Applications of the Pauli algebra and other geometric algebras.

Shazia. Hadi
University of Windsor

Applications of the Pauli Algebra and other Geometric Algebras

by

Shazia Hadi

A Thesis
Submitted to the College of Graduate Studies and Research
through the Department of Physics
in Partial Fulfillment of the Requirements for
the Degree of Master of Science at the
University of Windsor

Windsor, Ontario, Canada
September 1999

0-612-52560-0

973540

Abstract

Relationships among the complex numbers, quaternions, and the Pauli algebra are developed by presenting them as geometrical (Clifford) algebras. Rotations are examined using both quaternions and the Pauli algebra, and in particular, algorithms that are used in three-dimensional simulations and video games are formulated in the Pauli algebra. Relativity is presented using a number of formalisms, and the treatment of De Leo and Rotelli is clarified. The relationship between spinors and spacetime vectors is explored using the Pauli algebra. Dirac theory is exhibited using the Pauli algebra, and neutrino oscillations are discussed.

Acknowledgements

I would like to thank my supervisor, Dr. W. E. Baylis, for providing me with encouragement and inspiration, and for his tremendous patience. Also, I would like to thank my officemate Jacob Alexander for numerous discussions, as well as all my teachers and fellow students for the things that I have learned from them. Finally, a very important thanks goes to my parents, and sister and brothers for their continuous support throughout this long journey.

Contents

Chapter 1

Introduction

The complex numbers, quaternions, and the complex quaternions are all important examples of associative algebras that have found extensive applications in theoretical physics. These are all examples of geometric or Clifford algebras. A geometric algebra can also be formed from vectors in space and their products. This geometric algebra of vectors, the Pauli algebra, is isomorphic to the algebra of complex quaternions, and contains the algebras of the quaternions and complex numbers as subalgebras. Therefore, all applications of the complex numbers, quaternions, and the complex quaternions can be formulated within the Pauli algebra. By doing this, the geometrical significance of these applications is emphasized. Also, spinors reside naturally in the Pauli algebra, and unlike the real quaternions, the Pauli algebra is quite useful for applications in relativity. The complex quaternions can also be used for relativity, but they do not provide the geometrical and physical insight provided by the Pauli algebra. This thesis explores these ideas as follows.

Chapter 2 introduces the quaternions from a historical perspective, uses quaternions to represent reflections and rotations in space, and complexifies them for later use in relativity. The third chapter starts with a presentation of general geometric algebras as Clifford algebras, and then moves on to the specific examples of the quaternions and the Pauli algebra as geometric algebras. Special relativity is presented in the next chapter. Formulations are given using: the complex quaternions; operators on the real quaternions; and the Pauli algebra. Chapter 5 gives a Pauli algebra development of spinors and relates this development to the standard presentations of 2-component spinors and 4-component Dirac spinors. These spinor developments are used in chapter six to write the Dirac equation in the Pauli algebra, which is then used to explore the possibilities of massless and massive neutrinos.

In this thesis units are used such that $c = \hbar = 1$. Lower-case Latin letters are used for spatial indices that take on the values 1,2,3; lower-case Greek letters are used for spacetime indices that take on the values 0,1,2,3; and upper case Latin letters are used for indices spinor that have values 0,1. Any index that is repeated in a product is to be summed over the appropriate range of values. Vectors and spacetime vectors are denoted with upper-case letters when represented by quaternions and complex quaternions respectively.

Chapter 2

Quaternions

2.1 Semi-historical Introduction

An extraordinary invention in mathematics was the algebra of quaternions by Sir William Rowan Hamilton[1]. Quaternions were created by extending the complex numbers. Hamilton had been interested in complex numbers since the early 1830's. By 1830, complex numbers were intuitively well understood through their representations as points or as directed line segments in the plane, but this intuitive foundation was not satisfactory to Hamilton. He was more interested in basing the complex numbers on the logic of arithmetic. In 1833, he was the first one to show that complex numbers can be viewed as merely ordered pairs of real numbers with rules for products. He also pointed out that a complex number $a + bi$ is NOT a genuine sum of two (real) numbers, since bi cannot be added to a. Hamilton incorporated the peculiar $i = \sqrt{-1}$ in the definition of operations with ordered pairs. In Hamilton's interpretation of complex numbers the usual associative, commutative and distributive properties are logically founded on the basis of real numbers. The geometrical representation of complex numbers given by

Wessel, Argand and Gauss, when combined with Hamilton's algebraic interpretation, gives an algebra for vectors and vector operations in a plane. The power of complex numbers is thus realized as a tool for handling vectors algebraically, which is sometimes much easier than performing operations geometrically, but at the same time it is limited to vectors in the plane.

Hamilton wanted to find an algebra capable of handling operations with vectors in three-dimensional space. Guided by the example of ordered pairs and complex numbers, he hoped to find an algebra for ordered triples of real numbers. Hence the search for the extension of complex numbers to the so-called hypercomplex numbers and its algebra was started. Hypercomplex numbers were required to do every thing that complex numbers can do, i.e., they must all have binary operations and obey the usual associative, commutative and distributive laws so that algebraic operations could be performed effectively.

For the next ten years Hamilton worked very hard, if not obsessively, with the problem of finding the hypercomplex numbers that could be used to represent multiplication of vectors in three-dimensional space. On Monday, the 16th of October 1843, one of the best documented days in the history of mathematics, Hamilton found this new number system, which he called quaternions. He found the quaternions after recognizing that the multiplication that he needed was not possible with triples of real numbers. Hamilton realized that his new numbers each had to have four components, and that he had to give up the commutative multiplication law satisfied by both the real and complex numbers. Both of these features were radical for mathematics. One real unit and three imaginary units, **i**, **j**, **k**, were needed and they had to have the following relations:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1 \tag{2.1}$$

$$\mathbf{ijk} = -1. \tag{2.2}$$

There is an obvious geometrical reason why quaternions have to have four components, namely because a quaternion can be regarded as an operator that rotates a given vector about a given axis in the space and also stretches or contracts the vector. There are two parameters needed to specify the fixed axis of rotation, one parameter to specify the angle of rotation, and a fourth parameter is needed to stretch or contract the given vector.

A Hamilton quaternion A is written as

$$A = a\mathbf{1} + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, \tag{2.3}$$

where the coefficients a, x, y, and z are all real. Consequently, in modern terminology, the quaternions are said to form a four-dimensional real vector space. The set $\mathbb{H}$ of all quaternions forms an associative algebra[1] when distributivity is assumed. If $A \neq 0$ and

$$B = \frac{a\mathbf{1} - x\mathbf{i} - y\mathbf{j} - z\mathbf{k}}{a^2 + x^2 + y^2 + z^2}, \tag{2.4}$$

then $AB = BA = 1$. Hence, $\mathbb{H}$ is a division algebra. The only thing that prevents the quaternions from forming a field is that multiplications are not commutative. Therefore, they can be used as the components over which quaternionic vector spaces are formulated. In particular, quaternionic versions of Hilbert spaces can be constructed, and Adler[2] uses these to implement an extension of quantum theory. Adler's version of quantum

[1] Algebras are defined formally in the next chapter.

mechanics differs from standard quantum mechanics only for energies near the Plank scale.

If a is zero, A is said to be a pure quaternion, or a vector. Then $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are associated with the three mutually perpendicular directions e_1, e_2, e_3 in space. A vector $\mathbf{v} = v^1 e_1 + v^2 e_2 + v^3 e_3$ will be written as

$$\mathbf{V} = v^1 \mathbf{i} + v^2 \mathbf{j} + v^3 \mathbf{k}. \tag{2.5}$$

If $\mathbf{V}$ and $\mathbf{W}$ are two pure quaternions, it follows directly that in modern notation

$$\mathbf{VW} = -\mathbf{v} \cdot \mathbf{w} + \mathbf{V} \times \mathbf{W}, \tag{2.6}$$

where $\mathbf{V} \times \mathbf{W}$ is a pure quaternion with components $(\mathbf{V} \times \mathbf{W})^j = (\mathbf{v} \times \mathbf{w}) \cdot \mathbf{e}_j$. From this it follows that two non-zero vectors commute iff they are parallel, and anti-commute iff they are perpendicular.

After quaternions were introduced, there was a long controversy between Tait and other disciples of Hamilton on one side, and Gibbs and Heaviside on the other, over whether the full quaternion algebra, or just the vector parts, would prove most useful for physics. That vectors prevailed is obvious, as may be seen in practically any physics textbook today. By considering applications such as rotations, relativity, and relativistic quantum mechanics, this thesis hopes to show the usefulness of quaternions and their more general cousins, Clifford (geometric) algebras.

2.2 Rotations and Quaternions

Quaternions are very useful for rotating and reflecting vectors without using matrices. To see this, it is handy to have a geometrical picture of what a

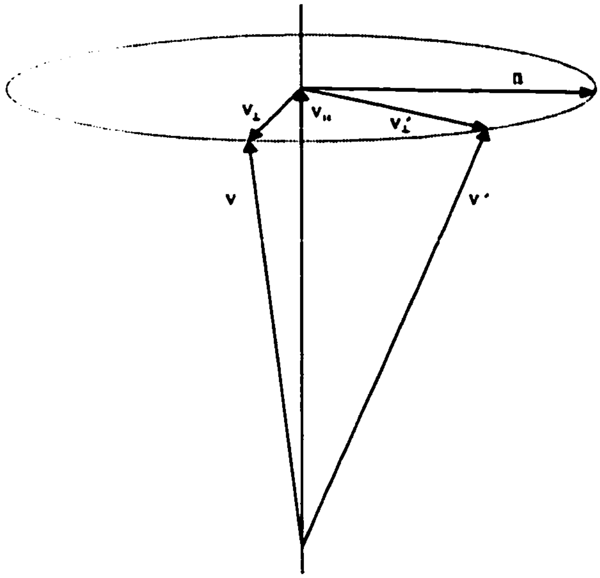

Figure 2.1.

rotation is. Let $\mathbf{v}$ be any vector that is to be rotated by the vector $\boldsymbol{\theta}$. This means that the unit vector $\widehat{\boldsymbol{\theta}}$ is the axis of rotation, and $\theta = \sqrt{\boldsymbol{\theta} \cdot \boldsymbol{\theta}}$ is the angle of rotation. A rotation of $\mathbf{v}$ by $\boldsymbol{\theta}$ means that the part of $\mathbf{v}$ parallel to $\boldsymbol{\theta}$, $\mathbf{v}_{\parallel} = \mathbf{v} \cdot \widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}$, remains unchanged, while the part of $\mathbf{v}$ perpendicular to $\boldsymbol{\theta}$, $\mathbf{v}_{\perp} = \mathbf{v} - \mathbf{v}_{\parallel}$, rotates in the plane perpendicular to $\boldsymbol{\theta}$. This plane is called the plane of rotation. Therefore, the rotated vector is $\mathbf{v}' = \mathbf{v}'_{\parallel} + \mathbf{v}'_{\perp}$, where $\mathbf{v}'_{\parallel} = \mathbf{v}_{\parallel}$. To find $\mathbf{v}'_{\perp}$, consider $\mathbf{n} = \widehat{\boldsymbol{\theta}} \times \mathbf{v}_{\perp}$, which is a vector perpendicular to both $\mathbf{v}$ and $\boldsymbol{\theta}$ that has the same length as $\mathbf{v}_{\perp}$. Also, $\mathbf{n}$ is lies in the plane of rotation.

From the geometry (see Figure 2.1), we can express the rotated vector $\mathbf{v}'_{\perp}$ in terms of the vectors $\mathbf{v}_{\perp}$ and $\mathbf{n}$:

$$\mathbf{v}'_\perp = \mathbf{v}_\perp \cos\theta + \mathbf{n}\sin\theta. \tag{2.7}$$

Consequently,

$$\mathbf{v}' = \mathbf{v}\cdot\widehat{\theta\theta} + \mathbf{v}_\perp\cos\theta + \widehat{\boldsymbol{\theta}}\times\mathbf{v}_\perp\sin\theta. \tag{2.8}$$

Rotations in three dimensional space can also be generated by the composition of two plane reflections. A reflection of a vector $\mathbf{v}$ through a plane proceeds as follows. Let $\widehat{\mathbf{n}}$ be the unit vector normal to the plane, and decompose $\mathbf{v}$ into parts parallel and perpendicular to $\widehat{\mathbf{n}}$. A reflection changes the sign of the part of $\mathbf{v}$ parallel to $\widehat{\mathbf{n}}$ (perpendicular to the plane), while leaving the part perpendicular to $\widehat{\mathbf{n}}$ (lying in the plane) unchanged. Now consider the quaternionic expression

$$\begin{aligned}
\widehat{\mathbf{N}}\mathbf{V}\widehat{\mathbf{N}} &= \widehat{\mathbf{N}}\,(\mathbf{V}_\parallel + \mathbf{V}_\perp)\,\widehat{\mathbf{N}} \\
&= (\mathbf{V}_\parallel - \mathbf{V}_\perp)\,\widehat{\mathbf{N}}\widehat{\mathbf{N}} \\
&= -\mathbf{V}_\parallel + \mathbf{V}_\perp \\
&= \mathbf{V} - 2\mathbf{V}_\parallel
\end{aligned}\; . \tag{2.9}$$

Therefore, $\mathbf{V}' = \widehat{\mathbf{N}}\mathbf{V}\widehat{\mathbf{N}}$ is the reflection of $\mathbf{v}$ through the plane that has normal $\widehat{\mathbf{n}}$.

Now consider two successive reflections of $\mathbf{v}$, first through a plane with normal $\hat{\mathbf{a}}$, and then through a plane with normal $\hat{\mathbf{b}}$. Thus, $\mathbf{v} \rightarrow \mathbf{v}'$, given by

$$\begin{aligned}
\mathbf{V}' &= \widehat{\mathbf{B}}\widehat{\mathbf{A}}\mathbf{V}\widehat{\mathbf{A}}\widehat{\mathbf{B}} \\
&= -\left(\widehat{\mathbf{A}}\cdot\widehat{\mathbf{B}} + \widehat{\mathbf{A}}\times\widehat{\mathbf{B}}\right)\mathbf{V}\left(-\widehat{\mathbf{A}}\cdot\widehat{\mathbf{B}} + \widehat{\mathbf{A}}\times\widehat{\mathbf{B}}\right) \\
&= -\left(\cos\theta + \widehat{\Theta}\sin\theta\right)(\mathbf{V}_\parallel + \mathbf{V}_\perp)\left(-\cos\theta + \widehat{\Theta}\sin\theta\right) \\
&= -\mathbf{V}_\parallel\left(\cos\theta + \widehat{\Theta}\sin\theta\right)\left(-\cos\theta + \widehat{\Theta}\sin\theta\right) \\
&\qquad -\mathbf{V}_\perp\left(\cos\theta - \widehat{\Theta}\sin\theta\right)\left(-\cos\theta + \widehat{\Theta}\sin\theta\right) \\
&= \mathbf{V}_\parallel + \mathbf{V}_\perp\cos(2\theta) + \widehat{\Theta}\times\mathbf{V}_\perp\sin(2\theta)
\end{aligned}\; , \tag{2.10}$$

where $\hat{\boldsymbol{\theta}}$ is the unit vector in the direction $\hat{\mathbf{a}}\times\hat{\mathbf{b}}$, and θ is the angle between $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, which is also the dihedral angle of opening between the planes. Because $\hat{\boldsymbol{\theta}}$ is perpendicular to both $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, it lies in both planes, i.e., it lies along the line of intersection of the planes.

