

## 3.4: The Pauli Algebra

### 3.4.1 Introduction

Let us consider the set of all  $2 \times 2$  matrices with complex elements. The usual definitions of matrix addition and scalar multiplication by complex numbers establish this set as a four-dimensional vector space over the field of complex numbers  $\mathcal{V}(4, C)$ . With ordinary matrix multiplication, the vector space becomes, what is called an algebra, in the technical sense explained at the end of Section 2.3. The nature of matrix multiplication ensures that this algebra, to be denoted  $\mathcal{A}_2$  is associative and noncommutative, properties which are in line with the group-theoretical applications we have in mind.

The name “Pauli algebra” stems, of course, from the fact that  $\mathcal{A}_2$  was first introduced into physics by Pauli, to fit the electron spin into the formalism of quantum mechanics. Since that time the application of this technique has spread into most branches of physics.

From the point of view of mathematics,  $\mathcal{A}_2$  is merely a special case of the algebra  $\mathcal{A}_n$  of  $n \times n$  matrices, whereby the latter are interpreted as transformations over a vector space  $\mathcal{V}(n^2, C)$ . Their reduction to canonical forms is a beautiful part of modern linear algebra.

Whereas the mathematicians do not give special attention to the case  $n = 2$  the physicists, dealing with four-dimensional space-time, have every reason to do so, and it turns out to be most rewarding to develop procedures and proofs for the special case rather than refer to the general mathematical theorems. The technique for such a program has been developed some years ago.

The resulting formalism is closely related to the algebra of complex quaternions, and has been called accordingly a system of hypercomplex numbers. The study of the latter goes back to Hamilton, but the idea has been considerably developed in recent years. The suggestion that the matrices (1) are to be considered symbolically as generalizations of complex numbers which still retain “number-like” properties, is appealing, and we shall make occasional use of it. Yet it seems confining to make this into the central guiding principle. The use of matrices harmonizes better with the usual practice of physics and mathematics.

In the forthcoming systematic development of this program we shall evidently cover much ground that is well known, although some of the proofs and concepts of Whitney and Tisza do not seem to be used elsewhere. However, the main distinctive feature of the present approach is that we do not apply the formalism to physical theories assumed to be given, but develop the geometrical, kinematic and dynamic applications in close parallel with the building up of the formalism.

Since our discussion is meant to be self-contained and economical, we use references only sparingly. However, at a later stage we shall state whatever is necessary to ease the reading of the literature.

### 3.4.2 Basic Definitions and Procedures

We consider the set  $\mathcal{A}_2$  of all  $2 \times 2$  complex matrices

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad (3.4.1)$$

Although one can generate  $\mathcal{A}_2$  from the basis

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (3.4.2)$$

$$e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (3.4.3)$$

$$e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (3.4.4)$$

$$e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.4.5)$$

in which case the matrix elements are the expansion coefficients, it is often more convenient to generate it from a basis formed by the Pauli matrices augmented by the unit matrix.

Accordingly  $\mathcal{A}_2$  is called the Pauli algebra. The basis matrices are

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.4.6)$$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (3.4.7)$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (3.4.8)$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.4.9)$$

The three Pauli matrices satisfy the well known multiplication rules

$$\sigma_j^2 = 1 \quad j = 1, 2, 3 \quad (3.4.10)$$

$$\sigma_j \sigma_k = -\sigma_k \sigma_j = i \sigma_l \quad jkl = 123 \text{ or an even permutation thereof} \quad (3.4.11)$$

All of the basis matrices are Hermitian, or self-adjoint:

$$\sigma_\mu^\dagger = \sigma_\mu \quad \mu = 0, 1, 2, 3 \quad (3.4.12)$$

(By convention, Roman and Greek indices will run from one to three and from zero to three, respectively.)

We shall represent the matrix  $A$  of Equation 3.4.1 as a linear combination of the basis matrices with the coefficient of  $\sigma_\mu$  denoted by  $a_\mu$ . We shall refer to the numbers  $a_\mu$  as the components of the matrix  $A$ . As can be inferred from the multiplication rules, Equation 3.4.11, matrix components are obtained from matrix elements by means of the relation

$$a_\mu = \frac{1}{2} \text{Tr}(A \sigma_\mu) \quad (3.4.13)$$

where  $\text{Tr}$  stands for trace. In detail,

$$a_0 = \frac{1}{2} (a_{11} + a_{22}) \quad (3.4.14)$$

$$a_1 = \frac{1}{2} (a_{12} + a_{21}) \quad (3.4.15)$$

$$a_2 = \frac{1}{2} (a_{12} - a_{21}) \quad (3.4.16)$$

$$a_3 = \frac{1}{2} (a_{11} - a_{22}) \quad (3.4.17)$$

In practical applications we shall often see that a matrix is best represented in one context by its components, but in another by its elements. It is convenient to have full flexibility to choose at will between the two. A set of four components  $a_\mu$ , denoted by  $\{a_\mu\}$ , will often be broken into a complex scalar  $a_0$  and a complex “vector”  $\{a_1, a_2, a_3\} = \vec{a}$ . Similarly, the basis matrices of  $\mathcal{A}_2$  will be denoted by  $\sigma_0 = 1$  and  $\{\sigma_1, \sigma_2, \sigma_3\} = \vec{\sigma}$ . With this notation,

$$A = \sum_\mu a_\mu \sigma_\mu = a_0 1 + \vec{a} \cdot \vec{\sigma} \quad (3.4.18)$$

$$= \begin{pmatrix} a_0 + a_3 & a_1 - i a_2 \\ a_1 + i a_2 & a_0 - a_3 \end{pmatrix} \quad (3.4.19)$$

