Vectors, Spinors, and Complex Numbers in Classical and Quantum Physics

David Hestenes

Abstract Geometric calculus is shown to unite vectors, spinors, and complex numbers into a single mathematical system with a comprehensive geometric significance. The efficacy of this calculus in physical applications is explicitly demonstrated.

[**Note:** The terms "multivector algebra" and "multivector calculus" originally used in this paper have been replaced throughout by the terms "geometric algebra" and "geometric calculus," which have since become standard terms.]

INTRODUCTION

Among the many alternative mathematical systems in which the equations of physics can be expressed, two of the most popular are the matrix form of spinor calculus and the vector calculus formulated by Gibbs. These two systems in some measure complement one another and are often used together in problems concerning particles with spin. However, many theorems of vector calculus are equivalent to theorems of matrix calculus. So when the two systems are combined a great deal of redundancy occurs which makes computations more difficult than necessary, much time being taken up with mere translation between the two modes of expression. This paper shows how the matrix and vector algebra can be replaced by a single mathematical system, called geometric algebra, with which the tasks of theoretical physics can be carried out more efficiently.

Geometric algebra derives its power from the fact that both the elements and the operations of the algebra are subject to direct geometrical interpretation. It can be argued, further, that geometric algebra is the simplest system capable of providing a complete algebraic expression of geometric concepts, though only some examples of the effectiveness of the system are given in this paper.

Additional perspective on the cogency of geometric algebra can be gained by comparing it with other approaches. For example, it is important to note that covariant formulations of tensor and spinor algebra are by nature inefficient, because at the very beginning they introduce irrelevant coordinates. The coordinates function as a vehicle for the algebraic system. The actual irrelevancy of the coordinates is finally admitted by requiring that quantities which represent intrinsic geometric (and physical) entities be covariant under coordinate transformations. In contrast, geometric algebra is built out of objects with direct geometric interpretations; the properties of these objects are specified by introducing algebraic operations which

directly determine their interrelations. Coordinates are utilized only when they arise naturally in specific problems. Moreover, geometric algebra retains all the advantages of tensor algebra, because tensors can be introduced in a coordinate free fashion as multilinear functions on geometric algebra; the operations of tensor algebra are then already included in the operations of geometric algebra. However, a systematic account of this approach to tensors has not yet been published. [But see more recent references on the web site.]

Though the geometric algebra discussed here is isomorphic to the so-called "Pauli (matrix) algebra," the interpretations of the two systems differ considerably, and the practical consequences of this difference are not trivial. Thus, questions of the representation of Pauli matrices and of transformations among representations never arise in geometric algebra, because they are irrelevant. Matrix algebra was invented to describe linear transformations. So it should not be surprising to find that matrices have irrelevant features when they are used to represent objects of a different nature. From the geometric viewpoint of geometric algebra, matrices are seen to arise in the theory of linear geometric functions; that is, geometrics are taken to be more fundamental than matrices, rather than the other around. Simplifications which result from this reversal of viewpoint are manifest in text and references of this paper.

Just as important as the mathematical simplifications that accrue is the fact that the use of geometric algebra imbues many well known mathematical expressions with new meaning. Thus, complex numbers arise naturally with distinctive geometrical and physical interpretations depending on the situation. Of particular significance is the realization of a connection between spin and complex numbers in quantum theory and the surprising conclusion that the Schrodinger theory already describes a particle with spin. This will be discussed thoroughly in a subsequent paper (hereinafter referred to as II[1]). [Also discussed in many other papers on this web site.]

The purpose of this paper is to explain some of the advantages of geometric calculus. The object is not to give a complete and balanced account of the subject, but only to call attention to its major features and some of the minor points that are often overlooked. Other aspects of the subject are treated in Refs. 2–5.

In Sec. 1, the fundamentals of geometric calculus are discussed in preparation for applications to follow. Section 2 delineates the role of spinors. Section 3 illustrates the use of spinors in classical mechanics. Section 4 discusses the significance of complex numbers in electrodynamics.

This paper makes only brief and incomplete allusions to quantum theory. However, the full power of geometric methods in quantum theory is demonstrated in II [1].

1 GEOMETRIC ALGEBRA

The most distinctive feature of geometric algebra is the rule for multiplying vectors. The product **ab** of vectors **a** and **b** is designed to contain all the information about the relative directions of vectors **a** and **b**. To see how this is done, it is helpful to

decompose the product into symmetric and antisymmetric parts. Write

$$\mathbf{ab} = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b} \,, \tag{1.1}$$

where

$$\mathbf{a} \cdot \mathbf{b} \equiv \tfrac{1}{2}(\mathbf{ab} + \mathbf{ba}) = \mathbf{b} \cdot \mathbf{a} \,, \tag{1.2}$$

$$\mathbf{a} \wedge \mathbf{b} \equiv \tfrac{1}{2}(\mathbf{ab} - \mathbf{ba}) = -\mathbf{b} \wedge \mathbf{a} \,. \tag{1.3}$$

The symmetric part of $\mathbf{a} \cdot \mathbf{b}$ is taken to be the familiar way inner (or scalar) product of vectors, so no comment about its geometrical significance is required. The antisymmetric part of $\mathbf{a} \wedge \mathbf{b}$ is taken to be a bivector determined by vectors $\mathbf{a}$ and $\mathbf{b}$.

Lamentably, bivectors are not so widely appreciated as they deserve to be, though they were invented together with vectors well over a hundred years ago by Grassmann. Since in geometric algebra bivectors are as important as vectors and scalars, a few words on their behalf may be in order. Just as a vector can be interpreted as a directed number which describes an oriented line segment, with the direction of the vector representing the oriented line and the magnitude of the vector measuring the length of the segment; so a bivector can be interpreted as a directed number which describes an oriented plane segment, with the direction of the bivector representing the oriented plane and the magnitude of the bivector measuring the area of the plane segment. The direction of the bivector $B = \mathbf{a} \wedge \mathbf{b}$ represents the oriented plane "containing" $\mathbf{a}$ and $\mathbf{b}$, while the magnitude of $\mathbf{a} \wedge \mathbf{b}$ is equal to the area of a parallelogram with sides corresponding to $\mathbf{a}$ and $\mathbf{b}$. The bivector B should not be conceived merely as the product $\mathbf{a} \wedge \mathbf{b}$, for B is an entity unto itself. The product simply describes a relation of the vectors $\mathbf{a}$ and $\mathbf{b}$ to B, just as $2 \times 3 = 6$ describes a relation of 2 and 3 to 6.

The product $\mathbf{ab}$ acquires a geometrical significance from the interpretations given to $\mathbf{a} \cdot \mathbf{b}$ and $\mathbf{a} \wedge \mathbf{b}$. This product may be interpreted directly as a measure of the relative directions of vectors $\mathbf{a}$ and $\mathbf{b}$. Thus, if $\mathbf{a}$ and $\mathbf{b}$ are collinear, then they commute: $\mathbf{ab} = \mathbf{ba}$. Or, if $\mathbf{a}$ and $\mathbf{b}$ are orthogonal, they anticommute: $\mathbf{ab} = -\mathbf{ba}$. In general, $\mathbf{ab}$ describes a "degree of commutivity" somewhere between these two extremes. Neither of the products $\mathbf{a} \cdot \mathbf{b}$ or $\mathbf{a} \wedge \mathbf{b}$ suffices by itself to fully describe the geometrical relation of $\mathbf{a}$ and $\mathbf{b}$; the full product $\mathbf{ab}$ is needed.

While development of the full geometric algebra in a coordinate free manner is the ideal way to proceed, it is convenient here to introduce a basis of vectors at once, because it facilitates comparison with matrix representations of algebra. Furthermore, only the geometric algebra appropriate for the description of objects in three-dimensional physical space is discussed here, though the whole approach is easily adapted to spaces of higher dimension. Application of geometric algebra to space-time physics is carried out in Refs. 2 and 3.