Therefore, the product of two reflections is a rotation about an axis that is the line of intersection of the planes, with an angle that is twice the angle of opening between the planes, and every rotation can be expressed in this way. Also, it has been shown that rotations are easily represented algebraically using quaternions, i.e.,

$$\mathbf{V}' = R(\boldsymbol{\theta})\,\mathbf{V}\overline{R(\boldsymbol{\theta})}, \tag{2.11}$$

where

$$R(\theta) = \cos\left(\frac{\theta}{2}\right) + \hat{\boldsymbol{\theta}}\sin\left(\frac{\theta}{2}\right), \tag{2.12}$$

and the bar changes the sign of vectors but not that of scalars. No rotation matrices are necessary!

2.3 Complex Quaternions

Since the quaternions are four-dimensional, it might be thought that they are useful in relativity, and that boosts as well as rotations can be treated using them. This is not possible by staying strictly within the quaternions, because of the indefinite nature of the Lorentz inner product. However, by complexifying the quaternions, or by considering operators on quaternions, formulations of relativity are possible. These formulations of relativity are briefly outlined in this section 3.2. Complex numbers, quaternions, and the complex quaternions are all examples of Clifford algebras, which are considered in the next chapter.

The quaternions form a four-dimensional vector space over the scalar field $\mathbb{R}$. The complex quaternions are obtained when the scalar field $\mathbb{R}$ of real numbers is replaced by the scalar field $\mathbb{C}$ of complex numbers. Thus we would like the complex quaternions, like the real quaternions, to also possess a vector space structure, but now over the complex field. Therefore, if $\alpha \in \mathbb{C}$ and $q \in \mathbb{H}$, then we want αq to be a complex quaternion, and we also want the sum of two complex quaternions to be another complex quaternion. To see how these axioms of vector space structure can be satisfied, let $q_1, q_2 \in \mathbb{H}$, $\alpha_1, \alpha_2 \in \mathbb{C}$ with $\alpha_k = a_k + ib_k$, and consider

$$\begin{aligned} \alpha_1 q_1 + \alpha_2 q_2 &= (a_1 + ib_1)\, q_1 + (a_2 + ib_2)\, q_2 \\ &= (a_1 q_1 + a_2 q_2) + i(b_1 q_1 + b_2 q_2) \, , \\ &= q_1^{'} + i q_2^{''} \end{aligned} \tag{2.13}$$

where $q_1^{'} = a_1 q_1 + a_2 q_2$ and $q_2^{''} = b_1 q_1 + b_2 q_2$ are both elements of $\mathbb{H}$. Hence, the complex quaternions are not just of the form αq with $\alpha \in \mathbb{C}$, but are taken to be $\{q + iq^{'} \mid q, q^{'} \in \mathbb{H}\}$, which forms a four-dimensional vector space over $\mathbb{C}$. As a real vector space, the space of complex quaternions is an eight-dimensional real vector space $\mathbb{H} \oplus \mathbb{H}$.

The quaternions are closed under multiplication, so a natural question is whether the complex quaternions have only a vector space structure, or if they inherit higher-level structure as well. In order to find this out let's see what happens when two complex quaternions are multiplied,

$$\begin{aligned} (q_1 + iq_2)(q_3 + iq_4) &= (q_1 q_3 - q_2 q_4) + i(q_1 q_4 + q_2 q_3) \\ &= q_5 + iq_6, \end{aligned} \tag{2.14}$$

where $q_n \in \mathbb{H}$ for $n = 1, .., 6$. Thus, the product of two elements of the form $q + iq^{'}$ gives another element of the same form, that is, the set of complex quaternions is closed under multiplication. Therefore, the set of complex quaternions forms an algebra. However, unlike the real quaternions, the complex quaternions to do not form a division algebra. Fom Frobenius' theorem, any division algebra over $\mathbb{R}$ is isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. Because the real dimension of the complex quaternions is greater than the dimensions of all of these, they can't form a division algebra. In the complex quaternions, divisors of zero exist. For example, $(1+i\mathbf{i})(1-i\mathbf{i})=0$.

Chapter 3

Clifford Algebras

3.1 General Clifford Algebras

Clifford algebras[3, 4, 5] are generated by inner product spaces and are ideally suited for studying the geometry and symmetries of these inner product spaces. An inner product space is a pair (V, g) where V is a vector space and the inner product is a symmetric bilinear map $g : V \times V \to \mathbb{R}$. The inner product g is said to be nondegenerate if $g(u, v) = 0\ \forall v \in V$ implies $u = 0$. If V is finite dimensional and g is nondegenerate, then there exists an orthonormal basis $\{e_1, \ldots, e_n\}$ for V such that

$$g(e_i, e_j) = \begin{cases} \pm 1; & i = j \\ 0; & i \neq j \end{cases} . \tag{3.1}$$

The inner product space is said to have signature (p, q) if p elements of an orthonormal basis give $+1$ in (3.1) and q elements of an orthonormal basis give -1 in (3.1). For $1 < n = p + q$ many different orthonormal bases exist. The signature is independent of which orthonormal basis is used.

A real associative algebra A is a vector space over $\mathbb{R}$ together with an algebraic product that is a bilinear map

$$\begin{aligned} A \times A &\rightarrow A \\ (a,b) &\longmapsto ab, \end{aligned} \tag{3.2}$$

which for $a, b, c \in A$ and $\alpha \in \mathbb{R}$ satisfies the following two conditions:

1) bilinearity implies multiplication is distributive over addition,

$$\begin{aligned} (a + \alpha b)\, c &= ac + \alpha bc \\ a\,(b + \alpha c) &= ab + \alpha ac, \end{aligned} \tag{3.3}$$

2) the product is associative,

$$a(bc) = (ab)c. \tag{3.4}$$

Note that multiplication is not necessarily commutative. In a complex associative algebra, $\mathbb{R}$ is replaced by $\mathbb{C}$. If there exists $1 \in A$ such that $1a = a1 = a$ for every $a \in A$, then 1 is called a unit or the identity for A.

A real Clifford algebra, denoted $Cl\,(V, g)$, for an inner product space (V, g) is an associative algebra with unit 1 that contains copies of V and $\mathbb{R} = \mathbb{R}1$ as distinct subspaces such that

1) $v^2 = g(v, v)$, $\forall\, v \in V$

2) V generates $Cl\,(V, g)$ as an algebra over $\mathbb{R}$

3) $Cl\,(V, g)$ is not generated by any proper subspace of V.

These axioms then uniquely define a Clifford algebra if the bilinear form g is nondegenerate. Often $Cl\,(V, g)$ is written $Cl_{p,q}$ when (V, g) is an inner product space that has signature (p, q).

In condition 1), the square of v denotes the product of v with itself in this algebra, and on the right hand side $g(v, v)$ is a real number which lies in the vector subspace of the algebra spanned by the identity. Now consider $g(w, w)$, where $w = u + v$:

$$g(w, w) = w^2$$

$$g(u, u) + 2g(u, v) + g(v, v) = u^2 + uv + vu + v^2$$

$$g(u, v) = \frac{1}{2}(uv + vu). \tag{3.5}$$

Therefore, for elements of the orthonormal basis,

$$e_i e_j = \begin{cases} \pm 1; & i = j \\ -e_j e_i; & i \neq j \end{cases}. \tag{3.6}$$

Condition 2) means that every element of $Cl(V, g)$ can be written as a linear combination of products of elements of V. The third condition, called the universal property of the Clifford algebras, is needed to guarantee that the algebra is the largest possible one that satisfies 1) and 2). Without condition 3), it is sometimes possible to generate a lower dimensional nonuniversal Clifford algebra. This is only possible for odd dimensional inner product spaces that satisfy[3] $p - q = 1 \bmod 4$.

Every element of V can be written as a linear combination of basis elements, and consequently every element of $Cl(V, g)$ can be written as a linear combination of products of the orthonormal basis elements $W = \{e_1, \ldots, e_n\}$ for V. Therefore, a maximal linearly independent set of products of the basis elements for V is a basis for $Cl(V, g)$. If a product of elements of W contains an element of W more than once, then (3.6) can be used to reduce it to a product that contains this element at most once. Therefore, a basis for $Cl(V, g)$ can be chosen that consists of products that contain elements of V at most once. Hence, if V is n-dimensional, the number of elements in each product must be less than or equal to n. Equation (3.6) can also be used to reorder the elements in the products such that the indices increase from left to right. For example,

$$e_3e_2e_3e_1 = -e_2e_3e_3e_1 = -e_2e_1 = e_1e_2, \tag{3.7}$$

assuming $e_3^2 = 1$. There are $\binom{n}{m}$ products of m elements that satisfy these conditions. Therefore, a basis for $Cl\,(V, g)$ has $\sum\limits_{m=0}^{n} \binom{n}{m} = 2^n$ elements, with $m = 0$ corresponding to the scalars.

3.2 Quaternions and Complex Numbers

As an example of a Clifford algebra, consider $Cl_{0,2}$, the Clifford algebra generated by a two-dimensional inner product space (V, g) that has a basis $\{c_1, e_2\}$ with

$$g\,(e_1, e_1) = g\,(e_2, e_2) = -1. \tag{3.8}$$

A basis for the real Clifford algebra $Cl_{0,2}$ is thus $\{1, e_1, e_2, e_1e_2\}$. Defining

$$\mathbf{i} := e_1,\ \mathbf{j} := e_2,\ \mathbf{k} := e_1e_2, \tag{3.9}$$

we find the following properties:

1) $\mathbf{ij} = \mathbf{k} = -\mathbf{ji}$, $\mathbf{jk} = \mathbf{i} = -\mathbf{kj}$, and $\mathbf{ki} = \mathbf{j} = -\mathbf{ki}$
2) $\mathbf{i}^2 = \mathbf{j}^2 = -1$, $\mathbf{k}^2 = \mathbf{ijij} = -\mathbf{i}^2\mathbf{j}^2 = -1$
3) $\mathbf{ijk} = \mathbf{k}^2 = -1$.

Note that any of the above two relation implies the third, and also notice that 2) implies that the algebraic product is not commutative.

Properties 2) and 3) are just Hamilton's defining relations for the quaternions. Therefore, the $2^2 = 4$-dimensional Clifford algebra $Cl_{0,2}$ is just the real algebra of the quaternions viewed from a slightly different perspective. Hence any quaternion q can be viewed as an element of $Cl_{0,2}$ and written as a real linear combination of its basis elements, i.e., $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. We define the Clifford conjugate of q by

$$\bar{q} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k} \tag{3.10}$$

and the real valued norm of q by

$$N(q) = q\bar{q} = \bar{q}q = a^2 + b^2 + c^2 + d^2. \tag{3.11}$$

Note that $N(q) = 0$ if and only if $q = 0$. Thus every nonzero element of $Cl_{0,2}$ has an inverse given by $q^{-1} = \frac{\bar{q}}{N(q)}$. Hence, as above, the algebra of real quaternions is a division algebra. When working with the quaternions it is useful to identify $\mathbb{R}$ and $\mathbb{R}^3$ with the subspaces spanned by $\{1\}$ and $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ respectively, so that each quaternion is uniquely expressed as

$$q = a + \mathbf{v} \tag{3.12}$$

where,

$$a = < q >_S := \frac{1}{2}(q + \bar{q}) \tag{3.13}$$

$$\mathbf{v} = < q >_V := \frac{1}{2}(q - \bar{q}). \tag{3.14}$$

The sum and product of two quaternions q and q' is then given by

$$q + q' = (a + a') + (\mathbf{v} + \mathbf{v}') \tag{3.15}$$

$$qq' = aa' - \mathbf{v} \cdot \mathbf{v}' + a'\mathbf{v} + a\mathbf{v}' + \mathbf{v} \times \mathbf{v}'. \tag{3.16}$$

The quaternions contain various subalgebras that can be regarded as the algebra of complex numbers. For example, consider the subspace $\mathbf{C} = \text{span}\{1, \mathrm{i}\}$. Because i commutes with all elements of $\mathbf{C}$ and $\mathrm{i}^2 = -1$, $\mathbf{C}$ is a subalgebra of $\mathbb{H}$ that is isomorphic to $\mathbb{C}$ as an algebra, with the isomorphism given by

$$a + b\mathrm{i} \rightarrow a + ib. \tag{3.17}$$

For any q in $\mathbb{H}$,

$$\begin{aligned} q &= q^0 + q^1\mathrm{i} + q^2\mathrm{j} + q^3\mathbf{k} \\ &= q^0 + q^1\mathrm{i} + \mathrm{j}\left(q^2 - q^3\mathrm{i}\right) \end{aligned}. \tag{3.18}$$

Thus, any quaternion can be expressed as the sum of an element of $\mathbf{C}$ with j times an element of $\mathbf{C}$. Therefore, the identification (3.17) induces the vector space isomorphisms[3]

$$\begin{gathered} \mathbb{H} \rightarrow \mathbf{C} \oplus \mathbf{C} \rightarrow \mathbb{C}^2 \\ q \rightarrow \left(q^0 + q^1\mathrm{i}, q^2 - q^3\mathrm{i}\right) \rightarrow \left(q^0 + q^1 i, q^2 - q^3 i\right) \end{gathered}. \tag{3.19}$$

These isomorphisms are used in 4.3 to make precise and clarify the work of De Leo and Rotelli.

3.3 Pauli Algebra

The Pauli algebra[7] $\mathcal{P}$ is the Clifford algebra Cl_3 of three-dimensional space with g being the ordinary dot product of vectors. The algebraic structure of this Clifford algebra enables us to construct inverses, square roots and functions of vectors, just as one is able to with fields. The main differences between fields such as $\mathbb{R}$ and $\mathbb{C}$, and the Pauli algebra of vectors are that the product in $\mathcal{P}$ is noncommutative and that divisors of zero exist, i.e., there exist Pauli elements $u \neq 0$, $v \neq 0$, with $uv = 0$. This allows us to construct projectors which give the Pauli algebra a richer structure then fields possess.

The Pauli algebra provides very natural ways for expressing geometrical relationships, e.g. rotations and reflections, as well as for four dimensional geometrical relationships in relativity[8, 9, 10]. Because of this close relationship between algebra and geometry, Clifford algebras are sometimes called geometric algebras. Calculations in geometric algebra often avoid the explicit basis dependence of coordinate systems. This is true for any geometric algebra, but what is interesting about the Pauli algebra is that it is the smallest possible Clifford algebra that can incorporate all possible relativistic phenomena in spacetime, from relativistic quantum theory[11, 12], to electrodynamics[7, 13, 14], to general relativity[10].