We associate with each matrix the half trace and the determinant

$$\frac{1}{2} \text{Tr} A = a_0 \quad (3.4.20)$$

$$|A| = a_0^2 - \vec{a}^2 \quad (3.4.21)$$

The extent to which these numbers specify the properties of the matrix  $A$ , will be apparent from the discussion of their invariance properties in the next two subsections. The positive square root of the determinant is in a way the norm of the matrix. Its nonvanishing:  $|A| \neq 0$  is the criterion for  $A$  to be invertible.

Such matrices can be normalized to become unimodular:

$$A \rightarrow |A|^{-1/2} A \quad (3.4.22)$$

The case of singular matrices

$$|A| = a_0^2 - \vec{a}^2 = 0 \quad (3.4.23)$$

calls for comment. We call matrices for which  $|A| = 0$ , but  $A \neq 0$ , null-matrices. Because of their occurrence,  $\mathcal{A}_2$  is not a division algebra. This is in contrast, say, with the set of real quaternions which is a division algebra, since the norm vanishes only for the vanishing quaternion.

The fact that null-matrices are important, stems partly from the indefinite Minkowski metric. However, entirely different applications will be considered later.

We list now some practical rules for operations in  $\mathcal{A}_2$ , presenting them in terms of matrix components rather than the more familiar matrix elements.

To perform matrix multiplications we shall make use of a formula implied by the multiplication rules, Equation 3.4.11:

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = \vec{a} \cdot \vec{b} I + i(\vec{a} \times \vec{b}) \cdot \vec{\sigma} \quad (3.4.24)$$

where  $\vec{a}$  and  $\vec{b}$  are complex vectors.

Evidently, for any two matrices A and B

$$[A, B] = AB - BA = 2i(\vec{a} \times \vec{b}) \cdot \vec{\sigma} \quad (3.4.25)$$

The matrices A and B commute, if and only if

$$\vec{a} \times \vec{b} = 0 \quad (3.4.26)$$

that is, if the vector parts  $\vec{a}$  and  $\vec{b}$  are “parallel” or at least one of them vanishes.

In addition to the internal operations of addition and multiplication, there are external operations on  $\mathcal{A}_2$  as a whole, which are analogous to complex conjugation. The latter operation is an involution, which means that  $(z^*)^* = z$ . Of the three involutions any two can be considered independent.

In  $\mathcal{A}_2$  we have two independent involutions which can be applied jointly to yield a third:

$$A \rightarrow A = a_0 I + \vec{a} \cdot \vec{\sigma} \quad (3.4.27)$$

$$A \rightarrow A^\dagger = a_0^* I + \vec{a}^* \cdot \vec{\sigma} \quad (3.4.28)$$

$$A \rightarrow \tilde{A} = a_0 I - \vec{a} \cdot \vec{\sigma} \quad (3.4.29)$$

$$A \rightarrow \tilde{A}^\dagger = \bar{A} = a_0^* I - \vec{a}^* \cdot \vec{\sigma} \quad (3.4.30)$$

The matrix  $A^\dagger$  is the Hermitian adjoint of A. Unfortunately, there is neither an agreed symbol, nor a term for  $\tilde{A}$  Whitney called it Pauli conjugate, other terms are quaternionic conjugate or hyper-conjugate  $A^\dagger$  (see Edwards, l.c.). Finally  $\bar{A}$  is called complex reflection.

It is easy to verify the rules

$$(AB)^\dagger = B^\dagger A^\dagger \quad (3.4.31)$$

$$(\tilde{A}B) = \tilde{B}\tilde{A} \quad (3.4.32)$$

$$(\overline{AB}) = \bar{B}\bar{A} \quad (3.4.33)$$

According to Equation 3.4.33 the operation of complex reflection maintains the product relation in  $\mathcal{A}_2$  it is an automorphism. In contrast, the Hermitian and Pauli conjugations are anti-automorphic.

It is noteworthy that the three operations  $\sim, \dagger, -$ , together with the identity operator, form a group (the four-group, “Viergruppe”). This is a mark of closure: we presumably left out no important operator on the algebra.