Let $\{\boldsymbol{\sigma}_i, \ i = 1, 2, 3\}$ be a right-handed orthonormal basis of vectors representing directions in physical space. Then Eq. (1.1) gives

$$\boldsymbol{\sigma}_1^2 = \boldsymbol{\sigma}_2^2 = \boldsymbol{\sigma}_3^2 = 1 \,, \tag{1.4}$$

$$\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2 = \boldsymbol{\sigma}_1 \wedge \boldsymbol{\sigma}_2 = -\boldsymbol{\sigma}_2\boldsymbol{\sigma}_1\,,$$

$$\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3 = \boldsymbol{\sigma}_2 \wedge \boldsymbol{\sigma}_3 = -\boldsymbol{\sigma}_3\boldsymbol{\sigma}_2\,,$$

$$\boldsymbol{\sigma}_3\boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_3 \wedge \boldsymbol{\sigma}_1 = -\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3\,. \tag{1.5}$$

Just as any vector **a** can be expressed as a linear combination of basis vectors:

$$\mathbf{a} = a_i\boldsymbol{\sigma}_i \qquad \text{(sum over repeated indices)}, \tag{1.6}$$

so any *bivector* S can be expressed as a linear combination of basis bivectors $\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j$:

$$S = \tfrac{1}{2}S_{ij}\,\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j\,. \tag{1.7}$$

The scalars $S_{ij} = -S_{ji}$ are called components of S with respect to the basis $\{\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j\}$. From (1.5) it follows that the bivectors compose a three-dimensional linear space.

Besides the bivectors $\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j$, only one other new entity can be constructed by multiplying together the $\boldsymbol{\sigma}_i$, namely the *pseudoscalar*

$$i = \boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3 = \boldsymbol{\sigma}_1 \wedge \boldsymbol{\sigma}_2 \wedge \boldsymbol{\sigma}_3\,. \tag{1.8}$$

Just as $\boldsymbol{\sigma}_1 \wedge \boldsymbol{\sigma}_2$ can be interpreted as an oriented plane segment with unit area, so the pseudoscalar i ean be interpreted as an oriented three-space segment with unit volume. And i should not be conceived as simply an abbreviation for the product $\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3$; the pseudoscalar is an entity unto itself, as significant as scalars, vectors, and bivectors. Equation (1.8) expresses the fact that i can be factored into a product of three vectors; actually, any three orthonormal vectors will do.

Every element of the geometric algebra of physical space can be expressed as a linear combination of the basis elements $\{1, \boldsymbol{\sigma}_i, \boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j, i\}$, composed of the vectors $\boldsymbol{\sigma}_i$ and their products. It follows that every geometric M can be expressed as the sum of a scalar M_S, a vector M_V, a bivector M_B, and a pseudoscalar M_P:

$$M = M_S + M_V + M_B + M_P\,. \tag{1.9}$$

It is fair to ask what it means to add scalars, vectors, bivectors, and pseudoscalars. Mathematically, the operation of addition is well defined by the usual rules, including the distributive rule, which relates addition to multiplication; together with the consistency requirement that scalars are linearly independent of vectors, while vectors are linearly independent of bivectors, and so on. On the other hand, the geometrical and physical meanings of geometric addition are many. One example has already been given by Eq. (1.1), which shows that the geometric properly describing the relative direction of two vectors consists of the sum of a scalar and a bivector. Another example is given by the fact that electromagnetic field is best represented by a geometric consisting of a vector and a bivector. Every combination of scalar, vector, bivector, and pseudoscalar turns out to be significant in one application or another.

If there is one message which needs repeated emphasis in these days when the abstract nature of mathematics is so avidly cultivated, it is that geometric algebra derives its potency from its geometric interpretation. Anyone who thinks that the interpretation is a mere adornment to the algebra is invited to examine the literature on Clifford algebra. He will find that where Clifford algebra is treated as a purely abstract system, the mathematics is hardly developed beyond the most trivial theorems. But when geometrical considerations intrude, the mathematical theory is correspondingly enriched. Geometric algebra is Clifford algebra with a thoroughgoing geometric interpretation; it is entirely a geometric algebra. This paper and Refs. 2–5 only begin to show what can be accomplished by this union of algebra and geometry.

To emphasize geometrical interpretations as well as to increase algebraic efficiency, it is desirable to introduce more concepts and definitions than are strictly needed for a formulation of the fundamentals of geometric algebra.

A geometric which can be factored into the product of k orthogonal vectors is called a k-vector. It is convenient to augment this definition by calling scalars 0-vectors. A k-vector is said to be even if k is even and odd if k is odd. Thus, scalars and bivectors are even k-vectors, while vectors and pseudoscalars are odd k-vectors. In general, a geometric will be said to be even (or odd) if it can be expressed as the sum of even (or odd) k-vectors.

For a k-vector A_k and an m-vector B_m, the *inner product* $A_k \cdot B_m$ is defined by

$$A_k \cdot B_m = (A_k B_m)_{|\,k-m\,|}, \tag{1.10}$$

where the subscript on the right means take the $|\,k-m\,|$-vector part of the products $A_k B_m$; analogously, the *outer product* $A_k \wedge B_m$ is defined by

$$A_k \wedge B_m = (A_k B_m)_{k+m}. \tag{1.11}$$

The inner and outer products of a vector $\mathbf{a}$ with a bivector B arise quite frequently, and it can be shown that

$$\mathbf{a} \cdot B = \tfrac{1}{2}(\mathbf{a}B - B\mathbf{a}). \tag{1.12}$$

This is a vector. Also,

$$\mathbf{a} \wedge B = \tfrac{1}{2}(\mathbf{a}B + B\mathbf{a}). \tag{1.13}$$

This is a three-vector. Equivalently, it is a pseudoscalar, so

$$\mathbf{a} \wedge B = \lambda i, \tag{1.14}$$

where λ is a scalar, and as always in this paper, i is the unit pseudoscalar. Adding (1.12) to (1.13) gives

$$\mathbf{a}B = \mathbf{a} \cdot B + \mathbf{a} \wedge B. \tag{1.15}$$

Note the similarity to Eq. (1.1).

In geometric algebra multiplication is noncommutative, yet the discussion of Eq. (1.1) already shows that the order of terms in a product has geometric significance. To represent symbolically the important operation of reversing the order of terms in a product, it is convenient to define a geometric $M^\dagger$ which is said to be the *adjoint* of a given geometric M. The adjoint of every geometric is determined by specifying the adjoint of every vector and the rule for determining the adjoint of a product from the adjoints of its factors; to wit,

$$\mathbf{a}^\dagger = \mathbf{a} \tag{1.16}$$

and

$$(MN)^\dagger = N^\dagger M^\dagger \tag{1.17}$$

for every vector $\mathbf{a}$ and all geometrics M and N. For example, when the rules (1.16) and (1.17) are applied, it follows from (1.3) that $(\mathbf{a} \wedge \mathbf{b})^\dagger = \mathbf{b} \wedge \mathbf{a} = -\mathbf{a} \wedge \mathbf{b}$, and from (1.5) and (1.8) that $i^\dagger = -i$. More generally, (1.9) yields

$$M^\dagger = M_S + M_V - M_B - M_P. \tag{1.18}$$

The *magnitude* (or *modulus*) $|M|$ of a geometric M is determined by the equation

$$|M| = [(MM^\dagger)_S]^{1/2}, \tag{1.19}$$

where the subscript S means take the scalar part of the product $MM^\dagger$. Substitution of (1.9) and (1.18) into (1.19) relates the magnitude of M to that of its k-vector parts; thus

$$
\begin{aligned}
|M|^2 &= (MM^\dagger)_S \\
&= |M_S|^2 + |M_V|^2 + |M_B|^2 + |M_P|^2 \\
&= M_S^2 + M_V^2 - M_B^2 - M_P^2.
\end{aligned}
\tag{1.20}
$$

It follows that $|M| \geq 0$, and $|M| = 0$ only if $M = 0$. A geometric M is said to be *unitary* if $|M| = 1$.

As usual, the *inverse* M^{-1} of a geometric M is defined by the equation

$$MM^{-1} = 1. \tag{1.21}$$

If $MM^\dagger = |M|^2$, the inverse of M exists and can be written in the simple form

$$M^{-1} = M^\dagger / |M|^2. \tag{1.22}$$

It follows that every vector has an inverse, so it is possible to divide by vectors. This one feature alone gives geometric calculus a great advantage over "vector algebras" in which vector division is not defined. Note that vector division is made possible by (1.1) which combines inner and outer products into a single product.