The Pauli algebra is generated by three-dimensional Euclidean space $\mathbb{R}^3$. Using the axioms in section 3.1 we construct the structure of the Pauli algebra. The product of any two vectors can be expressed as the sum of symmetric and antisymmetric parts:

$$\begin{aligned} \mathbf{ab} &= \tfrac{1}{2}(\mathbf{ab} + \mathbf{ba}) + \tfrac{1}{2}(\mathbf{ab} - \mathbf{ba}) \\ &=: \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b}, \end{aligned} \tag{3.20}$$

where from the first axiom for Clifford algebras, the symmetric part is just $\mathbf{a} \cdot \mathbf{b}$. Note that if $\mathbf{a}$ and $\mathbf{b}$ are parallel then

$$\mathbf{ab} = \mathbf{ba} = \mathbf{a} \cdot \mathbf{b}, \tag{3.21}$$

and if $\mathbf{a}$ and $\mathbf{b}$ are orthogonal then

$$\mathbf{ab} = -\mathbf{ba} = \mathbf{a} \wedge \mathbf{b}. \tag{3.22}$$

It follows from the section on general Clifford algebras that a basis for

Cl_3 is

$$\{1, e_1, e_2, e_3, e_1e_2, e_1e_3, e_2e_3, e_1e_2e_3\}, \tag{3.23}$$

where 1 is the unit scalar, $\{e_j\}$ is an orthonormal vector basis for $\mathbb{R}^3$, $\{e_1e_2, e_1e_3, e_2e_3\}$, is the basis for the bivector space, and $\{e_1e_2e_3\}$ is the trivector representing the volume. Geometrically, the elements of the basis for bivector space represent planes, e.g., e_1e_2 represents the plane spanned by e_1 and e_2. Hence any element of the Pauli algebra can be written in geometrical terms as the linear combination of scalars, vectors (lines), bivectors (planes) and a trivector (volume).

Note that $e_1e_2e_3$ squares to -1 and commutes with all other basis of Cl_3. Hence $\{1, e_1e_2e_3\}$ lies in the centre of the algebra, and forms a subalgebra isomorphic to the algebra of complex numbers $\mathbb{C}$, where $e_1e_2e_3$ plays the role of the imaginary unit i. This gives the Pauli algebra a natural complex structure. Bivectors, with this identification, are written as imaginary vectors, i.e., $e_je_k = i\varepsilon_{jkl}e_l$, and $e_j \rightarrow ie_j$ is an isomorphism between vectors and bivectors. Thus, over the field of complex numbers, Cl_3 is a four-dimensional algebra with basis $\{1, e_1, e_2, e_3\}$. Hence, any element $p \in Cl_3$ is written as the sum of a scalar p^0 and a vector $\mathbf{p}$, i.e., $p = p^0 + \mathbf{p}$, where p^0 and $\mathbf{p}$ may be complex.

The antisymmetric part of the product $\mathbf{ab}$ can now be written as a usual cross product:

$$\begin{aligned} \mathbf{a} \wedge \mathbf{b} &= a^j b^k e_j e_k \\ &= i a^j b^k \varepsilon_{jkl} e_l \\ &= i \mathbf{a} \times \mathbf{b} \end{aligned} \tag{3.24}$$

Therefore, i represents handedness in the algebra, since the association of i

with the volume element $e_1e_2e_3$ can be assumed in right handed coordinate system, i.e., $e_1 \times e_2 = e_3$ and cyclic permutations.

There are three types of involutions for Clifford algebras, with each having natural actions on elements of the Pauli algebra. An involution is an invertible transformation mapping an algebra onto itself which, when composed with itself, gives the identity.

The involution of spatial reversal is an antiautomorphism that reverses the sign of the vector part of the element

$$- : p \rightarrow \bar{p} = p_0 - \mathbf{p}. \tag{3.25}$$

It is easy to verify that the action of bar on the product reverses the order of the product, i.e., $\overline{pq} = \bar{q}\bar{p}$. Any element of Cl_3 can be split into vector and scalar parts using the bar involution:

$$\begin{aligned} p &= \tfrac{1}{2}(p+\bar{p}) + \tfrac{1}{2}(p-\bar{p}) \\ &=: \ < p >_S + < p >_V \end{aligned}, \tag{3.26}$$

and also the product of two elements of Cl_3 can be split into scalar and vector products that generate the dot and cross products of vectors, respectively:

$$< p\bar{q} >_S = \frac{1}{2}(p\bar{q} + \overline{p\bar{q}}) \tag{3.27}$$

$$< p\bar{q} >_V = \frac{1}{2}(p\bar{q} - \overline{p\bar{q}}). \tag{3.28}$$

Note that $< p\bar{q} >_S = < \bar{q}p >_S$.

Hermitian conjugation is another involution which is also an antiautomorphism. Hermitian conjugation is denoted by dagger; it reverses the order of multiplication and changes the sign of i. Hence, the dagger only affects the imaginary part of a Pauli element.

$$\dagger : p \rightarrow p^\dagger = p^{*\mu}e_\mu, \tag{3.29}$$

where $\{e_\mu\}$ with $e_0 \equiv 1$ is the basis (taken over $\mathbb{C}$) for Cl_3 and $p^{*\mu}$ is the complex conjugation of the components p^μ for $\mu = 0, 1, 2, 3$. The Hermitian conjugate involution is used to split the elements of Cl_3 into real and imaginary parts:

$$\begin{aligned} p &= \tfrac{1}{2}(p + p^\dagger) + \tfrac{1}{2}(p - p^\dagger) \\ &=: \ < p >_\Re + < p >_\Im \end{aligned} \tag{3.30}$$

The combination of the involutions spatial reversal and Hermitian conjugation gives another involution, but one which is an automorphism. The map is denoted by bar-dagger and preserves the order of the multiplication. This involution is used to split the Pauli algebra into even and odd parts, i.e., into elements constructed from even and odd products of vectors:

$$\begin{aligned} p &= \tfrac{1}{2}(p + \bar{p}^\dagger) + \tfrac{1}{2}(p - \bar{p}^\dagger) \\ &:= \ < p >_+ + < p >_- \end{aligned} . \tag{3.31}$$

The bar-dagger involution represents the parity transformation. In general Clifford algebras, this transformation is called the grade automorphism or involution. All three involutions are very important and are used to split the Pauli algebra into important parts and subalgebras. For example, $< Cl_3 >_{\Re S} = \left\{ p \in Cl_3 : \tfrac{1}{4}(p + p^\dagger + \bar{p} + \bar{p}^\dagger) \right\} \equiv \mathbb{R}$, the field of real numbers, and $< Cl_3 >_S \equiv \mathbb{C}$, the field of complex numbers.

Now consider $< Cl_3 >_+$ with basis $\{1, -ie_1, -ie_2, -ie_3\}$ Note:

$$(-ie_1)(-ie_2) = -ie_3, \ (-ie_2)(-ie_3) = -ie_1, \ (-ie_3)(-ie_1) = -ie_2, \tag{3.32}$$

and $(-ie_1)^2 = (-ie_2)^2 = (-ie_3)^2 = -1$. Therefore, the identifications

$$1 \to 1, \mathbf{i} \to -ie_1, \mathbf{j} \to -ie_2, \mathbf{k} \to -ie_3 \tag{3.33}$$

give an isomorphism between the even subalgebra $< Cl_3 >_+$ and $\mathbb{H}$, the algebra of real quaternions, that extends to a complex isomorphism of the complex quaternions with the Pauli algebra. Note that this isomorphism does not preserve the representation of vectors. In $\mathbb{H}$, a vector $\mathbf{v}$ is represented as $\mathbf{V} = v^1\mathbf{i} + v^2\mathbf{j} + v^3\mathbf{k}$, while in the Pauli algebra the same vector is represented as $\mathbf{v} = v^1 e_1 + v^2 e_2 + v^3 e_3$. Therefore, the two representations are related by $\mathbf{v} = i\mathbf{V}$, which is just the isomorphism between vectors and bivectors discussed earlier. Thus, Hamiltons "vectors" are realized as bivectors in Cl_3.

A faithful matrix representation of the Pauli algebra is obtained by associating the Pauli spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.34}$$

with the elements of a right handed orthonormal basis $\{e_1, e_2, e_3\}$, and the identity matrix with the unit element of $\mathcal{P}$. Then a vector $\mathbf{v}$ is represented as the 2×2 matrix

$$\mathbf{v} \cdot \boldsymbol{\sigma} := v^1\sigma_1 + v^2\sigma_2 + v^3\sigma_3, \tag{3.35}$$

and (3.20) becomes the familiar

$$\mathbf{a} \cdot \boldsymbol{\sigma}\mathbf{b} \cdot \boldsymbol{\sigma} = \mathbf{a} \cdot \mathbf{b}1 + i\mathbf{a} \times \mathbf{b} \cdot \boldsymbol{\sigma}. \tag{3.36}$$

The dagger involution corresponds to taking the Hermitian conjugate (complex conjugate transpose) in this matrix representation or any unitary transformation of it.

This matrix representation gives the Pauli algebra its name and has proved useful for physics[15, 16], but it should be emphasized that no matrix representation is needed for the algebra. Many representations are possible; what they have in common is the algebra.

3.4 Rotations with the Pauli Algebra

Reflections and rotations have slightly different representations in the Pauli algebra than they do using quaternions. For a reflection of $\mathbf{v}$ through the plane having normal unit $\mathbf{n}$, consider

$$\begin{aligned} -\widehat{\mathbf{n}}\mathbf{v}\widehat{\mathbf{n}} &= -(\mathbf{v}_{\parallel} - \mathbf{v}_{\perp})\,\widehat{\mathbf{n}}\widehat{\mathbf{n}} \\ &= \mathbf{v} - 2\mathbf{v}_{\parallel}. \end{aligned} \tag{3.37}$$

As demonstrated with quaternions, any rotation is the product of two reflections. Therefore, the rotation of $\mathbf{v}$ about the axis $\widehat{\boldsymbol{\theta}}$ by an angle θ is

$$\mathbf{v}' = R(\boldsymbol{\theta})\,\mathbf{v}\overline{R(\boldsymbol{\theta})}, \tag{3.38}$$

where

$$\begin{aligned} R &= \mathbf{n}_2\mathbf{n}_1 \\ &= \mathbf{n}_1 \cdot \mathbf{n}_2 - i\mathbf{n}_1 \times \mathbf{n}_2 \\ &= \cos\left(\tfrac{\theta}{2}\right) - i\widehat{\boldsymbol{\theta}}\sin\left(\tfrac{\theta}{2}\right), \end{aligned} \tag{3.39}$$

and $\mathbf{n}_1$ and $\mathbf{n}_2$ are unit vectors normal to the reflection planes that intersect along $\widehat{\boldsymbol{\theta}}$ and have a dihidral angle of $\frac{\theta}{2}$. Note that $R^{-1} = \overline{R} = R^{\dagger}$. Hence R is unimodular and unitary. The plane perpendicular to the axis of rotation is given by the bivector $\Theta = -i\widehat{\boldsymbol{\theta}}$. For example, if $\widehat{\boldsymbol{\theta}} = e_2$ then $\Theta = -e_1e_2e_3e_2 = e_1e_3$. Only the components of a vector that lie in Θ get rotated by R.

Rotations are used extensively in three-dimensional graphics programming for commercial games and professional flight simulators. These computationally intensive applications require simple, efficient algorithms, which the Pauli algebra can provide. As an example, consider the situation where all the objects in a three-dimensional image are to be rotated from initial orientations to final orientations, and that the images on the screen must move smoothly. What continuous sequence of rotations should be chosen to do the job? The technique used in practice is called spherical linear interpolation[17] (SLERP), and is usually implemented with quaternions. The Pauli algebra equivalent is presented below after disadvantages of other methods are discussed.

An orientation in space is given by a right-handed triad of orthonormal vectors. In many graphics applications orientations are specified by the unique rotation that transforms a fixed reference orientation into a required orientation. Hence, interpolating between orientations amounts to interpolating between rotations. When 3×3 rotation matrices are used, interpolation is difficult. A linear interpolation between rotation matrices results in intermediate matrices that are not rotation (orthogonal) matrices. Using a Gram-Schmidt algorithm to orthogonalize each interpolation matrix is computationally expensive, and may result in jerky motion.

Euler angles can also be used to represent the relationship between a fixed reference orientation and a required orientation. Suppose $\{e_1, e_2, e_3\}$ are the axes of fixed reference orientation. Any rotation can be expressed as the product of a rotation about the e_3 axis followed by a rotation about the e_2 axis followed by a rotation about the e_1 axis[1]. The Euler angles that represent the rotation are the angle ψ of the first rotation, the angle θ of the second rotation, and the angle ϕ of the last rotation. Let the Euler angles of the initial and final orientations be $\{\psi_i, \theta_i, \phi_i\}$ and $\{\psi_f, \theta_f, \phi_f\}$ respectively. Then

$$\left\{(1-t)\,\psi_i + t\psi_f, (1-t)\,\theta_i + t\theta_f, (1-t)\,\phi_i + t\phi_f\right\} \tag{3.40}$$

[1] In most physics texts, the final rotation is again about e_3. However, one can equally use the "gimbal " choice specified here.[18]

for $0 \le t \le 1$ interpolates the Euler angles. However, this interpolation may develop singularities[18] because the correspondence between Euler angles and rotations is not always unique. For example, consider the case when $\theta = \frac{\pi}{2}$ and ψ and ϕ are arbitrary. Then

$$\begin{aligned} R &= \left[\cos\left(\frac{\phi}{2}\right) - ie_1 \sin\left(\frac{\phi}{2}\right)\right]\left[\frac{1}{\sqrt{2}}(1 - ie_2)\right]\left[\cos\left(\frac{\psi}{2}\right) - ie_3 \sin\left(\frac{\psi}{2}\right)\right] \\ &= \frac{1}{\sqrt{2}}\left[\cos\frac{1}{2}(\psi+\phi) - ie_1 \sin\frac{1}{2}(\psi+\phi) - ie_2 \cos\frac{1}{2}(\psi+\phi)\right. \\ &\qquad \left. - ie_3 \sin\frac{1}{2}(\psi+\phi)\right]. \end{aligned} \tag{3.41}$$

Therefore, the sets of Euler angles $\left\{a - \phi, \frac{\pi}{2}, \phi\right\}$ represent the same rotation for any fixed value of a, and interpolation may behave badly when passing through certain values of the Euler angles.