In various contexts any of the three conjugations appears as a generalization of ordinary complex conjugation. Here are a few applications of the conjugation rules.

$$A\tilde{A} = (a_0^2 - \vec{a}^2) 1 = |A| 1 \quad (3.4.34)$$

For invertible matrices

$$A^{-1} = \frac{\tilde{A}}{|A|} \quad (3.4.35)$$

For unimodular matrices we have the useful rule:

$$A^{-1} = \tilde{A} \quad (3.4.36)$$

A Hermitian matrix  $A = A^\dagger$  has real components  $h_0, \vec{h}$ . We define a matrix to be positive if it is Hermitian and has a positive trace and determinant:

$$h_0 > 0, \quad |H| = (h_0^2 - \vec{h}^2) > 0 \quad (3.4.37)$$

If  $H$  is positive and unimodular, it can be parametrized as

$$H = \cosh(\mu/2) 1 + \sinh(\mu/2) \hat{h} \cdot \vec{\sigma} = \exp\{(\mu/2) \hat{h} \cdot \vec{\sigma}\} \quad (3.4.38)$$

The matrix exponential is defined by a power series that reduces to the trigonometric expression. The factor 1/2 appears only for convenience in the next subsection.

In the Pauli algebra, the usual definition  $U^\dagger = U^{-1}$  for a unitary matrix takes the form

$$u_0^* 1 + \vec{u}^* \cdot \vec{\sigma} = |U|^{-1} (u_0 1 - \vec{u} \cdot \vec{\sigma}) \quad (3.4.39)$$

If  $U$  is also unimodular, then

$$u_0^* = u_0 = \text{real} \quad (3.4.40)$$

$$\vec{u}^* = -\vec{u} = \text{imaginary} \quad (3.4.41)$$

and

$$u_0^2 - \vec{u} \cdot \vec{u} = u_0^2 + \vec{u} \cdot \vec{u}^* = 1 \quad (3.4.42)$$

$$U = \cos(\phi/2) 1 - i \sin(\phi/2) \hat{u} \cdot \vec{\sigma} = \exp(-i(\phi/2) \hat{u} \cdot \vec{\sigma}) \quad (3.4.43)$$

A unitary unimodular matrix can be represented also in terms of elements

$$U = \begin{pmatrix} \xi_0 & -\xi_1^* \\ \xi_1 & \xi_0^* \end{pmatrix} \quad (3.4.44)$$

with

$$|\xi_0|^2 + |\xi_1|^2 = 1 \quad (3.4.45)$$

where  $\xi_0, \xi_1$  are the so-called Cayley-Klein parameters. We shall see that both this form, and the axis-angle representation, Equation 3.4.43, are useful in the proper context.

We turn now to the category of normal matrices  $N$  defined by the condition that they commute with their Hermitian adjoint:  $N^\dagger N = N N^\dagger$ . Invoking the condition, Equation 3.4.26, we have

$$\vec{n} \times \vec{n}^* = 0 \quad (3.4.46)$$

implying that  $n^*$  is proportional to  $n$ , that is all the components of  $\vec{n}$  must have the same phase. Normal matrices are thus of the form

$$N = n_0 1 + n \hat{n} \cdot \vec{\sigma} \quad (3.4.47)$$

where  $n_0$  and  $n$  are complex constants and  $\hat{n}$  is a real unit vector, which we call the axis of  $N$ . In particular, any unimodular normal matrix can be expressed as

$$N = \cosh(\kappa/2)1 + \sinh(\kappa/2)\hat{n} \cdot \vec{\sigma} = \exp((\kappa/2)\hat{n} \cdot \vec{\sigma}) \quad (3.4.48)$$

where  $\kappa = \mu - i\phi$ ,  $-\infty < \mu < \infty$ ,  $0 \leq \phi < 4\pi$ , and  $\hat{n}$  is a real unit vector. If  $\hat{n}$  points in the "3" direction, we have

$$N_0 = \exp\left[\left(\frac{\kappa}{2}\right)\sigma_3\right] = \begin{pmatrix} \exp(\frac{\kappa}{2}) & 0 \\ 0 & \exp(-\frac{\kappa}{2}) \end{pmatrix} \quad (3.4.49)$$

Thus the matrix exponentials, Equations 3.4.38, 3.4.43 and 3.4.49, are generalizations of a diagonal matrix and the latter is distinguished by the requirement that the axis points in the z direction.

Clearly the normal matrix, Equation 3.4.49, is a commuting product of a positive matrix like Equation 3.4.38 with  $\hat{h} = \hat{n}$  and a unitary matrix like Equation 3.4.43, with  $\hat{u} = \hat{n}$

$$N = HU = UH \quad (3.4.50)$$

The expressions in Equation 3.4.50 are called the polar forms of  $N$ , the name being chosen to suggest that the representation of  $N$  by  $H$  and  $U$  is analogous to the representation of a complex number  $z$  by a positive number  $r$  and a phase factor:

$$z = r \exp(-i\phi/2) \quad (3.4.51)$$

We shall show that, more generally, any invertible matrix has two unique polar forms

$$A = HU = UH' \quad (3.4.52)$$

but only the polar forms of normal matrices display the following equivalent special features:

1.  $H$  and  $U$  commute
2.  $\hat{h} = \hat{u} = \hat{n}$
3.  $H' = H$

We see from the polar decomposition theorem that our emphasis on positive and unitary matrices is justified, since all matrices of  $\mathcal{A}_2$  can be produced from such factors. We proceed now to prove the theorem expressed in Equation 3.4.52 by means of an explicit construction.