The *dual* of a geometric is obtained by multiplying it by the unit pseudoscalar i. Particularly significant is the fact that every bivector S is the dual of a unique vector $\mathbf{s}$, that is

$$S = i\mathbf{s}. \tag{1.23}$$

This definition of dual is equivalent to that given in covariant tensor analysis, namely

$$S_{ij} = \epsilon_{ijk} S_k \,. \tag{1.24}$$

Equation (1.24) can be obtained from (1.23) by taking the inner product with $\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j$. Thus

$$(\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j) \cdot S = S_{ji} = -S_{ij} \,, \tag{1.25}$$

$$(\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j) \cdot (i\mathbf{s}) = [\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j i s_k \boldsymbol{\sigma}_k]_S = -\epsilon_{ijk} s_k \,. \tag{1.26}$$

[In (1.26) and below, multiplications indicated by a symbol are meant to be carried out first, for instance

$$A \cdot BC = (A \cdot B)C \neq A \cdot (BC)$$

and

$$AB \wedge C = A(B \wedge C) \neq (AB) \wedge C \,.$$

This convention eliminates a lot of parentheses.] It should be noted that (1.25) can be obtained from (1.7) by using

$$(\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j) \cdot (\boldsymbol{\sigma}_m \wedge \boldsymbol{\sigma}_n) = \delta_{jm}\delta_{in} - \delta_{jn}\delta_{im} \,, \tag{1.27}$$

which can be derived by inserting the steps

$$\begin{aligned}
&= [\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j \boldsymbol{\sigma}_m \wedge \boldsymbol{\sigma}_n]_S = [\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j \boldsymbol{\sigma}_m \wedge \boldsymbol{\sigma}_n]_S \\
&= [\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j \cdot (\boldsymbol{\sigma}_m \wedge \boldsymbol{\sigma}_n)]_S \\
&= \boldsymbol{\sigma}_i \cdot (\delta_{jm}\boldsymbol{\sigma}_n - \delta_{jn}\boldsymbol{\sigma}_m) \,.
\end{aligned}$$

Equation (1.26) can be established by noting the following equivalent expressions for the components of the alternating tensor:

$$\begin{aligned}
\epsilon_{ijk} &= -i\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j \wedge \boldsymbol{\sigma}_k = i^\dagger \boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j \wedge \boldsymbol{\sigma}_k \\
&= -i \cdot (\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j \wedge \boldsymbol{\sigma}_k) = -(i\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j \wedge \boldsymbol{\sigma}_k)_S \\
&= -(i\boldsymbol{\sigma}_i \wedge \boldsymbol{\sigma}_j \boldsymbol{\sigma}_k)_S = -(i\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j \boldsymbol{\sigma}_k)_S \\
&= (\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_2 \boldsymbol{\sigma}_3 \boldsymbol{\sigma}_i \boldsymbol{\sigma}_j \boldsymbol{\sigma}_k)_S \,. \tag{1.28}
\end{aligned}$$

It should be obvious that the coordinate free formulation of the dual is far simpler than the usual tensor relation. For instance, to invert (1.23) one only needs to remember that $i^2 = -1$; thus

$$\mathbf{s} = -iS \,. \tag{1.29}$$

To invert (1.24), more formulas, more indices, and more bookkeeping is needed, all of which is not worth the trouble now that a much simpler approach is available.

Of course the difference in complexity between the two approaches increases as problems become more difficult.

The "cross product" introduced by Gibbs is just the dual of the outer product, thus

$$\mathbf{a} \times \mathbf{b} = -i\,\mathbf{a} \wedge \mathbf{b}\,. \tag{1.30}$$

Multiplication of (1.8) on the left by $\boldsymbol{\sigma}_2\boldsymbol{\sigma}_1$ yields

$$\boldsymbol{\sigma}_3 = \boldsymbol{\sigma}_2\boldsymbol{\sigma}_1 i = i\boldsymbol{\sigma}_2\boldsymbol{\sigma}_1 = -i\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2 = \boldsymbol{\sigma}_1 \times \boldsymbol{\sigma}_2\,. \tag{1.31}$$

which shows that (1.8) and (1.30) are consistent with the usual right-hand rule for the cross product.

As defined in (1.30), $\mathbf{a} \times \mathbf{b}$ is a vector. It is a common practice [6] to define the cross product so that $\mathbf{a} \times \mathbf{b}$ transforms as an "axial vector" under space inversions. Actually, this axial vector is nothing more than the bivector $\mathbf{a} \wedge \mathbf{b}$ disguised as a vector. Since geometric algebra has bivectors to represent bivectors, the confusing practice of admitting two kinds of vectors is wholly unneccesary. In this connection, it should be noted that (1.30) shows that under space inversions the transformtion of the cross product differs from that of the outer product, because the pseudoscalar i changes sign.

Since the $\boldsymbol{\sigma}_k$ represent fixed directions in physical space, they are associated with a set of Cartesian coordinates $\{\boldsymbol{\sigma}_k, k = 1, 2, 3\}$. The association can be explicitly formulated by introducing the *gradient* operator ∇ with the formula

$$\nabla = \boldsymbol{\sigma}_k \partial_k\,, \tag{1.32}$$

where

$$\partial_k = \boldsymbol{\sigma}_k \cdot \nabla = \partial/\partial x_k\,. \tag{1.33}$$

Clearly

$$\boldsymbol{\sigma}_k = \nabla x_k\,, \tag{1.34}$$

which says that $\boldsymbol{\sigma}_k$ measures the direction and rate of change of x_k. In truth, the $\boldsymbol{\sigma}_k$ provide a mathematical formulation of the notion of uniform increase in a fixed direction implicit in the definition of the x_k.

When operating on a scalar, ∇ is equivalent to the usual gradient operator. However, in geometric calculus the gradient of a vector, indeed of any geometric, is a well-defined quantity. Thus, by virtue of (1.1) and (1.30), for a vector function $\mathbf{A}$,

$$\nabla\mathbf{A} = \nabla \cdot \mathbf{A} + \nabla \wedge \mathbf{A} = \nabla \cdot \mathbf{A} + i\nabla \times \mathbf{A}\,. \tag{1.35}$$

$\nabla \cdot \mathbf{A}$ and $\nabla \times \mathbf{A}$ are precisely the divergence and curl of Gibbs. The divergence determines the flux of the normal component of $\mathbf{A}$ through a given surface while the curl determines the flux of the tangential component of $\mathbf{A}$. Thus the complete rate of change of $\mathbf{A}$ across a surface is described by $\nabla\mathbf{A}$.

In Ref. [4] the gradient is defined in a more fundamental fashion without introducing coordinates. The gradient was introduced here by (1.32) to expediate comparison with better known approaches. In Ref. [5] the inverse of the gradient is found and shown to be a natural generalization of the famous Cauchy integral formula. This result has many applications to physics. For instance, it greatly facilitates the solving of electrostatic and magnetostatic problems directly for the field strengths without introducing potentials.

An isomorphism of the geometric algebra of physical space to the algebra of complex 2×2 matrices can be established by means of the following correspondences:

$$\boldsymbol{\sigma}_1 \sim \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_2 \sim \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

$$\boldsymbol{\sigma}_3 \sim \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad i \sim iI = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \tag{1.36}$$

The symbol i, of course, means something different in each of the algebras, but in both cases $i^2 = -1$, and i commutes with all elements of the algebra, which is reason enough for using the same symbol in both cases. In the matrix algebra, the significance of i is exhausted by the algebraic properties just mentioned; this i has no geometric significance. But the geometric i is to be interpreted geometrically as a pseudoscalar; this gives it additional algebraic properties, such as (1.8). As (1.36) indicates, the matrix analog of the pseudoscalar is not simply the "unit imaginary" in the matrix theory; it is the unit imaginary times the unit matrix.

The so-called Pauli algebra is more than just the algebra of complex 2×2 matrices; it is matrix algebra together with a geometrical interpretation. This association is usually set up by interpreting the set of matrices $\{\sigma_k\}$ as the *components* of a single vector. This clumsy approach is in large measure responsible for sustaining the practice of using two algebras, a matrix algebra and a vector or tensor algebra, where one will do. The preferred approach is, of course, to associate the σ_k with vectors as in (1.36), and then it becomes apparent that the matrices ought to be dispensed with altogether. But this program cannot be consummated unless spinors can be represented without matrices. That this can be done, and with salutary effects, is explicitly demonstrated in the following pages and in II [1].