In the usual matrix representation of the Pauli algebra, the condition $R^{-1} = R^{\dagger}$ means that the matrix representative of R is unitary, and the condition $R\overline{R} = 1$ implies that the matrix representative of R is unimodular, i.e.. has unit determinant. Hence, the set of all rotation operators forms a group that is isomorphic to the group $SU(2)$. Suppose $R = a + \vec{c}$ is a rotation operator. Then $\overline{R} = R^{\dagger}$ gives that $a = a*$ and $-\vec{c} = \vec{c}*$. Therefore, any rotation operator can be written as $R = a + i\vec{b}$, where a and $\vec{b}$ are both real, and the group of rotation operators is a subset of the four-dimensional real vector space $\langle Cl_3\rangle_+ = \langle\mathcal{P}\rangle_{S\Re} \oplus \langle\mathcal{P}\rangle_{V\Im}$. Unimodularity of R gives that $1 = a^2 + \vec{b} \cdot \vec{b}$, which means that the group of rotation operators (and SU(2)) has the topology of S^3, the three-dimensional surface of a sphere in four-dimensional Euclidean space. This also suggests taking $\langle A\overline{B}\rangle_S$ as the (positive definite) real inner product of elements A and B of the space $\langle Cl_3\rangle_+$. Hence, rotations are "unit vectors" in this inner product space, and the "angle" ω between rotations R_1 and R_2 is given by

$$\cos\omega = \left\langle R_2\overline{R}_1\right\rangle_S. \tag{3.42}$$

Suppose R_1 and R_2 are rotations that transform the reference orientation into the initial and final orientations respectively. Then $R_2\overline{R}_1$ is a rotation that transforms the initial orientation into the final orientation. Because $R_2\overline{R}_1$ is a rotation, it be expressed as $R_2\overline{R}_1 = e^{-i\omega\hat{\mathbf{n}}}$, which is consistent with (3.42). This means that the "angle" in the inner product space is twice the angle of rotation in physical space. An appropriate interpolation path between R_1 and R_2 is the arc of the "great circle" that runs along the surface of the 3-sphere between them. If t is the curve parameter and $R(t)$ is a rotation operator on this curve such that $R(0) = R_1$ and $R(1) = R_2$, then $R\overline{R}_1$ is a rotation that transforms the initial orientation into an intermediate orientation. Then, $R\overline{R}_1 = e^{-i\omega t\hat{\mathbf{n}}}$ for $0 \leq t \leq 1$ is a linear interpolation of both the angle in physical space and the arc length in the inner product space. Thus

$$R(t) = e^{-i\omega t\hat{\mathbf{n}}}R_1 \tag{3.43}$$

is the required spherical linear interpolation.

The rotation R can also be expressed as a linear combination of R_1 and R_2, i.e.,

$$R(t) = a(t)\,R_1 + b(t)R_2, \tag{3.44}$$

where a and b are scalars. The functional forms of a and b are easily found by expanding the above exponentials. Then $R_2 = (\cos\omega - i\hat{\mathbf{n}}\sin\omega)\,R_1$ implies that $-i\hat{\mathbf{n}}R_1 = (\sin\omega)^{-1}(R_2 - \cos\omega R_1)$ which gives

$$R(t) = \frac{\sin\left[\omega(1-t)\right]R_1 + \sin(\omega t)\,R_2}{\sin\omega} \tag{3.45}$$

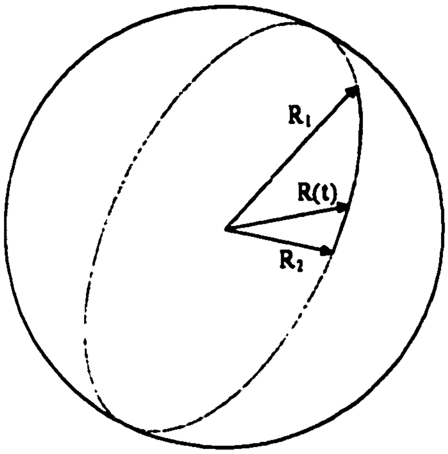

Figure 3.1.

when used in $R_2 = [\cos(\omega t) - i\hat{\mathbf{n}}\sin(\omega t)]\, R_1$. Equation (3.45) is the form of SLERP commonly used in industry, but (3.43) may offer a better implementation.

The bijective correspondence between the group $SU(2)$ and the space S^3 defines a group product on S^3. Therefore, any two elements of S^3 can be multiplied together to give another element of S^3. However, there is no visual geometrical picture of this multiplication.

The surface S^2 of the unit sphere in $\mathbb{E}^3$ can also be used to model $SU(2)$, and in this model the group multiplication is easy to visualize. Consider a rotation $R = \cos\left(\frac{\theta}{2}\right) - i\hat{\boldsymbol{\theta}}\sin\left(\frac{\theta}{2}\right)$ that has the bivector $\Theta = -i\hat{\boldsymbol{\theta}}$ as its plane of rotation. The plane Θ intersects S^2 in a great circle, and R is represented by any directed arc of this great circle that has length $\theta/2$. The direction of the arc is needed to determine the sense of the rotation. Hence, R is represented by an equivalence class of arcs called a spherical vector, where two arcs are equivalent if they lie on the same great circle and have the same direction and length. Any two antipodal points on a great circle are separated by an arc of length $2\pi/2$, which represents a rotation of 2π. An arc that has length $4\pi/2$ joins any point to itself represents a rotation of 4π. These two types of arcs are clearly distinguishable. This illustrates the difference between the $SU(2)$ element -1 that represents a rotation of 2π and the $SU(2)$ element $+1$ that represents a rotation of 4π.

To see the geometry of the group product, let $\overline{R}_1$ and R_2 be two rotations. Hence, $R_{21} = R_2\overline{R}_1$ is also a rotation. Any two great circles intersect in two antipodal points. Suppose c is an intersection of the great circles of the bivectors Θ_1 and Θ_{21}, and that the spherical vector for $\overline{R}_1$ starts at c (see Fig. 3.2). The spherical vector for $\overline{R}_1$ will end at a, an intersection of the great circles for Θ_1 and Θ_2. Since a is on the great circle for Θ_2, the spherical vector for R_2 can start there, and then will end at b, an intersection of the great circles of the bivectors Θ_2 and Θ_{21}. Consistency with the group multiplication law requires that the arc from c to b is a spherical vector for R_{21}. See Baylis[19] for details. Thus, geometrically, the group multiplication corresponds to just addition of spherical vectors. See Figure 3.2. The non-commutivity of the group product is associated with the fact that that addition of spherical vectors is not commutative.

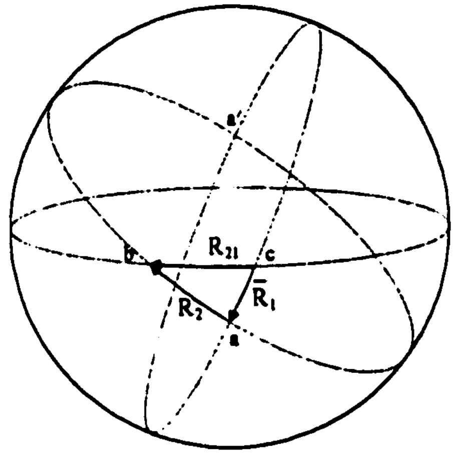

Figure 3.2.

Chapter 4

Special Relativity

4.1 Minkowski Spacetime

Minkowski spacetime[20] consists of the set of all possible events. Once an origin is fixed, Minkowski spacetime becomes the 4-D real vector space $\mathbf{M}$ of 4-positions relative to this origin. The elements of $\mathbf{M}$ are called spacetime vectors. Minkowski spacetime has a non-degenerate symmetric bilinear form g with signature $(1,3)$:

$$g : \mathbf{M} \times \mathbf{M} \longrightarrow \mathbb{R}, \tag{4.1}$$

called the Lorentz inner product. The Lorentz inner product is used to determine (observer dependent) elapsed times and spatial distances between events in spacetime. If $\{e_0, e_1, e_2, e_3\}$ is an orthonormal basis for $\mathbf{M}$, and x and y are the spacetime vectors of any physical events in spacetime, then $x = x^\mu e_\mu$ and $y = y^\mu e_\mu$, where the Einstein summation convention is used. When the speed of light is taken to be one, time and space coordinates are measured in the same units and the Lorentz inner product on $\mathbf{M}$ can be written

$$g(x,y) = x^0y^0 - x^1y^1 - x^2y^2 - x^3y^3. \tag{4.2}$$

A Lorentz transformation L is a linear operator on $\mathbf{M}$ that preserves the Lorentz inner product, i.e.,

1) $L : \mathbf{M} \rightarrow \mathbf{M}$ with $L(x + \alpha y) = Lx + \alpha Ly$

2) $g(x,y) = g(Lx, Ly)$,

for every $x, y \in \mathbf{M}$ and $\alpha \in \mathbb{R}$. In particular, if $\{e_\mu\}$ is an orthogonal basis for $\mathbf{M}$ then $\{e'_\mu = Le_\mu\}$ is also an orthogonal basis for $\mathbf{M}$. The metric coefficients are defined as

$$\eta_{\mu\nu} := g(e_\mu, e_\nu) = g(Le_\mu, Le_\nu), \tag{4.3}$$

and are Lorentz invariant. In component form (4.2) is $g(x,y) = \eta_{\mu\nu}x^\mu y^\nu$.

The following sections embed these features in associative algebras in various ways.

4.2 Relativity with the Complex Quaternions

The rich structure of the algebra of complex quaternions[21] enables us to define spacetime vectors and a Lorentz inner product within this algebra. Indeed, Minkowski spacetime, when complexified turns out to be isomorphic to the vector space formed by the complex quaternions. First, the most important thing is to specify what elements of the algebra are identified with the spacetime vectors. If $\{e_0, e_1, e_2, e_3\}$ is a spacetime vector basis, then the identifications

$$e_1 \rightarrow i\mathbf{i},\ e_2 \rightarrow i\mathbf{j},\ e_2 \rightarrow i\mathbf{k} \tag{4.4}$$

follow from section 2.1. Identifying e_0 with the scalar 1 completes the identifications. Then, a spacetime vector $v = v^\mu e_\mu$ has the form

$$\begin{aligned} v^0 + i\mathbf{v} &= v^0 + i\mathbf{V} \\ &= V^0 + i\left(V^1\mathbf{i}+V^2\mathbf{k}+V^3\mathbf{k}\right). \end{aligned} \tag{4.5}$$

where i is the usual imaginary scalar, not the quaternion unit $\mathbf{i}$. Therefore, in the complex quaternions, a spacetime vector splits into a part that is a real scalar and a part that is an imaginary pure quaternion.

In order to more clearly characterize the subset of the complex quaternions used to represent spacetime vectors, two conjugations are defined. First, the bar conjugation for the quaternions is extended naturally to the complex quaternions by defining

$$\overline{q_1 + iq_2} := \overline{q_1} + i\overline{q_2} \tag{4.6}$$

for q_1, q_2 elements of $\mathbb{H}$. Next, define the dagger conjugation

$$(q_1 + iq_2)^\dagger := q_1 - iq_2 \tag{4.7}$$

for q_1, q_2 elements of $\mathbb{H}$. It is easy to show that the dagger conjugate of a product is the product of the dagger conjugates in the same order, while the bar conjugate of a product is the product of the bar conjugates in the reverse order. Then, in the complex quaternions, any spacetime vector satisfies $V = \overline{V}^\dagger$.

The second most important concept in relativity is the Lorentz inner product. The Lorentz inner product has a very natural definition in the complex quaternions given by

$$\begin{aligned} V\overline{V} &= (v^0 + i\mathbf{V})(v^0 + i\overline{\mathbf{V}}) \\ &= \left(v^0\right)^2 - \mathbf{v}\cdot\mathbf{v} \\ &= \left(v^0\right)^2 - \left(v^1\right)^2 - \left(v^2\right)^2 - \left(v^3\right)^2. \end{aligned} \tag{4.8}$$

Now given that the Lorentz inner product exists in this algebra, we try to find whether Lorentz transformations have natural expressions using the complex quaternions. Consider the transformation

$$V \rightarrow V' = LV\overline{L}^{\dagger}, \tag{4.9}$$

where L is a complex quaternion that satisfies $L\overline{L} = 1$. Then

$$\overline{V'}^{\dagger} = \overline{LV\overline{L}^{\dagger}}^{\dagger} = LV\overline{L}^{\dagger}, \tag{4.10}$$

so V' represents a spacetime vector. Also,

$$V'\overline{V}' = LV\overline{L}^{\dagger}L^{\dagger}\overline{V}\,\overline{L} = LV\left(L\overline{L}\right)^{\dagger}\overline{V}\,\overline{L} = V\overline{V}. \tag{4.11}$$

Therefore (4.9) is a Lorentz transformation. This type of transformation is examined in more detail in 4.4.

4.3 Relativity with Operators on the Real Quaternions

The identifications (3.19) can be used to realize linear operators on $\mathbb{H}$ as 2×2 complex matrices. Examples that are relevant to the work of De Leo and Rotelli[6] are the linear operators Q_q[1] given by left multiplication: $Q_q(q') = qq'$ for all quaternions q and q'. Expressing these quaternions as $q = a + \mathbf{j}b$ and $q' = c + \mathbf{j}d$, where a, b, c, and d are all elements of $\mathbf{C} \equiv span\{1, \mathbf{i}\} \subset \mathbb{H}$ gives

$$Q_q(q') = ac - \overline{b}d + \mathbf{j}(bc + \overline{a}d). \tag{4.12}$$

[1]Actually, $Q : \mathbb{H} \rightarrow \mathrm{End}(\mathbb{H})$ is a representation of $\mathbb{H}$ in the algebra of operators on $\mathbb{H}$, with $Q_q : \mathbb{H} \rightarrow \mathbb{H}$ being the operator (element of $\mathrm{End}(\mathbb{H})$) that results when Q is evaluated at $q \in \mathbb{H}$.

Note that $\mathbf{ij} = -\mathbf{ji}$ implies that $a\mathbf{j} = \mathbf{j}\bar{a}$ for any a in $\mathbf{C}$. Therefore, the 2×2 complex matrix that represents the operator Q_q in the space with $q = a + \mathbf{j}b \to \begin{pmatrix} a \\ b \end{pmatrix}$, $a, b \in \mathbf{C}$ is given by[3]

$$Q_q \to \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}, \tag{4.13}$$

where no distinction has been made between elements of $\mathbf{C}$ and their images in $\mathbb{C}$. In particular,

$$Q_{\mathbf{i}} \to \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, Q_{\mathbf{j}} \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, Q_{\mathbf{k}} \to \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}, Q_{1} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{4.14}$$

which corresponds to the somewhat confusing equation (10) in De Leo and Rotelli[6].

The equally confusing (11) of De Leo and Rotelli is made clear by defining operators I_q by $I_q(q') := Q_q(q')\,\mathbf{i} = qq'\mathbf{i}$. Then

$$I_q \to \begin{pmatrix} ai & -\bar{b}i \\ bi & \bar{a}i \end{pmatrix} \tag{4.15}$$

gives

$$I_{\mathbf{i}} \to \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, I_{\mathbf{j}} \to \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, I_{\mathbf{k}} \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, I_{1} \to \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}. \tag{4.16}$$

Let $\mathcal{O}_1 = \{Q_q \mid q \in \mathbb{H}\}$. This set of operators is a subspace of the space of all operators on the quaternions because $0 = Q_0$ and $Q_{q_1} + aQ_{q_2} = Q_{q_1+aq_2}$. Now assume $Q_{q_1} = Q_{q_2}$. Then $q_1q = q_2q$ for every quaternion q. Taking $q = 1$ gives $q_1 = q_2$. Thus, $q \to Q_q$ is a vector space isomorphism

Then $q_1 q = q_2 q\mathbf{i}$ for every quaternion q. Taking: $q = \mathbf{i}$ gives $q_1\mathbf{i} = -q_2$; and $q = \mathbf{k}$ gives $q_1\mathbf{k} = q_2\mathbf{j}$, and right multiplying by $-\mathbf{j}$ implies $q_1\mathbf{i} = q_2$. Therefore, $q_1 = q_2 = 0$ and $\mathcal{O}_1 \cap \mathcal{O}_2 = \{0\}$. Hence, $\mathcal{O}_1 \oplus \mathcal{O}_2$ is an eight-dimensional subspace of the sixteen-dimensional space of all linear operators on the quaternions.