First we form the matrix  $AA^\dagger$ , which is positive by the criteria 3.4.36:

$$a_0 a_0^* + \vec{a} \cdot \vec{a}^* > 0 \quad (3.4.53)$$

$$|A| |A^\dagger| > 0 \quad (3.4.54)$$

Let  $AA^\dagger$  be expressed in terms of an axis  $\hat{h}$  and the hyperbolic angle  $\mu$ :

$$\begin{aligned} AA^\dagger &= b(\cosh \mu 1 + \sinh \mu \hat{h} \cdot \vec{\sigma}) \\ &= b \exp(\mu \hat{h} \cdot \vec{\sigma}) \end{aligned}$$

where  $b$  is a positive constant. We claim that the Hermitian component of  $A$  is the positive square root of ???

$$H = (AA^\dagger)^{1/2} = b^{1/2} \exp\left(\frac{\mu}{2} \hat{h} \cdot \vec{\sigma}\right) \quad (3.4.55)$$

with

$$U = H^{-1}A, \quad A = HU \quad (3.4.56)$$

That  $U$  is indeed unitary is easily verified:

$$U^\dagger = A^\dagger H^{-1}, \quad U^{-1} = A^{-1}H \quad (3.4.57)$$

and these expressions are equal by Equation 3.4.55.

From Equation 3.4.56 we get

$$A = U(U^{-1}HU) \quad (3.4.58)$$

and

$$A = UH' \quad \text{with} \quad H' = U^{-1}HU \quad (3.4.59)$$

It remains to be shown that the polar forms 3.4.56 are unique. Suppose indeed, that for a particular  $A$  we have two factorizations

$$A = HU = H_1U_1 \quad (3.4.60)$$

then

$$AA^\dagger = H^2 = H_1^2 \quad (3.4.61)$$

But, since  $AA^\dagger$  has a unique positive square root,  $H_1 = H$ , and

$$U = H_1^{-1}A = H^{-1}A = U \quad \text{q.e.d.} \quad (3.4.62)$$

Polar forms are well known to exist for any  $n \times n$  matrix, although proofs of uniqueness are generally formulated for abstract transformations rather than for matrices, and require that the transformations be invertible.

### 3.4.3 The restricted Lorentz group

Having completed the classification of the matrices of  $\mathcal{A}_2$ , we are ready to interpret them as operators and establish a connection with the Lorentz group. The straightforward procedure would be to introduce a 2-dimensional complex vector space  $\mathcal{V}(\mathbb{C})$ . By using the familiar bra-ket formalism we write

$$A|\xi\rangle = |\xi'\rangle \quad (3.4.63)$$

$$A^\dagger\langle\xi| = \langle\xi'| \quad (3.4.64)$$

The two-component complex vectors are commonly called spinors. We shall study their properties in detail in Section 5. The reason for this delay is that the physical interpretation of spinors is a

subtle problem with many ramifications. One is well advised to consider at first situations in which the object to be operated upon can be represented by a  $2 \times 2$  matrix.

The obvious choice is to consider Hermitian matrices, the components of which are interpreted as relativistic four-vectors. The connection between four-vectors and matrices is so close that it is often convenient to use the same symbol for both:

$$A = a_0 1 + \vec{a} \cdot \vec{\sigma} \quad (3.4.65)$$

$$A = \{a_0, \vec{a}\} \quad (3.4.66)$$

We have

$$a_0^2 - \vec{a}^2 = |A| = \frac{1}{2} \text{Tr}(A\bar{A}) \quad (3.4.67)$$

or more generally

$$a_0 b_0 - \vec{a} \cdot \vec{b} = \frac{1}{2} \text{Tr}(A\bar{B}) \quad (3.4.68)$$

A Lorentz transformation is defined as a linear transformation

$$\{a_0, \vec{a}\} = \mathcal{L} \{a'_0, \vec{a}'\} \quad (3.4.69)$$

that leaves the expression 3.4.68 and hence also 3.4.67 invariant. We require moreover that the sign of the “time component”  $a_0$  be invariant (orthochronic Lorentz transformation  $L^\uparrow$ ) and that the determinant of the  $4 \times 4$  matrix  $\mathcal{L}$  be positive (proper Lorentz transformation  $L_+$ ). If both conditions are satisfied, we speak of the restricted Lorentz group  $L_+^\uparrow$ . This is the only one to be of current interest for us, and until further notice “Lorentz group” is to be interpreted in this restricted sense.