The geometric algebra of physical space was called the Pauli algebra in Ref. [2], because of its similarity to the Pauli matrix algebra. Pauli's name is not the most appropriate to use in this connection, though Pauli was the first to use the matrix algebra in a discussion of electron spin. Nor is the use of Clifford's name better justified, though the name Clifford algebra is widely used in the literature. The names of Grassmann and Hamilton would be more appropriate, because they showed how to build geometric meaning into algebraic systems, but both failed to discover the rule (1.1) which unites their separate works into a harmonious whole. Clifford discovered the rule, but he failed to realize its geometric significance. The import of Clifford's innovation has remained unrealized to this day, though geometric algebras have been enriched by many men in manifold ways.

Instead of associating the name of any one man with a system which has evolved from the work of so many men, the descriptive name *geometric algebra* has been

adopted here. The word "multivector" has often been used in the literature in a sense similar to the one advocated here. Unfortunately, it is not a lovely word and is rather a mouthful to say.

2 REFLECTIONS, ROTATIONS, AND SPINORS

Consider the following function of a vector $\mathbf{v}$:

$$\mathbf{v} \to \mathbf{v}' = -\mathbf{ava}\,, \tag{2.1}$$

where $\mathbf{a}$ is a unit vector. Squaring, one finds that this mapping preserves the magnitude of $\mathbf{v}$:

$$\mathbf{v}'^2 = \mathbf{avaava} = \mathbf{avva} = \mathbf{v}^2 = |\,\mathbf{v}\,|^2\,. \tag{2.2}$$

Further insight into the nature of (2.1) can be obtained by decomposing $\mathbf{v}$ into a component $\mathbf{v}_\parallel$ collinear with $\mathbf{a}$ plus a component $\mathbf{v}_\perp$ orthogonal to $\mathbf{a}$; thus

$$\mathbf{v} = \mathbf{va}^2 = (\mathbf{v} \cdot \mathbf{a} + \mathbf{v} \wedge \mathbf{a})\mathbf{a} = \mathbf{v}_\parallel + \mathbf{v}_\perp\,, \tag{2.3}$$

where

$$\mathbf{v}_\parallel = (\mathbf{v} \cdot \mathbf{a})\mathbf{a}\,, \qquad \mathbf{v}_\perp = \mathbf{v} \wedge \mathbf{a}\,\mathbf{a} = (\mathbf{v} \wedge \mathbf{a}) \cdot \mathbf{a}\,, \tag{2.4}$$

and

$$\mathbf{v}_\parallel\mathbf{a} = \mathbf{v} \cdot \mathbf{a} = \mathbf{av}_\parallel\,, \qquad \mathbf{v}_\perp\mathbf{a} = \mathbf{v} \wedge \mathbf{a} = -\mathbf{av}_\perp\,, \tag{2.5}$$

So (2.1) yields

$$\mathbf{v}' = -\mathbf{a}(\mathbf{v}_\parallel + \mathbf{v}_\perp)\mathbf{a} = -\mathbf{v}_\parallel + \mathbf{v}_\perp\,. \tag{2.6}$$

This shows that the mapping (2.1) simply reverses the direction of any vector collinear with $\mathbf{a}$ and leaves any vector orthogonal to $\mathbf{a}$ unaffected; thus (2.10) is a *reflection* in the plane "orthogonal" to $\mathbf{a}$.

If the mapping (2.1) is followed by the mapping

$$\mathbf{v}' \to \mathbf{v}'' = -\mathbf{bv}'\mathbf{b}\,, \tag{2.7}$$

where $\mathbf{b}$ is a unit vector, the following composite mapping is obtained:

$$\mathbf{v} \to \mathbf{v}'' = U\mathbf{v}U^\dagger\,, \tag{2.8a}$$

where

$$U = \mathbf{ba} = \cos \tfrac{1}{2}|\,\theta\,| + \mathbf{i}\sin \tfrac{1}{2}|\,\theta\,|$$
$$= \exp\left(\tfrac{1}{2}\mathbf{i}|\,\theta\,|\right) = \exp\left(\tfrac{1}{2}\theta\right) \tag{2.8b}$$

To write the terms on the right of (2.8b), (1.1) was used together with the relations

$$\cos \tfrac{1}{2}|\theta| = \mathbf{a} \cdot \mathbf{b}, \tag{2.9a}$$

$$\sin \tfrac{1}{2}|\theta| = |\mathbf{b} \wedge \mathbf{a}|, \tag{2.9b}$$

$$\mathbf{i} = \mathbf{b} \wedge \mathbf{a}/|\mathbf{b} \wedge \mathbf{a}|, \tag{2.9c}$$

$$\theta = \mathbf{i}|\theta|. \tag{2.9d}$$

The scalar $|\theta|$ is readily identified as twice the magnitude of the angle between $\mathbf{b}$ and $\mathbf{a}$. The reason for introducing a half-angle is evident below. The unit bivector which represents the oriented "$\mathbf{b} \wedge \mathbf{a}$ plane" has been designated by the suggestive symbol $\mathbf{i}$, because $\mathbf{i}^2 = -1$; boldface type has been used to avoid confusion of the bivector $\mathbf{i}$ with the pseudoscalar i. It should be noted that the angle between the vectors $\mathbf{b}$ and $\mathbf{a}$ is fully represented only by the bivector $\tfrac{1}{2}\theta$ and not just a scalar; the "direction of the angle" represents the plane in which the magnitude $|\theta|$ is measured; the "sign of the angle" designates the orientation of the bivector.

To get at the meaning of (2.8a) it is useful to decompose $\mathbf{v}$ into a component $\mathbf{v}_{\parallel}$ which lies in the "$\mathbf{i}$ plane" and a component $\mathbf{v}_{\perp}$ which is "perpendicular" to it. Thus, since $-\mathbf{i}^2 = \mathbf{i}^{\dagger}\mathbf{i} = 1$,

$$\mathbf{v} = \mathbf{v}\mathbf{i}^{\dagger}\mathbf{i} = (\mathbf{v} \cdot \mathbf{i}^{\dagger} + \mathbf{v} \wedge \mathbf{i}^{\dagger})\mathbf{i} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp}, \tag{2.10}$$

where

$$\mathbf{v}_{\parallel} = (\mathbf{v} \cdot \mathbf{i}^{\dagger})\mathbf{i} = -(\mathbf{v} \cdot \mathbf{i})\mathbf{i} = -(\mathbf{v} \cdot \mathbf{i}) \cdot \mathbf{i}, \tag{2.11a}$$

$$\mathbf{v}_{\perp} = (\mathbf{v} \wedge \mathbf{i}^{\dagger})\mathbf{i} = -(\mathbf{v} \wedge \mathbf{i})\mathbf{i} = -(\mathbf{v} \wedge \mathbf{i}) \cdot \mathbf{i}, \tag{2.11b}$$

and

$$\mathbf{v}_{\parallel}\mathbf{i} = \mathbf{v} \cdot \mathbf{i} = -\mathbf{v} \cdot \mathbf{i} = -\mathbf{i}\mathbf{v}_{\parallel}, \tag{2.12a}$$

$$\mathbf{v}_{\perp}\mathbf{i} = \mathbf{v} \wedge \mathbf{i} = \mathbf{i} \wedge \mathbf{v} = \mathbf{i}\mathbf{v}_{\perp}. \tag{2.12b}$$

From (2.12a) and (2.8b), it follows that

$$\mathbf{v}_{\parallel}U^{\dagger} = U\mathbf{v}_{\parallel}, \tag{2.13a}$$

$$\mathbf{v}_{\perp}U^{\dagger} = U^{\dagger}\mathbf{v}_{\perp}. \tag{2.13b}$$

So (2.8a) yields

$$\mathbf{v}'' = U(\mathbf{v}_{\parallel} + \mathbf{v}_{\perp})U^{\dagger} = U^2\mathbf{v}_{\parallel} + \mathbf{v}_{\perp} = e^{\theta}\mathbf{v}_{\parallel} + \mathbf{v}_{\perp}. \tag{2.15}$$

Note that

$$e^{\theta}\mathbf{v}_{\parallel} = \mathbf{v}_{\parallel}\cos|\theta| + \mathbf{i} \cdot \mathbf{v}_{\parallel}\sin|\theta|. \tag{2.16}$$

is a rotation by $|\theta|$, since $\mathbf{i}\mathbf{v}_{\parallel} = \mathbf{i} \cdot \mathbf{v}_{\parallel}$ is just $\mathbf{v}_{\parallel}$ rotated by $90°$ in the $\mathbf{i}$ plane.