A bijective correspondence between the complex quaternions and the set of all operators Q_q and $I_{q'}$ is thus given by

$$q + iq' \rightarrow Q_q + I_{q'}. \tag{4.17}$$

Note: $Q_{q_1}Q_{q_2} = Q_{q_1q_2}$, $Q_{q_1}I_{q_2} = I_{q_1}Q_{q_2} = I_{q_1q_2}$, and $I_{q_1}I_{q_2} = -Q_{q_1q_2}$. Therefore, this $\mathcal{O}_1 \oplus \mathcal{O}_2$ is an algebra and the correspondence (4.17) is an isomorphism of algebras provided the correspondence is a homomorphism. Hence, consider $(q_1 + iq_2)(q_3 + iq_4)$ and $(Q_{q_1} + I_{q_2})(Q_{q_3} + I_{q_4})$. First,

$$(q_1 + iq_2)(q_3 + iq_4) = q_1q_3 - q_2q_4 + i(q_2q_3 + q_1q_4). \tag{4.18}$$

Next

$$\begin{aligned}(Q_{q_1} + I_{q_2})(Q_{q_3} + I_{q_4}) &= Q_{q_1q_3} - Q_{q_2q_4} + I_{q_2q_3} + I_{q_1q_4} \\ &= Q_{q_1q_3 - q_2q_4} + I_{q_2q_3 + q_1q_4}\end{aligned}. \tag{4.19}$$

Thus, $(q_1 + iq_2)(q_3 + iq_4) \rightarrow (Q_{q_1} + I_{q_2})(Q_{q_3} + I_{q_4})$ is a homomorphism. Therefore, everything that can be done with the complex quaternions can be done with this algebra of operators. In particular, any spacetime vector v can be represented as the operator $Q_{v^0} + I_{\mathbf{v}}$. Also, the bar and dagger conjugations extend to this operator algebra, which allows for Lorentz transformations.

4.4 Relativity with the Pauli Algebra

When doing relativity in any formalism, it is essential to first define what set represents the elements of the set of all possible events in spacetime, which forms four-dimensional vector space. It is also essential to define on this set the symmetric bilinear mapping g, via which Lorentz transformations are defined. Fortunately, the Pauli algebra possesses both of these structures in a most convenient manner. Thus, the Pauli algebra allows for a natural covariant formulation of special relativity.

Minkowski vector space is taken to be the four-dimensional subspace of Pauli algebra that contains only real elements of Cl_3. Hence, the spacetime vector v is represented by the Pauli algebra element

$$v = v^0 + \mathbf{v} = v^\mu e_\mu, \tag{4.20}$$

where v^μ are the real components of v relative to the basis $\{e_\mu\}$, and e_0 is a unit timelike spacetime vector along the time axis of the observer. Any sum of a scalar and a vector is called a paravector. The paravector v is associated with a contravariant spacetime vector, which can also expressed as $v = v_\mu e^\mu$, where $\{e^\mu := \eta^{\mu\nu} e_\mu\}$ is the reciprocal basis for Minkowski spacetime. The spatial reversal $\bar{v}$ is associated with the covariant spacetime vector $\bar{v} = v^\mu \bar{e}_\mu = v_\mu \bar{e}^\mu$, where $\{\bar{e}_\mu\}$ and $\{\bar{e}^\mu\}$ are considered to be reciprocal bases for the dual of Minkowski spacetime.

The Lorentz inner product g is defined on paravector space by observing that in the Pauli algebra the square of a paravector v is

$$v^2 = \left(v^0 + \mathbf{v}\right)\left(v^0 + \mathbf{v}\right) = \left(v^0\right)^2 + \mathbf{v}\cdot\mathbf{v} + 2v^0\mathbf{v}, \tag{4.21}$$

which, unlike the square of a vector, is not a scalar. However, $v\bar{v}$ is a scalar, and in fact,

$$v\bar{v} = \left(v^0 + \mathbf{v}\right)\left(v^0 - \mathbf{v}\right) = \left(v^0\right)^2 - \mathbf{v}\cdot\mathbf{v}. \tag{4.22}$$

Hence,

$$\bar{v}v = v\bar{v} = g\left(v, v\right) \tag{4.23}$$

is the desired Lorentz norm. As in (3.5), the Lorentz inner product is given by letting $v = x + y$ and using the symmetry and linearity of g in (4.23)

$$g(x, y) = \frac{1}{2}(\bar{x}y + \bar{y}x). \tag{4.24}$$

Notice that the right side of the (4.24) is just the scalar part of the product $< x\bar{y} >_S$. Given the Lorentz inner product, the Minkowski spacetime metric coefficients are determined explicitly in Cl_3 as

$$\eta_{\mu\nu} := \langle e_\mu \bar{e}_\nu \rangle_S = \frac{1}{2}\left(e_\mu \bar{e}_\nu + e_\nu \bar{e}_\mu\right) = \begin{cases} 1; & \mu = \nu = 0 \\ -1; & \mu = \nu = 1, 2, 3 \\ 0; & \mu \neq \nu \end{cases} . \tag{4.25}$$

The scalar part of the product is symmetric in Cl_3, thus $\langle e_\mu \bar{e}_\nu \rangle_S = \langle \bar{e}_\mu e_\nu \rangle_S = \langle \bar{e}_\nu e_\mu \rangle_S = \langle e_\nu \bar{e}_\mu \rangle_S$.

A Lorentz transformation is defined in section 4.1 as a linear transformation that preserves the Lorentz inner product. The set of all Lorentz transformations forms a group called Lorentz group. An important subgroup of a Lorentz group, called the restricted Lorentz group $\mathcal{L}_+^\uparrow$, consists of the proper orthochronous Lorentz transformations. Given an orthonormal basis $\{e_\mu\}$, a Lorentz transformation L is represented by the 4×4 matrix $L_\mu{}^\nu$ defined by $e'_\mu = Le_\mu = L_\mu{}^\nu e_\nu$. It follows from (4.25) that

$$\eta_{\mu\nu} = \left\langle e'_\mu \bar{e}'_\nu \right\rangle_S = L_\mu{}^\alpha L_\nu{}^\beta \left\langle e_\alpha \bar{e}_\beta \right\rangle_S = L_\mu{}^\alpha L_\nu{}^\beta \eta_{\alpha\beta}, \tag{4.26}$$

and $L_\mu{}^\nu$ is an orthogonal matrix with respect to η. A Lorentz transformation L is proper if the matrix $L_\mu{}^\nu$ has determinant $+1$, and is orthochronous if it preserves the time orientation of all timelike spacetime vectors. This matrix representation of the restricted Lorentz group is $SO_+(1,3)$. Restricted Lorentz transformations are particularly physically relevant for special relativity because they relate the orthonormal bases of inertial observers.

The two-fold covering group[2] of $\mathcal{L}_+^\uparrow$ is the spin group $SL(2,\mathbb{C})$ which is isomorphic to the subset[5] $\$pin_+(1,3)$ of the Pauli algebra that consists of all the unimodular elements of Cl_3:

$$\$pin_+(1,3) := \left\{ L \in Cl_3 \mid L\bar{L} = \bar{L}L = 1 \right\}. \tag{4.27}$$

This group is isomorphic to $\mathrm{SL}(2,\mathbb{C})$ in the usual matrix representation of the Pauli algebra. The spacetime vector $x = x^0 + \mathbf{x}$ transforms under $\mathcal{L}_+^\uparrow$ as

$$x \to x' = LxL^\dagger, \tag{4.28}$$

where L is a unimodular element of Cl_3. From (4.28), the corresponding covector $\bar{x}$ transforms like

$$\bar{x} \to \bar{x}' = \bar{L}^\dagger \bar{x} \bar{L}. \tag{4.29}$$

There are always two elements of $\$pin_+(1,3)$, $\pm L$, which map onto the same element in $\mathcal{L}_+^\uparrow$. Thus, there is a 2 to 1 map ϕ from the group of spin transformations to the group of restricted Lorentz transformations, i.e., the $\ker(\phi) = \{\pm 1\}$. Notice that while spin transformations exist as elements of the Pauli algebra, the restricted Lorentz transformations themselves do not exist as single elements of Cl_3.

[2] Locally, $\$pin_+(1,3)$ and $\mathcal{L}_+^\uparrow$ are alike, but globally they differ. The most important global differences are that $\$pin_+(1,3)$ is simply connected, while $\mathcal{L}_+^\uparrow$ is not, and there is 2 to 1 homomorphism from $\$pin_+(1,3)$ onto $\mathcal{L}_+^\uparrow$.

Any spin transformation can be written in the form

$$L = \pm \exp\left\{\frac{1}{2}(\mathbf{w} + i\boldsymbol{\theta})\right\}, \tag{4.30}$$

where $\mathbf{w}$ and $\boldsymbol{\theta}$ are real vectors. If $\mathbf{w} = 0$, then L is unitary and describes a rotation in $\mathbb{R}^3$ in the plane orthogonal to $\boldsymbol{\theta}$ by an angle $|\theta|$. If $\boldsymbol{\theta} = 0$, then L is real, and we have a boost at velocity

$$\mathbf{v} = \tanh(\mathbf{w}), \tag{4.31}$$

where $\mathbf{w}$ represents the rapidity.

However, unless a and b commute, e^{a+b} is not $e^a e^b$, and therefore, $\exp\left\{\frac{1}{2}(\mathbf{w} + i\boldsymbol{\theta})\right\}$ is not generally the product of the $e^{\frac{1}{2}\mathbf{w}}$ with the rotation $e^{\frac{i}{2}\boldsymbol{\theta}}$ unless $\mathbf{w}$ and $\boldsymbol{\theta}$ are parallel. In spite of this it is always possible to factor any transformation into the product of a boost and a rotation as follows. For any spin transformation L, the product $LL^\dagger$ is Hermitian, and gives the proper velocity of an object transformed from rest

$$u = Le_0L^\dagger = LL^\dagger = e^{\mathbf{w}}. \tag{4.32}$$

There exists a unique timelike square root of u with a positive time component which can be used to define a boost $B = \sqrt{LL^\dagger} = e^{\frac{\mathbf{w}}{2}}$. The product $\bar{B}L$ is unitary and unimodular hence it is a rotation $R = e^{-i\frac{\boldsymbol{\theta}}{2}}$. This shows that L can uniquely be written as the product of a boost and a rotation, $L = BR$.

The topology of the group $\{R : R\overline{R} = 1\}$ of rotations was shown in section 3.4 to be that of the 3-sphere S^3. The set $\{\mathbf{w} = \mathbf{w}^\dagger\}$ of boost parameters is a real three-dimensional vector, and thus has the topology of $\mathbb{R}^3$. Because the exponential is continuous with continuous inverse, the set $\{e^{\mathbf{w}} : \mathbf{w} = \mathbf{w}^\dagger\}$ of boosts is topologically identical[23] to the set of boost parameters, and thus also has the same topology as $\mathbb{R}^3$. Therefore, since each ordered pair consisting of a boost and a rotation is associated with a unique spin transformation, $\$pin_+(1,3)$ has the topology of $\mathbb{R}^3 \times S^3$[22]. The space $\mathbb{R}^3 \times S^3$ is easily given a group product by considering it to be the direct product of the groups $\mathbb{R}^3$ and S^3. Therefore, $\$pin_+(1,3)$ and $\mathbb{R}^3 \times S^3$ are both topological spaces and groups, and are equivalent as topological spaces. Are they equivalent also as groups? The answer is no, because the mapping

$$\begin{gathered} \mathbb{R}^3 \times S^3 \rightarrow \$pin_+(1,3) \\ \left(\mathbf{w}, e^{i\boldsymbol{\theta}}\right) \rightarrow e^{\mathbf{w}} e^{i\boldsymbol{\theta}} \end{gathered} \tag{4.33}$$

that makes them isomorphic as topological spaces maps the commutative subgroup $\mathbb{R}^3 \times \{1\}$ bijectively into the set $\{e^{\mathbf{w}}\}$ of boosts, which has a non-commutative product. Thus, this mapping is not an isomorphism of groups, and cannot be used to impose group structure on the set of boosts.

Since the spin transformations L and $-L$ correspond to the same Lorentz transformation, it is interesting to find out how their boost rotation decompositions differ. If $L = e^{\frac{\mathbf{w}}{2}} e^{-i\frac{\boldsymbol{\theta}}{2}}$ then

$$\begin{aligned} -L &= e^{\frac{\mathbf{w}}{2}} e^{-i\frac{\boldsymbol{\theta}}{2}} e^{i\pi\hat{\boldsymbol{\theta}}} \\ &= e^{\frac{\mathbf{w}}{2}} e^{-\frac{i}{2}(\theta - 2\pi)\hat{\boldsymbol{\theta}}} \end{aligned} \tag{4.34}$$

Therefore, the spin transformation that represents the boost is unchanged, and the negative sign is absorbed into the argument of the exponential that represents the rotation.

When $\Lambda^2 \neq 0$, it is always possible to find Λ' such that $e^{\frac{1}{2}\Lambda'} = -e^{\frac{1}{2}\Lambda}$ as follows. When transformed by a spin transformation L, this equation becomes $e^{\frac{1}{2}\mathbf{F}'} = -e^{\frac{1}{2}\mathbf{F}}$, where $\mathbf{F}' = L\Lambda'\overline{L}$ and $\mathbf{F} = L\Lambda\overline{L}$. If an L can be found such that the real and imaginary parts of $\mathbf{F}$ are parallel, then the minus sign can be absorbed into the exponential as in (4.34), so that $e^{\frac{1}{2}\mathbf{F}'} = e^{\frac{1}{2}(\mathbf{F}+2\pi i\hat{\mathbf{F}})}$. Applying the inverse transformation $\overline{L}$ then gives that $e^{\frac{1}{2}\Lambda'} = e^{\frac{1}{2}(\Lambda+2\pi i\overline{L}\hat{\mathbf{F}}L)}$. Then

$$\Lambda' = \Lambda + 2\pi i\overline{L}\widehat{\mathbf{F}}L. \tag{4.35}$$

To find the required L, note that Λ can be Lorentz transformed into $\mathbf{F}$ iff the Lorentz invariants Λ^2 and $\mathbf{F}^2$ are equal. Therefore, there exists a transformation(s) such that $\mathbf{F} = \sqrt{\Lambda^2}e_1$, which has real and imaginary parts parallel. Many different Lorentz transformations will do the job, but a unique boost B can always be found [16]. The direction of the boost is the vector $\hat{u}$ that is orthogonal to both the real and imaginary parts of $\mathbf{F} - \Lambda$ and the proper velocity is

$$u = B^2 = \frac{\mathbf{F}_\perp \Lambda_\perp}{\Lambda_\perp \cdot \Lambda_\perp}, \tag{4.36}$$

where $\perp$ is with respect to $\hat{u}$.