Note that  $A$  can be interpreted as any of the four-vectors discussed in Section 3.2:  $R = \{r, \vec{r}\}$

$$K = \{k_0, \vec{k}\}, \quad P = \{p_0, \vec{p}\} \quad (3.4.70)$$

Although these vectors and their matrix equivalents have identical transformation properties, they differ in the possible range of their determinants. A negative  $|P|$  can arise only for an unphysical imaginary rest mass. By contrast, a positive  $R$  corresponds to a time-like displacement pointing toward the future, an  $R$  with a negative  $|R|$  to a space-like displacement and  $|R| = 0$  is associated with the light cone. For the wave vector we have by definition  $|K| = 0$ .

To describe a Lorentz transformation in the Pauli algebra we try the “ansatz”

$$A' = VAW \quad (3.4.71)$$

with  $|V| = |W| = 1$  in order to preserve  $|A|$ . Reality of the vector, i.e., hermiticity of the matrix  $A$  is preserved if the additional condition  $W = V^\dagger$  is satisfied. Thus the transformation

$$A' = VAV^\dagger \quad (3.4.72)$$

leaves expression 3.4.67 invariant. It is easy to show that 3.4.68 is invariant as well.

The complex reflection  $\bar{A}$  transforms as

$$\bar{A}' = \bar{V}\bar{A}\bar{V} \quad (3.4.73)$$

and the product of two four-vectors:

$$\begin{aligned} (A\bar{B})' &= VAV^\dagger \bar{V}\bar{B}\bar{V} \\ &= V(A\bar{B})V^{-1} \end{aligned}$$

This is a so-called similarity transformation. By taking the trace of Equation ??? we confirm that the inner product 3.4.68 is invariant under 3.4.73. We have to remember that a cyclic permutation does not affect the trace of a product of matrices. Thus Equation 3.4.73 indeed induces a Lorentz transformation in the four-vector space of  $A$ .

It is well known that the converse statement is also true: to every transformation of the restricted Lorentz group  $L_+^\uparrow$  there are associated two matrices differing-only by sign (their parameters  $\phi$  differ by  $2\pi$ ) in such a fashion as to constitute a two-to-one homomorphism between the group of unimodular matrices  $\mathcal{SL}(2, C)$  and the group  $L_+^\uparrow$ . It is said also that  $\mathcal{SL}(2, C)$  provides a two-valued representation of  $L_+^\uparrow$ . We shall prove this statement by demonstrating explicitly the connection between the matrices  $V$  and the induced, or associated group operations.

We note first that  $A$  and  $\bar{A}$  correspond in the tensor language to the contravariant and the covariant representation of a vector. We illustrate the use of the formalism by giving an explicit form for the inverse of 3.4.73

$$A = V^{-1}A'V^{\dagger-1} \equiv \tilde{V}A'\tilde{V} \quad (3.4.74)$$

We invoke the polar decomposition theorem Equation 3.4.50 of Section 3.4.2 and note that it is sufficient to establish this connection for unitary and positive matrices respectively.

Consider at first

$$A' = UAU^\dagger \equiv UAU^{-1} \quad (3.4.75)$$

with

$$\begin{aligned} U\left(\hat{u}, \frac{\phi}{2}\right) &\equiv \exp\left(-\frac{i\phi}{2}\hat{u} \cdot \vec{\sigma}\right) \\ u_1^2 + u_2^2 + u_3^2 &= 1, \quad 0 \leq \phi < 4\pi \end{aligned}$$

The set of all unitary unimodular matrices described by Equation ??? form a group that is commonly called  $SU(2)$ .

Let us decompose  $\vec{a}$  :

$$\vec{a} = \vec{a}_\parallel + \vec{a}_\perp \quad (3.4.76)$$

$$\vec{a}_\parallel = (\vec{a} \cdot \hat{u})\hat{u}, \quad \vec{a}_\perp = \vec{a} - \vec{a}_\parallel = \hat{u} \times (\vec{a} \times \hat{u}) \quad (3.4.77)$$

It is easy to see that Equation 3.4.75 leaves  $a_0$  and  $a_\parallel$  invariant and induces a rotation around  $\hat{u}$  by an angle  $\phi$  :  $R\{\hat{u}, \phi\}$ .

Conversely, to every rotation  $R\{\hat{u}, \phi\}$  there correspond two matrices:

$$U\left(\hat{u}, \frac{\phi}{2}\right) \quad \text{and} \quad U\left(\hat{u}, \frac{\phi+2\pi}{2}\right) = -U\left(\hat{u}, \frac{\phi}{2}\right) \quad (3.4.78)$$

We have  $1 \rightarrow 2$  homomorphism between  $\mathcal{SO}(3)$  and  $\mathcal{SU}(2)$ , the latter is said to be a two-valued representation of the former. By establishing this correspondence we have solved the problem of parametrization formulated on page 13. The nine parameters of the orthogonal  $3 \times 3$  matrices are reduced to the three independent ones of  $U\left(\hat{u}, \frac{\phi}{2}\right)$ . Moreover we have the simple result

$$U^n = \exp\left(-\frac{in\phi}{2} \hat{u} \cdot \vec{\sigma}\right) \quad (3.4.79)$$

which reduces to the de Moivre theorem if  $\hat{n} \cdot \vec{\sigma} = \sigma_3$

Some comment is in order concerning the two-valuedness of the  $\mathcal{SU}(2)$  representation. This comes about because of the use of half angles in the algebraic formalism which is deeply rooted in the geometrical structure of the rotation group. (See the Rodrigues-Hamilton theorem in Section 2.2.)