The mapping (2.8) is simply a *rotation* by θ [i.e., a rotation by $|\theta|$ in the (oriented) $\mathbf{i}$ plane]. It will be proved below that every rotation can be expressed in the form (2.8). So the approach used above to arrive at (2.8) proves not only that the

composition of two reflections is equivalent to a rotation, but that every rotation can be expressed as a product of two reflections. Note that the fundamental rule (1.1) is responsible for the ease with which this result has been established and expressed.

The occurrence of half-angles in the fundamental expression (2.8) for a rotation is, of course, related to the bilinearity in U. The bilinear expression performs the function of separating, by virtue of the "commutation relations" (2.12), $\mathbf{v}_\parallel$ from $\mathbf{v}_\perp$ to give (2.15). As (2.16) shows, the bilinear form is not needed to represent rotation of vectors which lie entirely in the θ plane.

For obvious reasons, the bivector $\mathbf{i}$ is called the *generator of rotations* in the $\mathbf{i}$ plane. This name designates a geometrical property that was not ascribed to bivectors when they were first introduced. Thus, either or both of two fundamental geometrical interpretations can be ascribed to any bivector: an *operational* interpretation as the generator of rotations, or an *"extensive"* interpretation as an oriented plane segment. Grassmann based his "algebra of extension" on the extensive view, while Hamiltonian based his quaternions on the operational view. The rule (1.1) fuses these two complementary views into a single system.

It is worthwhile to observe that (2.15) expresses the solution of an eigenvalue problem. Both $\mathbf{v}_\parallel$ and $\mathbf{v}_\perp$ are eigenvectors of the linear transformation (2.8). Thus

$$\mathbf{v}_\perp \rightarrow \mathbf{v}'_\perp = U\mathbf{v}_\perp U^\dagger = \mathbf{v}_\perp\,, \qquad (2.17)$$

$$\mathbf{v}_\parallel \rightarrow \mathbf{v}'_\parallel = U\mathbf{v}_\parallel U^\dagger = \exp\left(\mathbf{i}\,|\,\theta\,|\right)\mathbf{v}_\parallel\,. \qquad (2.18)$$

The eigenvalue 1 corresponds to the eigenvector $\mathbf{v}_\perp$, and the "complex" eigenvalue $\exp\left(\mathbf{i}\,|\,\theta\,|\right)$ corresponds to the eigenvector $\mathbf{v}_\parallel$. This suggests that in the general theory of linear transformations complex eigenvalues can always be associated with rotations. The matrix approach to linear transformations fails to reveal the fact the "unit imaginary" which is needed to express a complex root of the characteristic equation has a geometric significance; (2.18) indicates that it can be interpreted as the generator of rotations in the plane of eigenvectors corresponding to the complex eigenvalue. A systematic exploitation of this fact ought to enrich the theory of linear transformations.

A rotation takes an orthonormal set of vectors $\{\boldsymbol{\sigma}_i\}$ into an orthonormal set $\{\mathbf{e}_i\}$. The $\mathbf{e}_i$ can be expressed as a linear combination of the $\{\boldsymbol{\sigma}_i\}$, that is,

$$\mathbf{e}_i = a_{ij}\boldsymbol{\sigma}_j\,. \qquad (2.19)$$

The rotation is completely determined bv specifying the $\{\boldsymbol{\sigma}_i\}$ and either the $\{\mathbf{e}_i\}$ or the rotation matrix a_{ij}. This representation of a rotation is evidently inferior to that displayed by (2.8). For the geometric $U = \exp\left(\frac{1}{2}\theta\right)$ describes the rotation completely and directly in terms of the angle and plane of rotation, whereas the matrix representation requires reference to a basis which typically has no significance.

Just the same, the matrix a_{ij} is sometimes a given quantity, so the problem of finding an equivalent U arises. This can be solved by applying (2.8) to the $\{\boldsymbol{\sigma}_i\}$ to get

$$\mathbf{e}_i = U\boldsymbol{\sigma}_i U^\dagger\,. \qquad (2.20)$$

This equation can be solved for U. First observe that since $\sigma_i\sigma_i = 3$ and $\boldsymbol{\sigma}_i B\boldsymbol{\sigma}_i = -B$ for any bivector B,

$$\boldsymbol{\sigma}_i U^\dagger \boldsymbol{\sigma}_i = \boldsymbol{\sigma}_i(U_S - U_B)\boldsymbol{\sigma}_i = 3U_S + U_B = 4U_S - U^\dagger \,. \qquad (2.21)$$

where U_S is the scalar part of U. Hence, by multiplying (2.20) on the right by $\boldsymbol{\sigma}_i$ and summing, one finds

$$4UU_S = 1 + M \,, \qquad (2.22)$$

where

$$M \equiv \mathbf{e}_i\boldsymbol{\sigma}_i = a_{ij}\boldsymbol{\sigma}_j\boldsymbol{\sigma}_i \,. \qquad (2.23)$$

Multiplying (2.22) by its adjoint, one obtains

$$16U_S^2 = (1 + M)(1 + M^\dagger) \,.$$

This quantity is a positive scalar so there is no problem in taking its square root and using the result to eliminate U_S from (2.22). Thus

$$U = \pm(1 + M)/[(1 + M)(1 + M^\dagger)]^{1/2} \,. \qquad (2.24)$$

Actually, an expression for U which is even simpler than (2.24) can be obtained by solving the scalar part of (2.22) directly for U_S; thus

$$4U_S^2 = 1 + M_S = 1 + a_{kk} \,. \qquad (2.25)$$

So,

$$U = \pm\tfrac{1}{2}(1 + M)/(1 + M_S)^{1/2} \,. \qquad (2.26)$$

By virtue of the definition for M in (2.23), either of the equations (2.24) or (2.26) provides the desired expression for U in terms of the transformation matrix a_{ij} and the basis $\{\boldsymbol{\sigma}_i\}$. That the U in (2.24) can be written in the explicit forms given by (2.8b) follows from the fact that M is composed of scalar and bivector parts only, as (2.23) shows. The ambiguity in sign in (2.22) merely reflects the bilinearity of (2.8); the negative root can be ignored without danger.

The inverse problem of finding the transformation matrix from U and the $\boldsymbol{\sigma}_i$ is easily solved by using (2.19) and (2.20). Thus,

$$a_{ij} = \mathbf{e}_j \cdot \boldsymbol{\sigma}_j = (\mathbf{e}_j\boldsymbol{\sigma}_j)_S = (U\boldsymbol{\sigma}_i U^\dagger \boldsymbol{\sigma}_j)_S \,. \qquad (2.27)$$

This completes an explicit demonstration that (2.8) and (2.19) are equivalent descriptions of a rotation. Since any rotation can obviously be described by (2.19), it follows, as promised, that any rotation can be written in the form (2.8).

Even geometrics have a special geometric significance which the above discussion of rotations makes it possible to appreciate. To emphasize this significance, the word *spinor* will be used hereafter to designate an even geometric. Justification for this usage will appear in due course.

From its definition, it follows that any spinor can be expressed as the sum of a scalar S and a bivector B:

$$\psi = S + B\,. \tag{2.28}$$

The decomposition can be used to prove that

$$|\,\psi\,|^2 = \psi\psi^\dagger\,, \tag{2.29}$$

for

$$\begin{aligned}
\psi\psi^\dagger = (S+B)(S-B) &= S^2 - B^2 \\
&= S^2 + |\,B\,|^2 = (\psi\psi^\dagger)_S\,,
\end{aligned}$$

Now factor ψ into the product of a scalar $\lambda = |\,\psi\,|$ and a *unitary spinor* U:

$$\psi = \lambda U\,. \tag{2.30}$$

Clearly

$$UU^\dagger = U^\dagger U = 1\,. \tag{2.31}$$

In fact, it is a simple matter to show that U can be expressed in any of the forms given by (2.8b).