This construction fails when $\Lambda^2 = 0$. Hence, the negative sign in front of the exponential in (4.30) is essential[4, 5] and cannot always be absorbed within the exponential. To see this consider $L = -e^{\frac{1}{2}\Lambda}$, where $\Lambda = \mathbf{w} + i\boldsymbol{\theta}$ and

$$\Lambda^2 = w^2 - \theta^2 - 2i\mathbf{w}\cdot\boldsymbol{\theta} = 0. \tag{4.37}$$

Thus $w = \theta$ and $\mathbf{w}\cdot\boldsymbol{\theta} = 0$. Then $\Lambda^2 = 0$ gives

$$e^\Lambda = 1 + \Lambda + \frac{1}{2!}\Lambda^2 + \ldots = 1 + \mathbf{w} + i\boldsymbol{\theta}. \tag{4.38}$$

Suppose $L = e^{\Lambda'}$, where $\Lambda' = \mathbf{w}' + i\boldsymbol{\theta}'$. If $(\Lambda')^2 = 0$, then

$$1 + \mathbf{w}' + i\boldsymbol{\theta}' = -(1 + \mathbf{w} + i\boldsymbol{\theta}), \tag{4.39}$$

which clearly is impossible. If $(\Lambda')^2 \neq 0$, then

$$\cosh \Lambda' + \frac{\sinh \Lambda'}{\Lambda'} \Lambda' = -(1 + \mathbf{w} + i\boldsymbol{\theta}). \tag{4.40}$$

Then $\cosh \Lambda' = -1$ gives $\Lambda' = \pm i\pi$, which in turn gives $\sinh \Lambda' = 0$, which clearly also is impossible for non-zero Λ. Therefore, this L cannot be written as $e^{\Lambda'}$ for some Λ'. However, L can be approached arbitrarily closely by elements of the form $e^{\Lambda'}$.

Every Λ with $\Lambda^2 = 0$ can be rotated into $\mathbf{A} = a(e_1 - ie_2)$. Hence, for simplicity and without loss of generality, consider $L = -e^{\frac{1}{2}\mathbf{A}}$. Define a one-parameter family

$$\mathbf{W}(\lambda) = ae_1 - i(a - \lambda)e_2, \tag{4.41}$$

where $0 < \lambda < a$. Then $\lim_{\lambda \to 0} \mathbf{W}(\lambda) = \mathbf{A}$ implies that $\lim_{\lambda \to 0} \left[-e^{\frac{1}{2}\mathbf{W}}\right] = -e^{\frac{1}{2}\mathbf{A}}$. But whenever $\lambda \neq 0$,

$$\mathbf{W}^2 = a^2 - (a - \lambda)^2 = \lambda(2a - \lambda) \neq 0. \tag{4.42}$$

Therefore, from the above, there exist $\mathbf{W}'(\lambda)$ with $e^{\frac{1}{2}\mathbf{W}'} = -e^{\frac{1}{2}\mathbf{W}}$, and $\lim_{\lambda \to 0} \left[e^{\frac{1}{2}\mathbf{W}'}\right] = -e^{\frac{1}{2}\mathbf{A}}$. Explicitly, take $\mathbf{F} = \alpha e_1$ where $\alpha = \sqrt{\lambda(2a - \lambda)}$. Then

$$\mathbf{F} - \mathbf{W} = (\alpha - a)e_1 + i(a - \lambda)e_2 \tag{4.43}$$

gives that $\mathbf{u}$ is in the e_3 direction. Therefore $\mathbf{F} = \mathbf{F}_\perp$ and $\mathbf{W} = \mathbf{W}_\perp$, and from (4.36),

$$\begin{aligned} u &= \frac{\mathbf{FW}}{\mathbf{W}^2} \\ &= \alpha^{-1}\left[a - i(a - \lambda)e_1e_2\right] \\ &= \alpha^{-1}\left[a + (a - \lambda)e_3\right]. \end{aligned} \tag{4.44}$$

Finally, (4.35) with $L = B = u^{\frac{1}{2}}$ gives

$$\begin{aligned} \mathbf{W}' &= \mathbf{W} + 2\pi i \overline{B} e_1 B \\ &= \mathbf{W} + 2\pi i e_1 u \\ &= a e_1 + 2\pi \alpha^{-1} (a - \lambda) e_2 + i \left[2\pi \alpha^{-1} a e_1 - (a - \lambda) e_2 \right] . \end{aligned} \tag{4.45}$$

Therefore

$$\Re \left\{ \mathbf{W}'^2 \right\} = a^2 + \left[2\pi \alpha^{-1} (a - \lambda) \right]^2 - \left(2\pi \alpha^{-1} a \right)^2 - (a - \lambda)^2 = \alpha^2 - 4\pi^2 \tag{4.46}$$

$$\Im \left\{ \mathbf{W}'^2 \right\} = 4\pi \alpha^{-1} a^2 - 4\pi \alpha^{-1} (a - \lambda)^2 = 4\pi \alpha, \tag{4.47}$$

and $\mathbf{W}'^2 = (\alpha + 2\pi i)^2$. These together with

$$e^{\frac{1}{2}\mathbf{W}'} = \cosh \left(\frac{1}{2} \sqrt{\mathbf{W}'^2} \right) + \frac{\mathbf{W}'}{\sqrt{\mathbf{W}'^2}} \sinh \left(\frac{1}{2} \sqrt{\mathbf{W}'^2} \right)$$

give

$$\begin{aligned} \lim_{\lambda \to 0} \left[e^{\frac{1}{2}\mathbf{W}'} \right] &= \cosh (\pi i) + (2\pi i)^{-1} \lim_{\lambda \to 0} \left[\mathbf{W}' \sinh \tfrac{1}{2} (\alpha + 2\pi i) \right] \\ &= - \left[1 + \tfrac{a}{2} (e_1 - i e_2) \right] \\ &= - e^{\frac{1}{2}\mathbf{A}} . \end{aligned} \tag{4.48}$$

Now find the boost B and rotation R such that $L = -e^{a(e_1 + i e_2)} = BR$. To find the boost, first calculate

$$\begin{aligned} B^2 &= L L^{\dagger} \\ &= (1 + a e_1 + i a e_2)(1 + a e_1 - i a e_2) \\ &= 1 + 2a^2 + 2a e_1 + 2a^2 e_3 \end{aligned} \tag{4.49}$$

This gives

$$B = \frac{1}{\sqrt{1 + a^2}} \left(1 + a^2 + a e_1 + a^2 e_3 \right), \tag{4.50}$$

and

$$
\begin{aligned}
R &= \overline{B}L \\
&= \tfrac{-1}{\sqrt{1+a^2}}\left(1+a^2-ae_1-a^2e_3\right)\left(1+ae_1+iae_2\right) \\
&= \tfrac{-1}{\sqrt{1+a^2}}\left(1+iae_2\right)
\end{aligned}
\tag{4.51}
$$

In terms of exponentials, $B = e^{\frac{\mathbf{w}}{2}} = \cosh\frac{w}{2} + \widehat{\mathbf{w}}\sinh\frac{w}{2}$ and $R = e^{-i\frac{\theta}{2}e_2} = \cos\frac{\theta}{2} - ie_2\sin\frac{\theta}{2}$ give

$$\widehat{\boldsymbol{\theta}} = e_2, \tag{4.52}$$

$$a = -\tan\frac{\theta}{2} = \sinh\frac{w}{2}, \tag{4.53}$$

$$\widehat{\mathbf{w}} = -e_1\cos\frac{\theta}{2} + e_3\sin\frac{\theta}{2}. \tag{4.54}$$

Thus, $L = -e^{a(e_1+ie_2)} = e^{\sinh^{-1}(a)\widehat{\mathbf{w}}}e^{i\tan^{-1}(a)e_2}$. Even though the boost and rotation are expressed as positive exponentials, their product can only be expressed as the negative of a single exponential. Notice that the boost direction and the rotation axis are perpendicular to each other, and in the limit $a \to \infty$, the boost parameter is infinite and rotation angle is π.

Note that (4.54) gives that $\widehat{\mathbf{w}}$ is a rotation of $-e_1$ about the e_2 axis by an angle $\frac{\theta}{2}$. This gives $\widehat{\mathbf{w}} = e^{-i\frac{\theta}{4}e_2}(-e_1)e^{i\frac{\theta}{4}e_2}$, and thus

$$
\begin{aligned}
L &= e^{e^{-i\frac{\theta}{4}e_2}\left(-\frac{w}{2}e_1\right)e^{i\frac{\theta}{4}e_2}}e^{-i\frac{\theta}{2}e_2} \\
&= e^{-i\frac{\theta}{4}e_2}e^{-\frac{w}{2}e_1}e^{i\frac{\theta}{4}e_2}e^{-i\frac{\theta}{2}e_2} \\
&= e^{-i\frac{\theta}{4}e_2}e^{-\frac{w}{2}e_1}e^{-i\frac{\theta}{4}e_2}
\end{aligned}
\tag{4.55}
$$

Hence, L is a rotation about the e_2 axis by an angle $\frac{\theta}{2}$, followed by a boost in the $-e_1$ direction with rapidity w, followed by a second rotation about the e_2 axis by an angle $\frac{\theta}{2}$. This gives geometrical significance to the vectors that appear in Λ.

Chapter 5

Spinors

Spinors are needed in physics because in relativistic quantum mechanics states are represented by rays in a Hilbert space[23]. The state of a quantum system depends on the reference frame of the observer. Suppose ψ and ψ' are the states with respect to reference frames related by the Lorentz transformation L. Then there is a unitary transformation U_L such that $\psi' = U_L\psi$. However, if ϕ is any normalized element of Hilbert space, then ϕ and $e^{i\theta}\phi$ represent the same physical state of the quantum system. Therefore, if L_1 and L_2 are two Lorentz transformations, then

$$U_{L_1}U_{L_2} = \omega U_{L_1L_2}, \tag{5.1}$$

where ω is a phase factor that depends on L_1 and L_2. Wigner[24] showed that the unitary operators can be chosen such that this phase factor is ± 1. This means a double-valued representation of the Lorentz group, or a single-valued representation of its covering group, $SL(2,\mathbb{C})$, is needed. These representations are the spinor representations. The elements of the representation spaces are called spinors, and are fundamental to physics. Spinors also provide useful insight in classical physics[14].

The smallest irreducible representations of $SL(2, \mathbb{C})$ are on the 2-component spinor spaces[20, 22], and every representation of $SL(2, \mathbb{C})$ is equivalent to a representation of $\$pin_+ (1, 3)$. As a matter of fact, it turns out that minimal left and right ideals of Cl_3 can be used as irreducible representation spaces of $\$pin_+ (1, 3)$[25]. Thus, spinors and spinor algebra can be fully incorporated in the Pauli algebra. This leads to a particularly elegant algebraic formulation of spinors that is largely index free.

Minimal left and right ideals of geometric algebras are created with projectors. Consider the Pauli element $P = \frac{1}{2}(1 + \mathbf{e})$ constructed from the unit spatial direction $\mathbf{e}$. This element is called a projector because P is real and $P = P^2$, which means that projectors are non-invertible elements of the Pauli algebra. A minimal left ideal S of the Pauli algebra is defined by any projector P:

$$S = \{aP \mid a \in Cl_3\}. \tag{5.2}$$

Define an irreducible representation of $\$pin_+ (1, 3)$ on S by the spinor transformation

$$\eta \rightarrow \Lambda\eta \tag{5.3}$$

for every $\eta \in S$ and $\Lambda \in \$pin_+ (1, 3)$. For any elements η and ξ of S, $\eta\bar{\xi} = 0$. A basis over $\mathbb{C}$ for the minimal left ideal S is $\{\alpha_0, \alpha_1\}$, with $\alpha_0 = 2^{\frac{1}{4}} P$ and $\alpha_1 = 2^{\frac{1}{4}} \mathbf{n} P = e^{\frac{\pi}{2}\mathbf{ne}} \alpha_0$, where $\mathbf{n}$ is a unit vector orthogonal to $\mathbf{e}$. The normalization has been chosen to make comparison with standard results easier. In terms of components, $\eta \in S$ is written as $\eta = \eta^A \alpha_A$; $A = 0, 1$. The η^A are the indexed quantities used in the standard formulation of spinors. Thus, the minimal left ideal S is identified with the space of contravariant 2-component spinors.

Notice that $\bar{\alpha}_0\alpha_0 = \bar{\alpha}_1\alpha_1 = 0$ and $\bar{\alpha}_0\alpha_1 = -\bar{\alpha}_1\alpha_0$. This allows for a skew symmetric inner product

$$\langle\ ,\ \rangle : S \times S \to \mathbb{C} \tag{5.4}$$

on S defined by

$$\begin{aligned} \bar{\eta}\xi &= \left(\eta^0\xi^1 - \eta^1\xi^0\right)\bar{\alpha}_0\alpha_1 \\ &= \langle\eta,\xi\rangle\,\bar{\alpha}_0\alpha_1 \qquad , \\ &= \sqrt{2}\,\langle\mathbf{n}\bar{\eta}\xi\rangle_S\,\bar{\alpha}_0\alpha_1 \end{aligned} \tag{5.5}$$

i.e., $\langle\eta,\xi\rangle = \sqrt{2}\,\langle\mathbf{n}\bar{\eta}\xi\rangle_S$. Note that $\langle\eta,\eta\rangle = 0$ for any spinor η. The basis $\{\alpha_0,\alpha_1\}$ satisfies $\langle\alpha_0,\alpha_1\rangle = -\langle\alpha_1,\alpha_0\rangle = 1$. Any basis with this property forms a spin frame. The indexed version of the inner product is given by

$$\epsilon_{AB} := \langle\alpha_A,\alpha_B\rangle\,,\ [\epsilon_{AB}] = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{5.6}$$

and $\langle\eta,\xi\rangle = \eta^A\xi^B\epsilon_{AB}$.

Since $\overline{P} = \overline{P}^2$, $\overline{P}$ is also a projector, it can be used to define a minimal right ideal

$$\begin{aligned} \overline{S} &= \{\overline{P}a \mid a \in Cl_3\} \\ &= \{\bar{\eta} \mid \eta \in S\}. \end{aligned} \tag{5.7}$$

This space is identified with the space of covariant spinors, the dual space to the minimal left ideal S. Under the spin transformation Λ, $\bar{\eta} \in \bar{S}$

$$\bar{\eta} \to \bar{\eta}\bar{\Lambda}. \tag{5.8}$$

Thus, under the spin transformation Λ,

$$\bar{\eta}\xi \to \bar{\eta}\overline{\Lambda}\Lambda\xi = \bar{\eta}\xi, \tag{5.9}$$

and the inner product (5.4) is invariant under all spin transformations. Suppose $\bar{S} = \text{span}\{\overline{\beta}^0, \overline{\beta}^1\}$, with $\{\overline{\beta}^0, \overline{\beta}^1\}$ dual to the basis $\{\alpha_0, \alpha_1\}$, and $\overline{\eta} = \eta_A \overline{\beta}^A$ for $\eta \in S$. The dual basis elements are defined as scalar valued linear functions of spinors in S, namely $\overline{\beta}^A(\alpha_B) = \delta^A_B$, and $\bar{\eta}(\xi) = \eta_A \xi^B \overline{\beta}^A(\alpha_B) = \eta_0 \xi^0 + \eta_1 \xi^1$. On the other hand,

$$\begin{aligned} <\eta, \xi> &= <\eta^A \alpha_A, \xi^B \alpha_B> \\ &= \eta^0 \xi^1 - \eta^1 \xi^0 \end{aligned} . \tag{5.10}$$

Hence, if $\bar{\eta}(\xi) = <\eta, \xi>$, then $\eta_0 = -\eta^1$ and $\eta_1 = \eta^0$ are implied. Thus, $\overline{\eta} = \eta_A \overline{\beta}^A = \eta^B \overline{\alpha_B}$ gives that the basis (over $\mathbb{C}$) for $\bar{S}$ is $\left\{\overline{\beta}^0 = -\overline{\alpha_1}, \overline{\beta}^1 = \overline{\alpha_0}\right\}$.