Whereas the two-valuedness of the  $\mathcal{SU}(2)$  representation does not affect the transformation of the A vector based on the bilateral expression 3.4.75, the situation will be seen to be different in the spinorial theory based on Equation 3.4.63, since under certain conditions the sign of the spinor  $|\xi\rangle$  is physically meaningful.

The above discussion of the rotation group is incomplete even within the classical theory. The rotation  $R\{\hat{u}, \phi\}$  leaves vectors along  $\hat{u}$  unaffected. A more appropriate object to be rotated is the Cartesian triad, to be discussed in Section 5.

We consider now the case of a positive matrix  $V = H$

$$A' = HAH \quad (3.4.80)$$

with

$$H = \exp\left(\frac{\mu}{2} \hat{h} \cdot \sigma\right) \quad (3.4.81)$$

$$h_1^2 + h_2^2 + h_3^2 = 1, \quad -\infty < \mu < \infty \quad (3.4.82)$$

We decompose  $\vec{a}$  as

$$\vec{a} = a\hat{h} + \vec{a}_\perp \quad (3.4.83)$$

and using the fact that  $(\vec{a} \cdot \vec{\sigma})$  and  $(\vec{b} \cdot \vec{\sigma})$  commute for  $\vec{a} \parallel \vec{b}$  and anticommute for  $\vec{a} \perp \vec{b}$ , we obtain

$$A' = \exp\left(\frac{\mu}{2} \hat{h} \cdot \sigma\right) (a_0 1 + a\hat{h} \cdot \sigma + \vec{a}_\perp \cdot \sigma) \exp\left(\frac{\mu}{2} \hat{h} \cdot \sigma\right) \quad (3.4.84)$$

$$= \exp(\mu \hat{h} \cdot \sigma) (a_0 1 + \vec{a} \hat{h} \cdot \sigma) + \vec{a}_\perp \cdot \sigma \quad (3.4.85)$$

Hence

$$a'_0 = \cosh \mu a_0 + \sinh \mu a \quad (3.4.86)$$

$$a' = \sinh \mu a_0 + \cosh \mu a \quad (3.4.87)$$

$$\vec{a}'_\perp = \vec{a}_\perp \quad (3.4.88)$$

This is to be compared with Table 3.1, but remember that we have shifted from the passive to the active interpretation, from alias to alibi.

Positive matrices with a common axis form a group (Wigner's "little group"), but in general the product of Hermitian matrices with different axes are not Hermitian. There arises a unitary factor, which is the mathematical basis for the famous Thomas precession.

Let us consider now a normal matrix

$$V = N = H\left(\hat{k}, \frac{\mu}{2}\right) U\left(\hat{k}, \frac{\phi}{2}\right) = \exp\left(\frac{\mu - i\phi}{2} \hat{n} \cdot \sigma\right) \quad (3.4.89)$$



where we have the commuting product of a rotation and a Lorentz transformation with the same axis  $\hat{n}$ . Such a constellation is called a Lorentz 4-screw

An arbitrary sequence of pure Lorentz transformations and pure rotations is associated with a pair of matrices  $V$  and  $-V$ , which in the general case is of the form

$$H\left(\hat{h}, \frac{\mu}{2}\right) U\left(\hat{u}, \frac{\phi}{2}\right) = U\left(\hat{u}, \frac{\phi}{2}\right) H'\left(\hat{h}', \frac{\mu}{2}\right) \quad (3.4.90)$$

According to Equation 3.4.59 of Section 3.4.2,  $H$  and  $H'$  are connected by a similarity transformation, which does not affect the angle  $\mu$  but only the axis of the transformation. (See the next section.)

This matrix depends on the 6 parameters,  $\hat{h}, \mu, \hat{u}, \phi$  and thus we have solved the general problem of parametrization mentioned above.

For a normal matrix  $\hat{h} = \hat{u} = \hat{n}$  and the number of parameters is reduced to 4.

Our formalism enables us to give a closed form for two arbitrary normal matrices and the corresponding 4-screws.

$$[N, N'] = 2i \sinh \frac{\kappa}{2} \sinh \frac{\kappa'}{2} (\hat{n} \times \hat{n}') \cdot \vec{\sigma} \quad (3.4.91)$$

where  $\kappa = \mu - i\phi, \kappa' = \mu' - i\phi'$

In the literature the commutation relations are usually given in terms of infinitesimal operators which are defined as follows:

$$U\left(\hat{u}_k, \frac{d\phi}{2}\right) = 1 - \frac{i}{2} d\phi \sigma_k = 1 + d\phi I_k \quad (3.4.92)$$

$$I_k = -\frac{i}{2} \sigma_k \quad (3.4.93)$$

$$H\left(\hat{h}_k, \frac{d\mu}{2}\right) = 1 + \frac{d\mu}{2} \sigma_k = 1 + d\mu L_k \quad (3.4.94)$$

$$L_k = \frac{1}{2} \sigma_k \quad (3.4.95)$$

The commutation relations are

$$[I_1, I_2] = I_3 \quad (3.4.96)$$

$$[L_1, L_2] = -I_3 \quad (3.4.97)$$

$$[L_1, I_2] = L_3 \quad (3.4.98)$$

and cyclic permutations.