Now consider the following function of a vector $\mathbf{v}$:

$$\mathbf{v} \to \mathbf{v}' = \psi\mathbf{v}\psi^\dagger = \lambda^2 U\mathbf{v}U^\dagger\,. \tag{2.32}$$

Comparison with (2.8) shows that (2.32) consists of a rotation and a dilation by λ^2. Thus, by virtue of (2.32), every spinor uniquely determines a rotation dilatation. Therein lies the operational geometric significance of spinors.

The spinors form a subalgebra of geometric algebra. The algebra of spinors was discovered independent of the full geometric algebra by Hamilton, who gave it the name quaternion algebra. Some readers may want to say, rather, that we have here two isomorphic algebras, but there is no call for any such distinction. A quaternion is a spinor. The identification of quaternions with spinors is fully justified not only because they have equivalent algebraic properties, but more important, because they have the same geometric significance.

It is unfortunate that Hamilton chose the name quaternion, for this name merely refers to the comparatively insignificant fact that the quaternions compose a linear space of four dimensions. The name quaternion diverts attention from the key fact that Hamilton had invented a geometric algebra. Hamilton's work itself shows clearly the crucial role of geometry in his invention. Hamilton was consciously looking for a system of numbers to represent rotations in three dimensions. He was looking for a way to describe geometry by algebra, and so he found a geometric algebra.

In the twentieth century quaternions seem to be generally regarded as something of a mathematical museum piece, a curious artifact from the nineteenth century or

14

a diversion from the mainstream of modern mathematics. One reason for this unfortunate circumstance might be found in the fact that Frobenius proved in 1878 that the quaternions constitute the largest associative *division* algebra. Thus, if division is a characteristic feature of quaternions, then quaternion algebra cannot be generalized. This is a typical diversion from the geometrical spirit of Hamilton's inquiry. The predominance of such diversions from the geometric to the algebraic features of quaternions goes far to explain why quaternions have not yet found the central place in mathematics which Hamilton had forseen.

The treatment of reflections and rotations given above originated with Hamilton. Many opportunities to make use of Hamilton's incisive conceptions are missed, because quaternion algebra is something apart from the usual vector or tensor algebra. But this barrier is completely removed by adopting the more comprehensive geometric viewpoint of geometric algebra. Moreover, identification with spinors gives quaternions a prominent place in modern physics. Spinors are commonly held to be an invention of the twentieth century, the credit often going to Cartan, though he did not contribute to their emergence in physics. The relation of spinors to quaternions has remained rather obscure since their inception. The "new" definition of a spinor given here goes back over a hundred years to Hamilton, so it is actually the oldest definition of a spinor rediscovered and renamed.

When the Hamiltonian spirit is restored by interpreting quaternions as rotation dilatations, and infused with the spirit of Grassmann by recognizing its place in geometric algebra, then the appropriate generalization of quaternions is easily found to be as follows: The vectors of an n-dimension vector space $\mathcal{R}_n$ generate a geometric algebra $\mathcal{M}_n$ (Refs. [4], [2], [1]); spinors are identified as elements of the even subalgebra of $\mathcal{M}_n$, and they can be shown to represent rotation dilatations of $\mathcal{R}_n$ by a direct and simple generalization of the procedure used above. A reformulation of Dirac theory using this conception of spinor has already been carried out in Ref. [3]. Thus, the quaternion representing a rotation dilatation in three space is generalized to a spinor representing a rotation dilatation in n space. Of course, the fourness and the universal invertibility of quaternions is lost in this generalization, but the geometric meaning is retained, as is its most significant algebraic realization, namely, the unitary spinors form a group under multiplication which is homomorphic to the group of rotations in n dimensions. Thus, the unitary spinors characterize the rotations in n space simply and completely. This feature is so important that it is worthwhile to express it in another fashion: Every unitary spinor is a *number* which represents a rotation. The word number is used advisedly here, because multiplication has a significant meaning. The product of two such numbers to produce another directly and completely describes the composition of two rotations.

Though the quaternions are well known to many physicists, they have been regarded, heretofore, as something different than spinors. So the relation of the quaternions to other representations of spinors must be explained.

In Ref. [2] spinors $\{\phi_+\}$ are identified with elements of minimal left ideals of geometric algebra. Each ϕ_+ is related to a quaternion ψ by the formula

$$\phi_+ = \psi(1 + \boldsymbol{\sigma}_3)(2)^{-1/2}, \tag{2.33}$$

15

where $\boldsymbol{\sigma}_3$ is a fixed but definite unit vector. The trouble with the "ideal" approach is that ψ and $\boldsymbol{\sigma}_3$ have different geometric significance, but ϕ_+ mixes them up, mixes them so well, in fact, that even authors well versed in spinor lore have failed to realize their existence and disentangle them. In Ref. [3] the separation of ψ from $\boldsymbol{\sigma}_3$ was first carried out and shown to have physical significance in the Dirac theory. The roles of ψ and $\boldsymbol{\sigma}_3$ in the Pauli theory are thoroughly discussed in II [1].

The representation of spinors as ideal elements is almost identical to their representation as column matrices; the connection is discussed in Ref. [2]. However, the correspondence of a quaternion ψ with a column spinor Ψ is most easily accomplished by means of the following algorithm: Express ψ as a polynomial of the basis vectors $\{\boldsymbol{\sigma}_k\}$; then replace the $\{\boldsymbol{\sigma}_k\}$ by their matrix representations (1.36). Define the column matrix u by

$$\boldsymbol{\sigma}_3 u = u, \qquad u = \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \tag{2.34}$$

Then, regarding ψ as a matrix operator, Ψ is given by

$$\Psi = \psi u . \tag{2.35}$$

This enables one to obtain Ψ from ψ. To obtain ψ from Ψ is more difficult. Further details are given in II [1].

One sometimes finds in the literature the cryptic assertion that spinors are more fundamental than vectors, because a spinor is a kind of square root of a vector. This view reveals an incomplete grasp of the geometric meaning of spinors. Spinors cannot be defined without reference to vectors. As has been shown above, the geometric meaning of a spinor is operational: it transforms one vector (or frame of vectors) into another. The resulting vector (or frame) is a bilinear function (a kind of square) of the spinor as in (2.30), but it also depends on the choice of the "initial" vector (or frame) to be transformed. The dependence of column spinors of a specification of an "initial" frame of vectors is subtly disguised (though, of course, not on purpose). It is certainly not apparent in (2.33), but (2.33) has no significance apart from a specification of the matrix representations of the $\{\boldsymbol{\sigma}_k\}$. In this covert manner the $\{\boldsymbol{\sigma}_k\}$ have been chosen implicitly as an "initial" frame. In physical applications, the spinor ψ (or Ψ) transforms $\{\boldsymbol{\sigma}_k\}$ into some vector with physical significance, such as the spin. By maintaining an explicit separation of the geometrically distinct elements of spinor theory from the beginning, the "operational" approach to spinors recommended in this paper broadens the range, simplifies the application, and enhances the geometric content of the mathematical apparatus.

3 SPINORS IN CLASSICAL MECHANICS

Geometric calculus makes it possible to take full advantage of spinors in classical mechanics. Since unitary spinors are numbers which directly represent rotations, they can be used to advantage whenever rotations play a significant role in a physical

problem. A wider use of spinors in classical mechanics ought to dispel the pervasive mistaken impression that spinors are a special feature of quantum mechanics. As a simple significant example what should be one of the most common applications of spinors is discussed in this section.

Let the vector $\mathbf{x}$ which designates a point of physical space be called the *coordinate* of the point. A coordinate is more than a name for a point; it is an element of a geometric algebra which is a kind of special language invented to express relations among "physical points" in a most efficient manner. Let a specified procedure which (in principle) puts the points of physical space (or some subset of them, at least) in one to one correspondence with coordinates be called a *coordinate system*. Write $\mathbf{x} = \mathbf{x}(t)$ to denote that the correspondence may depend on the time t. Any smooth invertible function $f(\mathbf{x})$ determines a new system of coordinates $\mathbf{x}'$ by the rule

$$f: \quad \mathbf{x} \to \mathbf{x}' = f(\mathbf{x}), \quad \mathbf{x} = f^{-1}(\mathbf{x}'). \tag{3.1}$$

In practical problems, one frequently encounters coordinate systems related by a rotation. In that case the function $f(\mathbf{x})$ can be written explicitly in the form

$$\mathbf{x}' = U\mathbf{x}U^{\dagger}. \tag{3.2}$$

Here $U = U(t)$ is a unitary spinor which may be a function of time.

