{x}' - \mathbf{x}} d\mathbf{x}' E(x'),
\]

where \( i \) is the unit bivector for the plane. In the absence of sources, the first integral on the right vanishes and the field within \( \mathcal{R} \) is given entirely by a line
integral of its value over the boundary. The resulting equation is precisely equivalent to the celebrated Cauchy integral formula, as is easily shown by changing to the complex variable \( z = xa \), where \( a \) is a fixed unit vector in the plane that defines the “real axis” for \( z \). Thus GA automatically incorporates the full power of complex variable theory into electromagnetic theory. Indeed, formula (61) generalizes the Cauchy integral to include sources and the generalization can be extended to 3D with arbitrary sources.\(^{16,17}\) But this is not the place to discuss such matters.

An electromagnetic field \( F = F(x, t) \) with charge density \( \rho = \rho(x, t) \) and charge current \( J = J(x, t) \) as sources is determined by Maxwell's Equation

\[
\left( \frac{1}{c} \partial_t + \nabla \right) F = \rho - \frac{1}{c} J. \tag{62}
\]

To show that this is equivalent to the standard set of four equations, we employ (57), (58) and (30) to separate, respectively, its scalar, vector, bivector, and pseudoscalar parts:

\[
\begin{align*}
\nabla \cdot E &= \rho, \tag{63} \\
\frac{1}{c} \partial_t E - \nabla \times B &= -\frac{1}{c} J, \tag{64} \\
i \frac{1}{c} \partial_t B + \nabla \times E &= 0, \tag{65} \\
i \nabla \cdot B &= 0 \tag{66}
\end{align*}
\]

Here we see the standard set of Maxwell's equations as four geometrically distinct parts of one equation. Note that this separation into several parts is similar to separating equating real and imaginary parts in an equation for complex variables.

V. GA in the Mathematics Curriculum

My monograph on Geometric Algebra and Calculus as a unified language for mathematics\(^{17}\) is the most comprehensive reference on the subject, but it is not suitable as a textbook, except perhaps for a graduate course in mathematics. Besides there have been some important developments since it was first published. Some GA textbooks for physics students have recently been published.\(^3,33\) Of course, it will take more than some books to define a full curriculum.

As it does for physics, GA provides a framework for critique of the current math curriculum. I mention only courses that are mainstays of mathematical physics. A full critique of these courses requires much more space than we can afford here. By first introducing GA as the basic language, the course in linear algebra can be simplified and enriched.\(^{45}\) For example, we have seen how GA facilitates the treatment of rotations and reflections. GA will then
supplant matrix algebra as the basic computation system. Of course, matrix algebra is a very powerful and well-developed system, but it is best developed from GA rather than the other way around. Courses on advanced calculus and multivariable calculus with differential forms and differential geometry are unified and simplified by geometric calculus. Likewise, GA unifies courses on real and complex analysis. Group theory can also be developed within the GA framework, but much work remains to incorporate the full range of available methods and results.

VI. Outlook

I challenge educators, scientists and engineers to critically examine the following claims supported by the argument in this paper:

- GA provides a unified language for the whole of physics and for much of mathematics and its applications that is conceptually and computationally superior to alternative mathematical systems in many application domains.
- GA can enhance student understanding and accelerate student learning.
- GA is ready to incorporate into curricula, especially the physics curriculum.
- GA provides new insight into the structure and interpretation of mathematical applications.
- Research on the design and use of mathematical tools is equally important for instruction and theory.
- Reforming the mathematical language of physics is the single most essential step toward simplifying physics education at all levels from high school to graduate school. Similar benefits can be expected for other disciplines.

Note. Most of my papers listed in the references are available online. Physics education research papers can be accessed from <http://modeling.asu.edu>. GA papers can be accessed from <http://modelingnts.asu.edu>. Many fine papers on GA applications in physics and engineering are available at the Cambridge website <http://www.mrao.cam.ac.uk/~clifford/>.

This paper is an extract from my Oersted Medal Lecture. I thank Eckhard Hitzer for help in preparing it.

References


[40] L. Dorst, C. Doran & J. Lasenby (Eds.), Applications of Geometrical Algebra in Computer Science and Engineering (Birkhäuser, Boston, 2002)