It is a well known result of the Lie-Cartan theory of continuous group that these infinitesimal generators determine the entire group. Since we have represented these generators in  $\mathcal{SL}(2, C)$ , we have completed the demonstration that the entire group  $L_+^\uparrow$  is accounted for in our formalism.

### 3.4.4 Similarity classes and canonical forms of active transformations

It is evident that a Lorentz transformation induced by a matrix  $H\left(\hat{h}, \frac{\mu}{2}\right)$  assumes a particularly simple form if the z-axis of the coordinate system is placed in the direction of  $\hat{h}$ . The diagonal matrix  $H\left(\hat{z}, \frac{\mu}{2}\right)$  is said to be the canonical form of the transformation. This statement is a special case of the problem of canonical forms of linear transformations, an important chapter in linear algebra.

Let us consider a linear mapping in a vector space. A particular choice of basis leads to a matrix representation of the mapping, and representations associated with different frames are connected by similarity transformations. Let  $A_1$  be an arbitrary and  $S$  an invertible matrix. A similarity transformation is effected on  $A$ , by

$$A_2 = S A_1 S^{-1} \quad (3.4.99)$$

Matrices related by similarity transformation are called similar, and matrices similar to each other constitute a similarity class.

In usual practice the mapping-refers to a vector space as in Equation 3.4.63 of Section 3.4.3:

$$A_1 |\xi\rangle_1 = |\xi'\rangle_1 \quad (3.4.100)$$

The subscript refers to the basis “1.” A change of basis  $\Sigma_1 \rightarrow \Sigma_2$  is expressed as

$$|\xi\rangle_2 = S|\xi\rangle_1, \quad |\xi'\rangle_2 = S|\xi'\rangle_1 \quad (3.4.101)$$

Inserting into Equation 3.4.100 we obtain

$$A_1 S^{-1} |\xi\rangle_2 = S^{-1} |\xi\rangle_2 \quad (3.4.102)$$

and hence

$$A_2 |\xi\rangle_2 = |\xi\rangle_2 \quad (3.4.103)$$

where  $A_2$  is indeed given by Equation 3.4.99

The procedure we have followed thus far to represent Lorentz transformations in  $A_2$  does not quite follow this standard pattern.

We have been considering mappings of the space of fourvectors which in turn were represented as  $2 \times 2$  complex matrices. Thus both operators and operands are matrices of  $A_2$ . In spite of this difference in interpretation, the matrix representations in different frames are still related according to Equation \label{100}.

This can be shown as follows. Consider a unimodular matrix  $A$ , that induces a Lorentz transformation in P-space, whereby the matrices refer to the basis  $\Sigma_1$ :

$$P'_1 = A_1 P_1 A_1^\dagger \quad (3.4.104)$$

We interpret Equation 3.4.104 in the active sense as a linear mapping of P-space on itself that corresponds physically to some dynamic process that alters P in a linear way.

We shall see in Section 4 that the Lorentz force acting on a charged particle during the time  $dt$  can be indeed considered as an active Lorentz transformation. (See also page 26.)

The process has a physical meaning independent of the frame of the observer, but the matrix representations of  $P$ ,  $P'$  and of  $A$  depend on the frame. The four-momenta in the two frames are connected by a Lorentz transformation interpreted in the passive sense:

$$P_2 = S P_1 S^\dagger \quad (3.4.105)$$

$$P_2 = S P'_1 S^\dagger \quad (3.4.106)$$

with  $|S| = 1$ . Solving for  $P$ ,  $P'$  and inserting into Equation 3.4.104 we obtain

$$S^{-1} P'_2 \tilde{S}^\dagger = A_1 S^{-1} P_2 \tilde{S}^\dagger A_1^\dagger S \quad (3.4.107)$$

or

$$P'_2 = A_2 P_2 A_1^\dagger \quad (3.4.108)$$

where  $A_2$  and  $A_1$  are again connected by the similarity transformation 3.4.99

We may apply the polar decomposition theorem to the matrix  $S$ . In the special case that  $S$  is unitary, we speak of a unitary similarity transformation corresponding to the rotation of the coordinate system discussed at the onset of this section. However, the general case will lead us to less obvious physical applications.