The aim of this section is to show how to use the spinor U to relate dynamics in the $\mathbf{x}$ system to dynamics in the $\mathbf{x}'$ system. The key to the whole approach is the formula

$$dU/dt = \tfrac{1}{2}\Omega U, \tag{3.3}$$

where Ω is a bivector.

First it is necessary to show that Ω is, in fact, a bivector. To this end $UU^{\dagger} = 1$ may be used to solve (3.3) for Ω:

$$\Omega = 2(dU/dt)U^{\dagger}. \tag{3.4}$$

Differentiation of $UU^{\dagger} = 1$ yields

$$(dU/dt)U^{\dagger} = -U(dU^{\dagger}/dt) = -[(dU/dt)U^{\dagger}]^{\dagger},$$

so

$$\Omega = -\Omega^{\dagger}. \tag{3.5}$$

Since U is even, by (3.4) Ω is even. And, by (3.5) Ω cannot have a scalar part. Hence Ω is a bivector.

It is appropriate to call Ω the *angular velocity* of the rotation U, though this name is usually reserved for its dual

$$\omega = i\Omega. \tag{3.6}$$

The bivector Ω is more fundamental than the vector ω, because it describes the rate of rotation *in a plane*, while an equivalent description by ω as the rate of rotation *about an axis* is possible only in a space of three dimensions. However, the intuition of physicists is more highly developed for vectors than for bivectors, and of course, ω is more common in the literature. So both characterizations are used here, especially to emphasize how easy it is to go from one to the other with (3.6).

For some readers it may not be superfluous to remark that Ω is equal to the time rate of change of angle if and only if the plane of rotation is a constant; that is, for $U = \exp\left(\frac{1}{2}\theta\right)$, (3.3) is equivalent to $\Omega = d\theta/dt$ if and only if

$$\frac{d(\theta/|\,\theta\,|)}{dt} = 0\,.$$

It is often convenient to express the constancy of the plane of rotation by $\Omega\theta = \theta\Omega$.

Now observe that the adjoint of (3.3) is

$$dU^\dagger/dt = \tfrac{1}{2}U^\dagger\Omega^\dagger = -\tfrac{1}{2}U^\dagger\Omega\,. \tag{3.7}$$

So, with dots for time derivatives, the time derivative of (3.2) is

$$\dot{\mathbf{x}}' = U\dot{\mathbf{x}}U^\dagger + \dot{U}\mathbf{x}U^\dagger + U\mathbf{x}\dot{U}^\dagger$$
$$= U(\dot{\mathbf{x}} + U^\dagger\dot{U}\mathbf{x} + \mathbf{x}\dot{U}^\dagger U)U^\dagger$$
$$= U[\dot{\mathbf{x}} + \tfrac{1}{2}(\Omega\mathbf{x} - \mathbf{x}\Omega)]U^\dagger$$
$$= U[\dot{\mathbf{x}} - i\tfrac{1}{2}(\omega\mathbf{x} - \mathbf{x}\omega)]U^\dagger$$

Recalling (1.12) and (1.30), this can be written

$$\dot{\mathbf{x}}' = U(\dot{\mathbf{x}} + \Omega\cdot\mathbf{x})U^\dagger$$
$$= U(\dot{\mathbf{x}} + \omega\times\mathbf{x})U^\dagger \tag{3.8}$$

With the initial condition at time t,

$$U(t) = 1 \quad\text{or}\quad \mathbf{x}'(t) = \mathbf{x}(t)\,, \tag{3.9}$$

(3.8) becomes

$$\dot{\mathbf{x}}' = \dot{\mathbf{x}} + \Omega\cdot\mathbf{x} = \dot{\mathbf{x}} + \omega\times\mathbf{x} \tag{3.10}$$

Differentiation of (3.8) yields

$$\ddot{\mathbf{x}}' = U[\ddot{\mathbf{x}} + \dot{\Omega}\cdot\mathbf{x} + \Omega\cdot\dot{\mathbf{x}} + \Omega(\dot{\mathbf{x}} + \Omega\cdot\mathbf{x})]U^\dagger$$
$$= U[\ddot{\mathbf{x}} + \dot{\Omega}\cdot\mathbf{x} + 2\Omega\cdot\dot{\mathbf{x}} + \Omega\cdot(\Omega\cdot\mathbf{x})]U^\dagger$$
$$= U[\ddot{\mathbf{x}} + \dot{\omega}\times\mathbf{x} + 2\omega\times\dot{\mathbf{x}} + \omega\times(\omega\times\mathbf{x})]U^\dagger \tag{3.11}$$

With the initial condition (3.9), this becomes

$$\ddot{\mathbf{x}}' = \ddot{\mathbf{x}} + \dot{\Omega}\cdot\mathbf{x} + 2\Omega\cdot\dot{\mathbf{x}} + \Omega\cdot(\Omega\cdot\mathbf{x})$$
$$= \ddot{\mathbf{x}} + \dot{\omega}\times\mathbf{x} + 2\omega\times\dot{\mathbf{x}} + \omega\times(\omega\times\mathbf{x}) \tag{3.12}$$

The vector forms of (3.10) and (3.12) are, of course, well known and often used. The matrix equivalent of (3.3) is also used, though less frequently than it deserves. What is new here is the simple connection between these equations made possible by the comprehensive geometric approach of geometric calculus. This unifies vector and spinor desciptions of rotating systems into a single simple approach. And it may not be superfluous to point out that, in contrast to other approaches, no arbitrary vector basis or set of scalar coordinates is needed to carry out computations. Moreover, as shown in II [1], this unified approach applies to quantum as well as to classical physics.

As a typical illustration of the missed opportunities to exploit spinors in classical mechanics the chapter on rigid body motion in the widely used textbook by Goldstein [7] may be cited. Goldstein formulates Eq. (3.2) with the customary clumsy matrix approach, but he does not use it to obtain (3.10) and (3.12), or anything else for that matter. Instead, to get (3.10) and (3.12) he describes rotations with 3×3 matrices and then translates some of the results into vector notation. Altogether, the chapter is a jumble of three different descriptions of rotations only loosely tied together. Moreover, it is unnecessarily restricted in generality. It does not apply unless the plane of rotation is constant, and it implicitly assumes the initial condition (3.9) with no hint of how this condition can be relaxed.

4 COMPLEX NUMBERS IN ELECTRO-DYNAMICS

The electromagnetic field can be described by eight equations for six scalars, or four equations for two vectors, or two equations for an antisymmetric tensor, or as shown in Ref. [2], by a single equation for a single geometric. The last of these formulations is simpler than the others not just because it deals with only one equation, but more because geometric calculus has a more comprehensive geometric significance than the usual vector or tensor calculus. To illustrate this point, the role of complex numbers in electrodynamics is briefly discussed here.

Every physicist is aware of the great utility of complex numbers for describing electromagnetic waves. For this reason alone, in vector and tensor calculus the electromagnetic field is commonly represented *ad hoc* by complex quantities. The phrase *ad hoc* is advisable here, because only the real part of a complex quantity is purported to have physical significance. In contrast, when the electromagnetic field is described by geometric calculus, "complex quantities" arise naturally with definite geometric and physical meanings. Thus it appears that the utility of complex numbers is something more than an accident.

As shown in Ref. [2], *Maxwell's Equation* for the electromagnetic field in empty space can be written

$$(c^{-1}\partial_t + \nabla)F = 0 \,, \tag{4.1}$$

where

$$F = \mathbf{E} + i\mathbf{B} \,. \tag{4.2}$$

The curious convenience of combining the electric and magnetic vectors $\mathbf{E}$ and $\mathbf{B}$ into a single complex vector has been noted time and again for more than 50 yr but with no inkling of its geometric significance. The expression (4.2) is actually more than a "complex vector"; for i is more than a "unit imaginary"; it is the unit pseudoscalar introduced in Sec. 1, and it appears in (4.2) because the magnetic field is correctly described by the bivector $i\mathbf{B}$ and not its dual $\mathbf{B}$. Representation of the magnetic field by a vector dates back to the days when the difference between a vector and a bivector was not appreciated.