Similarly, the spaces of contravariant and covariant conjugate (or dotted) spinors can also be identified within Cl_3. The contravariant space is identified with the minimal right ideal

$$\begin{aligned} S^\dagger &= \{Pa \mid a \in Cl_3\} \\ &= \{\eta^\dagger \mid \eta \in S\}. \end{aligned} , \tag{5.11}$$

with the transformation law

$$\eta^\dagger \rightarrow \eta^\dagger \Lambda^\dagger. \tag{5.12}$$

This space is spanned by $\left\{\alpha_{\dot{0}} = \alpha_0^\dagger, \alpha_{\dot{1}} = \alpha_1^\dagger\right\}$, and for every $\eta \in S$, $\eta^\dagger \in S^\dagger$ is written as $\eta^\dagger = \eta^{\dot{X}} \alpha_{\dot{X}}$, $\dot{X} = \dot{0}, \dot{1}$. The covariant space is identified with the minimal left ideal space (covariant conjugate spinors)

$$\begin{aligned} \overline{S}^\dagger &= \{a\overline{P} \mid a \in Cl_3\} \\ &= \{\overline{\eta}^\dagger \mid \eta \in S\}. \end{aligned} \tag{5.13}$$

Under the spin transformation Λ, $\bar{\eta}^\dagger \in \bar{S}^\dagger$

$$\bar{\eta}^\dagger \rightarrow \overline{\Lambda}^\dagger \bar{\eta}^\dagger. \tag{5.14}$$

The basis for $\bar{S}^{\dagger}$ is $\left\{\overline{\beta}^{\dot{0}} = \overline{\alpha}^{0\dagger}, \overline{\beta}^{\dot{1}} = \overline{\alpha}^{1\dagger}\right\}$ and $\forall\ \eta \in S$, $\bar{\eta}^{\dagger}$ is written as $\bar{\eta}^{\dagger} = \eta_{\dot{X}}\overline{\beta}^{\dot{X}}$. The obvious dual relationship between covariant and contravariant conjugate spinors is given by

$$\bar{\eta}^{\dagger}\left(\xi^{\dagger}\right) = <\eta, \xi>^{*}. \tag{5.15}$$

From (5.8), $\bar{S}$ is a representation space for $\$pin_{+}(1,3)$. The representations of $Spin_{+}(1,3)$ on S and $\bar{S}$ are equivalent because $: S \rightarrow \bar{S}$ is an invertible linear mapping. The map $S \rightarrow S^{\dagger}$ is a conjugate isomorphism, i.e.,

$$\eta + \xi \rightarrow \eta^{\dagger} + \xi^{\dagger} \tag{5.16}$$

$$c\eta \rightarrow c^{*}\eta^{\dagger} \tag{5.17}$$

$\forall\ \eta, \xi \in S$ and $c \in \mathbb{C}$. Therefore, the representations of $\$pin_{+}(1,3)$ on S and $S^{\dagger}$ are inequivalent, while representations on $S^{\dagger}$ and $\overline{S}^{\dagger}$ are equivalent.

The product of the spinor η with the conjugate spinor $\xi^{\dagger}$ transforms as

$$\eta\xi^{\dagger} \rightarrow \Lambda\eta\xi^{\dagger}\Lambda^{\dagger} \tag{5.18}$$

under a spin transformation Λ. This is the same transformation law as that of a spacetime vector. Also, $\overline{\eta\xi^{\dagger}}\eta\xi^{\dagger} = 0$. Thus, $\eta\xi^{\dagger}$ is a candidate for a lightlike spacetime vector. However, since in general $\eta\xi^{\dagger}$ is not real, it is an element of complexified Minkowski spacetime. The basis

$$\left\{l = \alpha_0\alpha_0^{\dagger}, m = \alpha_0\alpha_1^{\dagger}, m^{\dagger} = \alpha_1\alpha_0^{\dagger}, n = \alpha_1\alpha_1^{\dagger}\right\} \tag{5.19}$$

for Cl_3 is a basis of lightlike spacetime vectors with

$$\bar{l}l = \overline{m}m = \overline{m}^{\dagger}m^{\dagger} = \bar{n}n = 0 \tag{5.20}$$

$$\langle \bar{l}n\rangle_S = -\left\langle \overline{m}m^\dagger \right\rangle_S = 1. \tag{5.21}$$

Such a basis is often called a null tetrad. Any real spacetime vector v can be expanded with respect to a standard orthonormal basis, or with respect to the null tetrad:

$$v = v^\mu e_\mu = V^{A\dot{B}}\alpha_A\alpha_{\dot{B}}^\dagger. \tag{5.22}$$

Note that the v^μ are real, while the $V^{A\dot{B}}$ may be complex.

In relativity the choice of a metric with signature $(1,3)$ or with signature $(3,1)$ is usually considered a matter of convention. However, in light of (5.22), where a spacetime vector is expressed in terms of both a spacetime vector basis and a spinor basis, the choice of signature $(1,3)$ seems much more natural and convenient[23], as we now show. For the rest of this paragraph assume that the choice $(3,1)$ is made. Then the scalar $\overline{v}v = -v_\mu v^\mu$. Also

$$\begin{aligned}
\overline{v}v &= V^{A\dot{B}}V^{C\dot{D}}\left\langle \overline{\alpha_A\alpha_{\dot{B}}^\dagger}\alpha_C\alpha_{\dot{D}}^\dagger \right\rangle_S \\
&= V^{A\dot{B}}V^{C\dot{D}}\left\langle \overline{\alpha}_{\dot{B}}^\dagger\overline{\alpha}_A\alpha_C\alpha_{\dot{D}}^\dagger \right\rangle_S \\
&= V^{A\dot{B}}V^{C\dot{D}}\epsilon_{AC}\left\langle \overline{\alpha}_{\dot{B}}^\dagger\overline{\alpha}_0\alpha_1\alpha_{\dot{D}}^\dagger \right\rangle_S \\
&= V^{A\dot{B}}V^{C\dot{D}}\epsilon_{AC}\left\langle (\overline{\alpha}_B\alpha_D)^\dagger\,\overline{\alpha}_0\alpha_1 \right\rangle_S \\
&= V^{A\dot{B}}V^{C\dot{D}}\epsilon_{AC}\epsilon_{BD} \\
&= V_{C\dot{D}}V^{C\dot{D}}
\end{aligned} \tag{5.23}$$

Hence, $\overline{v}v = -v_\mu v^\mu = V_{C\dot{D}}V^{C\dot{D}}$. Thus lowering one spacetime index or lowering two spinor indices are both natural ways to change a contravariant spacetime vector into a covariant spacetime vector. When the signature $(3,1)$ is used, these two natural ways of moving to the dual space differ by a minus sign! When doing calculations that involve spinor and spacetime vector indices it is necessary to keep track of which method is used. Choosing the opposite sign for the spinor metric doesn't help because it is used twice for every use of the spacetime metric. When the signature $(1,3)$ is used, no such problems occur because the very natural relation $\bar{v}v = v_\mu v^\mu = V_{C\dot{D}}V^{C\dot{D}}$ is satisfied.

So far, 2-component spinor spaces have been identified with certain left and right ideals in Cl_3. But in order to write the Dirac equation in Cl_3 we must express 4-component Dirac spinors in Cl_3 as well[12]. Dirac spinor space is the direct sum of the two inequivalent 2-component spinor spaces S and $\overline{S}^\dagger$. Now note that the projectors P and $\overline{P}$ satisfy $P\bar{P} = 0$ and $P + \bar{P} = 1$. Hence, the minimal left ideal

$$\tilde{S} = \{a\overline{P} \mid a \in Cl_3\} \tag{5.24}$$

is complementary to the minimal left ideal S. That is, an element a of the Pauli algebra can be expressed uniquely as a sum of an element of S and an element of $\tilde{S}$: $a = aP + a\overline{P}$. In other words, the Pauli algebra is the direct sum $S \oplus \tilde{S}$, which is used to embed Dirac spinors in the Pauli algebra as follows.

First define a spin action on the whole Pauli algebra by $\Psi \rightarrow \Lambda\Psi$ for any Pauli element Ψ. Now,

$$\begin{aligned} \Psi &= \Psi P + \Psi\overline{P} \\ &= \Psi P + \Psi\mathbf{n}\mathbf{n}\overline{P} \\ &= \Psi P + \Psi\mathbf{n}P\mathbf{n} \\ &= \eta + \xi\mathbf{n}, \end{aligned} \tag{5.25}$$

where $\eta = \Psi P$ and $\xi = \Psi\mathbf{n}P$ are both elements of S. The spin action on the Pauli algebra is thus consistent with the spin action (5.3) on S. The space of Dirac spinors is taken to be the Pauli algebra with this spin action and the decomposition into (5.25).

Now that Dirac spinors have been formulated within the Pauli algebra, the Dirac equation is treated in the next chapter.

Chapter 6

Pauli Algebra Dirac Equation

6.1 Dirac Equation

The Dirac equation[11, 12] can be motivated from spacetime momentum. The spacetime momentum of a particle of rest mass m is given by $p = mu$, where u is the proper velocity of the particle with respect to the lab frame. Here, $c = 1$ gives that $u\bar{u} = 1$. The particle's proper velocity with respect to its own rest frame is just 1. These momenta are related by the Lorentz transformation that takes the particle rest frame to the lab frame, i.e., $p = mu = m\Lambda 1 \Lambda^{\dagger}$. This, together with $\Lambda\bar{\Lambda} = 1$ gives

$$p\bar{\Lambda}^{\dagger} = m\Lambda, \tag{6.1}$$

where Λ is the eigenspinor relating the rest frame to the lab frame. Equation (6.1) looks very much like the standard formulation Dirac equation $p^{\mu}\gamma_{\mu}\psi = m\psi$ except that: 1) in a matrix representation of the Pauli algebra, the 4-momentum is expressed using 2×2 matrices, and in the standard formulation it is written using 4×4 matrices, 2) in the standard formulation, ψ is a Dirac spinor with four independent complex degrees of freedom, while in (6.1),

Λ is a unimodular element of the Pauli algebra specified by three complex numbers. The first difference is an advantage over the standard formulation. The second difference means that (6.1) cannot be a Pauli algebra formulation of the Dirac equation. To overcome this difficulty, replace Λ by Ψ, a Pauli algebra version of a Dirac spinor as discussed at the end of the previous chapter. Actually, Ψ is spinor field that depends on spacetime position. Under a spin transformation (that doesn't change the spacetime origin) Λ, $\Psi'(x') = \Lambda\Psi(x)$, which is the same as $\Psi'(x) = \Lambda\Psi\left(\overline{\Lambda}x\overline{\Lambda}^{\dagger}\right)$.

The differential form of the standard formulation Dirac equation is obtained by replacing the conjugate momentum $p^{\mu} + eA^{\mu}$ by $i\partial^{\mu}$. In the Pauli algebra formulation, $(p + eA)\overline{\Psi}^{\dagger}$ is replaced by $i\partial\overline{\Psi}^{\dagger}\mathbf{e}$, where $\mathbf{e}$ is the spin axis in the electron's rest frame, and $\partial := e^{\mu}\partial_{\mu}$ is the Pauli algebra differential operator. This results in the Pauli algebra Dirac equation

$$i\partial\overline{\Psi}^{\dagger}\mathbf{e} - eA\overline{\Psi}^{\dagger} = m\Psi, \tag{6.2}$$

or equivalently, after taking bar-dagger

$$i\overline{\partial}\Psi\mathbf{e} - e\overline{A}\Psi = m\overline{\Psi}^{\dagger}. \tag{6.3}$$

Under a spin transformation Λ, $\overline{\partial} \rightarrow \overline{\Lambda}^{\dagger}\overline{\partial}\,\overline{\Lambda}$, $\Psi \rightarrow \Lambda\Psi$, $\overline{A} \rightarrow \overline{\Lambda}^{\dagger}\overline{A}\,\overline{\Lambda}$, and $\mathbf{e}$ is invariant because it always represents the spin in the rest frame of the electron. Lorentz covariance is easily shown by transforming these quantities in (6.3).

It is very easy to verify that the Pauli algebra Dirac equation is equivalent to the standard Dirac equation. This is done by multiplying (6.2) from the right by the projectors P and $\overline{P}$. Using $\mathbf{e}P = P$ and $\mathbf{e}\overline{P} = -\overline{P}$, this gives

$$i\partial\overline{(\Psi\overline{P})}^{\dagger} - eA\overline{(\Psi\overline{P})}^{\dagger} = m\Psi P \tag{6.4}$$

$$-i\partial\overline{(\Psi P)}^{\dagger} - eA\overline{(\Psi P)}^{\dagger} = m\Psi\overline{P}, \tag{6.5}$$

and with the identifications $\Psi P = \eta$ and $\Psi\bar{P} = \xi\mathbf{n}$ of the previous section these become

$$(i\partial - eA)\,\overline{(\xi\mathbf{n})}^{\dagger} = m\eta \tag{6.6}$$

$$-(i\partial + eA)\,\overline{\eta}^{\dagger} = m\xi\mathbf{n}. \tag{6.7}$$

Thus, these equations are two coupled equation with η and ξ, and from this we can write the Weyl bispinor as

$$\psi = \begin{pmatrix} -\bar{\xi}^{\dagger}\mathbf{n} \\ \eta \end{pmatrix}. \tag{6.8}$$

The standard Dirac equation $\gamma^{\mu}(i\partial_{\mu} - eA_{\mu})\psi = m\psi$, when expanded in the Weyl representation, leads to the same two coupled equations.

If $\Psi_1 = \eta_1 + \xi_1\mathbf{n}$ and $\Psi_2 = \eta_2 + \xi_2\mathbf{n}$ are two solutions of (6.2), then any real combination of Ψ_1 and Ψ_2 is also a solution. However, $c_1\Psi_1$ is not necessarily a solution for c_1 an arbitrary complex scalar, and thus the superposition principle does not hold for the Pauli algebra Dirac equation. This is because the appearance of dagger on only one side of (6.2) makes it a complex non-linear equation. There are Weyl bispinor solutions ψ_1 and ψ_2 to the standard Dirac equation that correspond to Ψ_1 and Ψ_2. The superposition principle for the standard Dirac equation then gives that

$$c_1\psi_1 = \begin{pmatrix} -\overline{(c_1{}^{*}\xi_1)}^{\dagger}\mathbf{n} \\ c_1\eta_1 \end{pmatrix} \tag{6.9}$$

is also a solution. Therefore, for $c_1 = r_1 e^{i\theta_1}$ with real r_1 and θ_1,

$$\begin{aligned} \phi &= c_1\eta_1 + c_1{}^{*}\xi_1\mathbf{n} \\ &= r_1\left(e^{i\theta_1}\eta_1 + e^{-i\theta_1}\xi_1\mathbf{n}\right) \end{aligned} \tag{6.10}$$

is the solution to the Pauli algebra Dirac equation that corresponds to $c_1\psi_1$. Now note that

$$\begin{aligned} Pe^{i\theta\mathbf{e}} &= \tfrac{1}{2}(1+\mathbf{e})(\cos\theta + i\mathbf{e}\sin\theta) \\ &= Pe^{i\theta}, \end{aligned} \tag{6.11}$$

and $\eta = \eta P$. Hence, $\phi = r_1\Psi_1 e^{i\theta_1\mathbf{e}}$ is a solution to the Pauli algebra Dirac equation whenever Ψ_1 is. This can be verified by substituting ϕ directly into (6.2). Note that $e^{i\theta_1\mathbf{e}}$ is rotation about the spin axis $\mathbf{e}$.