With the geometric significance of i clear, it is apparent that the physical separation of the complete electromagnetic field F into electric and magnetic parts corresponds exactly to the geometric separation into vector and bivector parts.

The limited geometric view of the usual form of vector calculus has preserved an artificial separation of electric and magnetic fields, though the fact that they compose a single physical entity has long been understood. This has helped to disguise the significance of complex numbers in electrodynoamics. Today, electric and magnetic fields are commonly represented as complex vectors without realizing that the real part of one can be related to the imaginary part of the other, so that only the single "complex vector" (4.2) is needed for a complete description of the field. This more compact description of the electromagnetic field also yields solutions with a more direct geometric significance. A study of plane waves brings this out nicely.

No physicist will be surprised to be told that

$$F_+(\mathbf{x}, t) = f \exp\left[\, i(\omega t - \mathbf{k} \cdot \mathbf{x})\,\right] \qquad (4.3)$$

is a solution of Maxwell's equation describing a monochromatic plane wave with frequency $\omega > 0$ and propagation vector $\mathbf{k}$. However, he may be surprised to learn that this wave is necessarily *right circularly polarized*, because that follows unequivocally only from the geometric meaning of geometric algebra. To establish the properties of (4.3), substitute it in (4.1) and use the facts that f, ω, and k are constant to get Maxwell's equation in the form

$$(\omega/c - \mathbf{k})F = 0 \,. \qquad (4.4)$$

Multiplication by $(\omega/c + \mathbf{k})$ gives

$$(\omega^2/c^2 - \mathbf{k}^2)F = 0 \,.$$

Since F is not zero, this implies that $|\mathbf{k}| = \omega/c$, as expected. So (4.4) can be written in the simpler form

$$\hat{\mathbf{k}}F = F \,. \qquad (4.5)$$

Substitution of (4.2) in (4.5) shows that both the even and odd parts of (4.5) give the same relation among the vectors $\mathbf{k}$, $\mathbf{E}$ and $\mathbf{B}$, namely,

$$\hat{\mathbf{k}}\mathbf{E} = i\mathbf{B} \,. \qquad (4.6)$$

20

Equation (4.6) can be used to eliminate the magnetic field from (4.2) giving

$$F = (1 + \hat{\mathbf{k}})\mathbf{E} = \mathbf{E}(1 - \hat{\mathbf{k}}). \tag{4.7}$$

This agrees with (4.5) since obviously $\hat{\mathbf{k}}(1 + \hat{\mathbf{k}}) = (1 + \hat{\mathbf{k}})$. The square of (4.6) gives immediately $\mathbf{E}^2 = \mathbf{B}^2$, so (4.6) can be written

$$\hat{\mathbf{E}}\hat{\mathbf{B}}\hat{\mathbf{k}} = i. \tag{4.8}$$

Comparison with (1.8) reveals that $\{\hat{\mathbf{E}}, \hat{\mathbf{B}}, \hat{\mathbf{k}}\}$, in that order, compose a right-handed orthornormal frame of vectors.

To make the time dependence of the field explicit, first note that

$$F(0,0) = f = \mathbf{E}_0 + i\mathbf{B}_0 = \mathbf{E}_0(1 - \hat{\mathbf{k}}). \tag{4.9}$$

Also note that

$$(1 - \hat{\mathbf{k}}) \exp(i\omega t) = (1 - \hat{\mathbf{k}}) \exp(-i\hat{\mathbf{k}}\omega t).$$

So at any point on the plane $\mathbf{x} \cdot \mathbf{k} = 0$, the field is

$$F = \mathbf{E}(t) + i\mathbf{B}(t) = f \exp(i\omega t)$$
$$= \mathbf{E}_0(1 - \hat{\mathbf{k}}) \exp(i\omega t) = \mathbf{E}_0(1 - \hat{\mathbf{k}}) \exp(-i\hat{\mathbf{k}}\omega t). \tag{4.10}$$

The vector part of this expression is

$$\mathbf{E}(t) = \mathbf{E}_0 \exp(-i\hat{\mathbf{k}}\omega t)$$
$$= \mathbf{E}_0(\cos\omega t - i\hat{\mathbf{k}}\sin\omega t)$$
$$= \mathbf{E}_0 \cos\omega t - \mathbf{B}_0 \sin\omega t. \tag{4.11}$$

This shows explicitly that as t increases the electric vector rotates *clockwise* in the plane as viewed by an observer facing the oncoming wave train. The magnetic vector follows 90° behind. Thus the usual picture of a circularly polarized wave arises, but the "unit imaginary" in the complex solution generates the rotation of $\mathbf{E}$ and $\mathbf{B}$.

To obtain the *left circularly polarized* solution F_- it is only necessary to change the orientation of the "generator" in (4.3). Thus

$$F_-(\mathbf{x}, t) = f \exp\left[-i(\omega t - \mathbf{k} \cdot \mathbf{x})\right]. \tag{4.12}$$

In both (4.3) and (4.12) the frequency is positive. Instead, it is convenient to allow the frequency to take both positive and negative values. The sign of the frequency is then associated with the polarization of the wave, F_+ being the positive frequency solution and F_- being the negative frequency solution. This is exactly the relation of frequency to polarization found in the theory of the neutrino.

In general, a wave packet propagating in the $\hat{\mathbf{k}}$ direction can be written in the form

$$F(\mathbf{x}, t) = F(s) = f z(s), \tag{4.13}$$

where

$$s = t - \mathbf{x} \cdot \hat{\mathbf{k}}/c \,, \tag{4.14}$$

$$f = (1 + \hat{\mathbf{k}})\mathbf{e} \,, \tag{4.15}$$

$$z(s) = \int_{-\infty}^{\infty} d\omega \, \alpha(\omega) \, e^{i\omega s} \,,$$

$$= \int_{0}^{\infty} d\omega \left[\alpha_{+} e^{i\omega s} + \alpha_{-} e^{-i\omega s} \right], \tag{4.16}$$

$$\alpha_{\pm}(\omega) = \alpha(\pm|\,\omega\,|) \,. \tag{4.17}$$

In (4.15), $\mathbf{e}$ can be any constant unit vector orthogonal to $\hat{\mathbf{k}}$; hence

$$|f|^2 = 2 \,. \tag{4.18}$$

The quantity $\alpha(\omega)$ is "formally complex" (i.e., it has only scalar and pseudoscalar parts). A change in the selection of the vector $\mathbf{e}$ is equivalent to a change in the overall phase of α. The energy density of the field is

$$\tfrac{1}{2}|\,F\,|^2 = \tfrac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) = |\,z\,|^2 \,. \tag{4.19}$$

The quantities $\alpha_{+}(\omega)$ and $\alpha_{-}(\omega)$ describe the components of the packet with frequency $|\,\omega\,|$ and, respectively, right and left polarization. Thus it is clear that the salient properties of the packet are described by the "complex" function $z(s)$ or its Fourier transform $\alpha(\omega)$.

The above discussion serves to illustrate how the geometric approach describes the electromagnetic field by a single quantity which can be readily analyzed to reveal its salient physical features. In contrast to more usual approaches, it is not compelled to break up the field into electric and magnetic pieces. Thus, despite the picturesqueness of rotating vectors, the decomposition of a propagating wave into electric and magnetic parts holds little of physical interest; the decomposition into orthogonal polarizations is more significant, and the multivector approach accomplishes this forthrightly. When "complex quantities" are appropriate they arise automatically and with a geometric interpretation. One example is given above; another is given in the discussion of quantum mechanics in II [1]. Is there a connection between them? That is a physical question, so it calls for a physical analysis.

References

[1] D. Hestenes and R. Gurtler, *Amer. J. Phys.* **39**, 1028 (1971) (following paper).

[2] D. Hestenes, *Space-Time Algebra*, (Gordon and Breach, New York, 1966).

[3] D. Hestenes, *J. Math. Phys.* **8**, 798 (1967).

[4] D. Hestenes, *J. Math. Anal. Appl.* **24**, 313 (1968).

[5] D. Hestenes, *J. Math. Anal. Appl.* **24**, 476 (1968).

[6] H. Gelman, *Amer. J. Phys.* **38**, 599 (1970).

[7] H. Goldstein, *Classical Mechanics*, (Addison-Wesley, Reading, Mass., 1950).