In the Weyl representation for the standard Dirac equation, the charge conjugation operator interchanges η and ξ. Therefore, in the Paul algebra the action of charge conjugation C is

$$\Psi = \eta + \xi\mathbf{n} \rightarrow \Psi_C = \xi + \eta\mathbf{n} = \Psi\mathbf{n}. \tag{6.12}$$

To confirm this calculate

$$\begin{aligned} m\overline{\Psi_C}^{\dagger} &= -m\overline{\Psi}^{\dagger}\mathbf{n} \\ &= -\left(-i\bar{\partial}\Psi_C\mathbf{e} - e\bar{A}\Psi_C\right) \\ &= i\bar{\partial}\Psi_C\mathbf{e} + e\bar{A}\Psi_C, \end{aligned} \tag{6.13}$$

which is (6.3) with the sign of the charge changed. Under parity P, all spatial vectors reverse direction, but products of spatial vectors do not change order. This is accomplished by taking the bar-dagger:

$$\Psi = \eta + \xi\mathbf{n} \rightarrow \Psi_P = \overline{\Psi}^{\dagger} = \bar{\eta}^{\dagger} - \bar{\xi}^{\dagger}\mathbf{n}. \tag{6.14}$$

Hence Ψ_P satisfies

$$\begin{aligned} m\overline{\Psi_P}^{\dagger} &= m\Psi \\ &= i\partial\overline{\Psi}^{\dagger}\mathbf{e} - eA\overline{\Psi}^{\dagger} \\ &= i\partial\Psi_P\mathbf{e} - eA\Psi_P. \end{aligned} \tag{6.15}$$

In other words $\overline{\partial} \rightarrow \partial$ and $\overline{A} \rightarrow A$ as expected. Time reversal T is achieved by

$$\Psi = \eta + \xi \mathbf{n} \rightarrow \Psi_T = \overline{\Psi}^\dagger e^{\mathbf{i}\frac{\pi}{2}\mathbf{n}} = \overline{\left(\eta e^{\mathbf{i}\frac{\pi}{2}\mathbf{n}}\right)}^\dagger - \overline{\left(\xi e^{\mathbf{i}\frac{\pi}{2}\mathbf{n}}\right)}^\dagger \mathbf{n}. \tag{6.16}$$

Therefore,

$$\begin{aligned} m\overline{\Psi_T}^\dagger &= m\Psi e^{\mathbf{i}\frac{\pi}{2}\mathbf{n}} \\ &= i\partial\overline{\Psi}^\dagger e^{-\mathbf{i}\frac{\pi}{2}\mathbf{n}}\mathbf{e} - eA\overline{\Psi}^\dagger e^{\mathbf{i}\frac{\pi}{2}\mathbf{n}} \; . \\ &= -i\partial\Psi_T\mathbf{e} - eA\Psi_T. \end{aligned} \tag{6.17}$$

Under time reversal $\mathbf{e} \rightarrow -\mathbf{e}$ and $\overline{p} \rightarrow p$. This is the same as rotating the spin axis about $\mathbf{n}$ by π followed by a parity transformation.

6.2 Neutrinos

Recently, there has been much interest in neutrinos, which are electrically neutral elementary particles that satisfy the Dirac equation. The current interest has largely centred on whether or not neutrinos have mass. Neutrinos interact with matter through the electroweak force. When the electromagnetic and weak forces decouple in the low-energy limit, neutrinos interact with matter only via the weak force. This makes neutrinos difficult, but not impossible, to detect experimentally. Also, this means that the electric charge $e = 0$ in the Dirac equation for neutrinos. Independent neutrinos exist for each of the three generations of leptons, the electron, muon and tau particles. A brief introduction to neutrino oscillations and to the treatment of Dirac-theory neutrinos in the Pauli algebra follows.

Neutrinos were originally introduced theoretically to conserve energy, momentum and angular momentum in beta decay. At the time, there was no direct experiment evidence for this particle, but the only other possibility seemed to be to give up the law of conservation of spacetime momentum. The indirect experimental evidence that existed indicated that the mass of a neutrino had to be very small[26], so the simplest assumption was to set its mass equal to zero. Since it needed to be a fermion of spin $\frac{1}{2}$, it should be governed by the Dirac equation (6.3) with $m = e = 0$

$$\begin{aligned} 0 &= i\overline{\partial}\Psi\mathbf{e} \\ &= i\overline{\partial}(\eta + \xi\mathbf{n}), \end{aligned} \tag{6.18}$$

which decouples into $i\overline{\partial}\eta = 0$ and $i\overline{\partial}\xi = 0$. One of these equations is sufficient, because they are mathematically identical, and η and ξ don't interact with each other. Hence, neutrino wavefunction space is $\mathcal{N} = \{\eta \in S : i\overline{\partial}\eta = 0\}$. If η is a neutrino wavefunction, then parity acting on η gives $\overline{\eta}^\dagger \in \overline{S}^\dagger$, which is not in $\mathcal{N}$, and thus not a neutrino wavefunction. Therefore, since parity is not even defined on $\mathcal{N}$, parity cannot be a symmetry of this formulation of the neutrino. This was initially thought to be a negative aspect of the theory, but was later confirmed by experiment[27].

Neutrinos produced in nuclear reactions in the sun are detectable on the earth. However, the number of solar neutrinos detected is smaller than expected. One possible explanation is that neutrino oscillations[28] are involved. These oscillations are possible only when one or more of the neutrino masses is nonzero. To show this, assume for simplicity that there are only two flavours (generations) of neutrinos, the electron neutrino $|\nu_e\rangle$ and the muon neutrino $|\nu_\mu\rangle$. The mechanism for neutrino oscillation also assumes that there is mass mixing, so that the flavour eigenstates are superpositions of orthogonal mass eigenstates. Hence, in the usual Dirac notation,

$$|\nu_e\rangle = c_1\,|m_1\rangle + c_2\,|m_2\rangle \tag{6.19}$$

$$|\nu_\mu\rangle = d_1\,|m_1\rangle + d_2\,|m_2\rangle\,. \tag{6.20}$$

Orthogonality of the flavour eigenstates gives $c_1d_1^* + c_2d_2^* = 0$. If an electron neutrino in state $|\nu_e\rangle$ is created at the spacetime origin and the mass eigenstates propagate as plane waves that have the same 3-momentum $\mathbf{p}$, then the state of the neutrino at all spacetime positions x is given by

$$|\psi\rangle = c_1 e^{-i\langle p_1\bar{x}\rangle_S}\,|m_1\rangle + c_2 e^{-i\langle p_2\bar{x}\rangle_S}\,|m_2\rangle\,. \tag{6.21}$$

The probability of detecting a muon neutrino at spacetime position x is then given by

$$\begin{aligned} |\langle\nu_\mu|\psi\rangle|^2 &= \left|d_1^*c_1e^{-i\langle p_1\bar{x}\rangle_S} + d_2^*c_2e^{-i\langle p_2\bar{x}\rangle_S}\right|^2 \\ &= |d_1c_1|^2 + |d_2c_2|^2 + 2\,\langle d_1^*c_1d_2c_2^*\exp\{-i\,\langle(p_2-p_1)\,\bar{x}\rangle_S\}\rangle_{\Re}\,. \end{aligned} \tag{6.22}$$

The probability at the spacetime origin is, as expected, $|c_1d_1^* + c_2d_2^*|^2 = 0$, but oscillates between 0 and 1 as spacetime position changes. Thus, an electron neutrino can "oscillate" into a muon neutrino as it propagates. Note that if the masses of the mass eigenstates are equal then equal 3-momenta implies that the spacetime momenta p_1 and p_2 are equal, and no oscillation occurs. In particular, both masses cannot be zero.

Consider the Dirac equation for the neutrino when $m \neq 0$, for which (6.3) is $i\bar{\partial}\Psi\mathbf{e} = m\overline{\Psi}^\dagger$. Therefore, a neutrino wavefunction in this formulation is $\Psi \in \mathcal{P}$ that satisfies the Pauli algebra Dirac equation with $e = 0$. Hence, from the above, parity acting on any neutrino wavefunction gives a wavefunction that is also a neutrino wavefunction. Thus, in this formulation, the theory of the massive neutrino may be made invariant under (intrinsic) parity.

For a formulation of the neutrino that is explicitly not invariant under parity, consider a massive Majorana neutrino. A massive Majorana neutrino is its own antiparticle, i.e., its wavefunction satisfies $\Psi = \Psi_C$. When $\Psi = \eta + \xi\mathbf{n}$, $\Psi_C = \xi + \eta\mathbf{n}$, and $\Psi = \Psi_C$ iff $\xi = \eta$. Then (6.3) reduces to $i\bar{\partial}\eta = m\overline{(\eta\mathbf{n})}^{\dagger}$, and each Majorana wavefunction has half the number of degrees of freedom of the massive Dirac neutrino. The Majorana formulation of the massive neutrino does not have parity as a symmetry because, as in the case for the massless neutrino, parity cannot be defined on the neutrino wavefunction space.

Both the massive Dirac neutrino and the massive Majorana neutrino approach the massless case smoothly as $m \to 0$. Therefore if the mass of the neutrino is small but nonzero, it could prove difficult to distinguish experimentally between them.

Chapter 7

Conclusions

Geometrical algebras are ideally suited for providing geometrical insight into various subjects and calculations. For example, the geometrical basis for rotations and Lorentz transformations is best illustrated using the Pauli algebra, which is the geometrical algebra for physical space. Geometrical algebras also unify many concepts and formalisms, including the complex numbers, quaternions, and complex quaternions. When using geometrical algebras, no matrix representation is needed. In fact, it is usually best to avoid such a representation.

The Pauli algebra contains the complex numbers and quaternions as subalgebras, and is itself isomorphic to the complex quaternions. Thus, it is no surprise that the geometrical concepts for space, and applications of complex numbers and quaternions, are all easiiy expressed using it. The geometric algebra for spacetime has twice the dimension of the Pauli algebra, and would seem to be a natural tool for relativity. Therefore, it is a pleasant surprise that all aspects of relativity can be formulated easily within the smaller and simpler Pauli algebra. This includes 2 and 4-component spinors, which appear as algebraic ideals of the Pauli algebra.

This thesis has investigated a number of examples in which the Pauli algebra unifies quaternionic, spinorial, and other approaches to problems involving rotations, boosts and relativistic quantum theory. These investigations have shown the versatility and power of the Pauli algebra.

Bibliography

[1] M. Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, Oxford 1972.

[2] S. Adler, *Quaternionic Quantum Mechanics and Quantum Fields*, Oxford University Press, Oxford 1995.

[3] I. R. Porteous, *Clifford Algebras and the Classical Groups*, Cambridge University Press, Cambridge 1997.

[4] M. Riesz, *Clifford Numbers and Spinors*, (eds.: E. F. Bolinder, P. Lounesto), Kluwer, Dordrecht 1993.

[5] P. Lounesto, *Clifford Algebras and Spinors*, Cambridge University Press, Cambridge 1997.

[6] S. De Leo and P. Rotelli, "Translations between Quaternion and Complex Quantum Mechanics," *Progress of Theoretical Physics* **92**(5), 917-926 1994.

[7] W. E. Baylis, "Introduction," in *Clifford (Geometric) Algebras*, edited by W. E. Baylis, Birkhäuser, Boston 1996.

[8] W. E. Baylis, *Electrodynamics, a Modern Geometrical Appraoch*, Birkhäuser, Boston 1998.

[9] W. E. Baylis and G. Jones, "The Pauli Algebra Approach to Special Relativity," *J. Phys.* A**22**, 1-15 1989.

[10] G. Jones, "The Pauli Algebra Approach to Relativity," Ph. D. dissertation, University of Windsor 1993.

[11] W. E. Baylis, "Classical Eigenspinors and the Dirac Equation," *Phys. Rev.* A**45**, 4293-4302 1992.

[12] W. E. Baylis, "Eigenspinors in Quantum Theory," in *Clifford (Geometric) Algebras*, edited by W. E. Baylis, Birkhäuser, Boston 1996.

[13] W. E. Baylis and G. Jones, "Relativistic Dynamics of Charges in External Fields: The Pauli Algebra Approach," *J. Phys.* **A22**, 17-29 1989.

[14] W. E. Baylis and Y. Yao, "Relativistic Dynamics of Charges in Electromagnetic Fields: an Eigenspinor Approach," *Phys. Rev.* **A60**, 785-795 1999.

[15] W. E. Baylis, "Special Relativity with 2×2 Matrices," *Am. J. Phys.* **48**, 918-925 1980.

[16] W. E. Baylis and G. Jones , "Special Relativity with Clifford Algebras and 2×2 Matrices, and the Exact Product of two Boosts", *J. Math Phys.*, **29** 57-62 1988.

[17] N. Bobick, "Rotating Objects using Quaternions" *Game Developer*, **2**(26), 1998.

[18] Waterloo CS 488 course web site:
http://www.undergrad.math.uwaterloo.ca/~cs488/ln.html.

[19] W. E. Baylis, "Transformation Parameters," unpublished 1999.

[20] G. Naber, *The Geometry of Minkowski Spacetime*, Springer-Verlag, New York 1992.

[21] J. Lambek, "If Hamilton had Prevailed: Quaternions in Physics," *Mathematical Intelligencer*, **17**(6), 7-15 1995.

[22] R. Penrose and W. Rindler, *Spinors and Spacetime Volume 1: Two Spinor Calculus and Relativistic Fields*, Cambridge University Press, Cambridge 1984.

[23] R. M. Wald, *General Relativity*, University of Chicago Press, Chicago 1984.

[24] E. P. Wigner, "On Unitary Representations of the Inhonogeneous Lorentz Group," *Ann. Math.*, **40**, 149-204 1939.

[25] G. Jones and W. E. Baylis, "Crumeyrolle-Chevalley-Riesz Spinors and Covariance," in *Clifford Algebras and Spinors*, edited by P. lounesto and R. Ablamowicz, Kluwer Academic 1994.

[26] D. Grifftiths, *Introduction to Elementary Particles*, John Wiley & Sons, New York 1987.

[27] C. S. Wu et al., *Phys. Rev.*, **105**, 1413 1957.

[28] B. Kayser with F. Gibrat-Debu and F. Perrier, *The Physics of Massive Neutrinos*, World Scientific, Singapore 1989.

Vita Auctoris

NAME:	Shazia Hadi
PLACE OF BIRTH:	Karachi, Pakistan
YEAR OF BIRTH:	1971
EDUCATION:	OAC, Nelson A. Boylen C. I., Toronto, 1990
	B. Sc., Physics, York University, Toronto, 1995