The above considerations provide sufficient motivation to examine the similarity classes of  $A_2$ . We shall see that all of them have physical applications, although the interpretation of singular mappings will be discussed only later

The similarity classes can be characterized in several convenient ways. For example, one may use two independent similarity invariants shared by all the matrices  $A = a_0 I + \vec{a} \cdot \vec{\sigma}$  in the class. We shall find it convenient to choose

1. the determinant  $|A|$ , and

2. the quantity  $\vec{a}^2$

The trace is also a similarity invariant, but it is not independent:  $a_0^2 = |A| + \vec{a}^2$

Alternatively, one can characterize the whole class by one representative member of it, some matrix  $A_0$  called the canonical form for the class (See Table 3.2).

We proceed at first to characterize the similarity classes in terms of the invariants 1 and 2. We recall that a matrix  $A$  is invertible if  $|A| \neq 0$  and singular if  $|A| = 0$ . Without significant loss of generality, we can normalize the invertible matrices of  $\mathcal{A}_2$  to be unimodular, so that we need discuss only classes of singular and of unimodular matrices. As a second invariant to characterize a class, we choose  $\vec{a} \cdot \vec{a}$  and we say that a matrix  $A$  is axial if  $\vec{a} \cdot \vec{a} \neq 0$ . In this case, there exists a unit vector  $\hat{a}$  (possibly complex) such that  $\vec{a} = a \cdot \hat{a}$  where  $a$  is a complex constant. The unit vector  $\hat{a}$  is called the axis of  $A$ . Conversely, the matrix  $A$  is non-axial if  $\vec{a} \cdot \vec{a} = 0$ , the vector  $\vec{a}$  is called isotropic or a null-vector, it cannot be expressed in terms of an axis.

The concept of axis as here defined is the generalization of the real axis introduced in connection with normal matrices on page 33. The usefulness of this concept is apparent from the following theorem:

**Theorem 1.** For any two unit vectors  $\hat{v}_1$ , and  $\hat{v}_2$ , real or complex, there exists a matrix  $S$  such that

$$\hat{v}_2 \cdot \vec{\sigma} = S \hat{v}_1 \cdot \vec{\sigma} S^{-1} \quad (3.4.109)$$

*Proof.* We construct one such matrix  $S$  from the following considerations. If  $\hat{v}_1$ , and  $\hat{v}_2$  are real, then let  $S$  be the unitary matrix that rotates every vector by an angle  $\pi$  about an axis which bisects the angle between  $\hat{v}_1$ , and  $\hat{v}_2$ :

$$S = -i \hat{s} \cdot \vec{\sigma} \quad (3.4.110)$$

where

$$\hat{s} = \frac{\hat{v}_1 + \hat{v}_2}{\sqrt{2\hat{v}_1 \cdot \hat{v}_2 + 2}} \quad (3.4.111)$$

Even if  $\hat{v}_1$ , and  $\hat{v}_2$  are not real, it is easily verified that  $S$  as given formally by Equations 3.4.110 and 3.4.111, does indeed send  $\hat{v}_1$  to  $\hat{v}_2$ . Naturally  $S$  is not unique; for instance, any matrix of the form

$$S = \exp \left\{ \left( \frac{\mu_2}{2} - i \frac{\phi_2}{2} \right) \vec{v}_2 \cdot \vec{\sigma} \right\} (-i \hat{s} \cdot \vec{\sigma}) \exp \left\{ \left( \frac{\mu_1}{2} - i \frac{\phi_1}{2} \right) \vec{v}_1 \cdot \vec{\sigma} \right\} \quad (3.4.112)$$

will send  $\hat{v}_1$  to  $\hat{v}_2$

This construction fails only if

$$\hat{v}_1 \cdot \hat{v}_2 + 1 = 0 \quad (3.4.113)$$

that is for the transformation  $\hat{v}_1 \rightarrow -\hat{v}_2$ . In this trivial case we choose

$$S = -i \hat{s} \cdot \vec{\sigma}, \quad \text{where} \quad \hat{s} \perp \vec{v}_1 \quad (3.4.114)$$

Since in the Pauli algebra diagonal matrices are characterized by the fact that their axis is  $\hat{x}_3$ , we have proved the following theorem:

**Theorem 2.** All axial matrices are diagonalizable, but normal matrices and only normal matrices are diagonalizable by a unitary similarity transformation.

The diagonal forms are easily ascertained both for singular and the unimodular cases. (See Table 3.2.) Because of their simplicity they are called also canonical forms. Note that they can be multiplied by any complex number in order to get all of the axial matrices of  $\mathcal{A}_2$

The situation is now entirely clear: the canonical forms show the nature of the mapping; a unitary similarity transformation merely changes the geometrical orientation of the axis. The angle of circular and hyperbolic rotation specified by  $a_0$  is invariant. A general transformation complexifies the axis. This situation comes about if in the polar form of the matrix  $A = HU$ , the factors have distinct real axes, and hence do not commute.

There remains to deal with the case of nonaxial matrices. Consider  $A = \vec{a} \cdot \vec{\sigma}$  with  $\vec{a}^2 = 0$ . Let us decompose the isotropic vector  $\vec{a}$  into real and imaginary parts:

$$\vec{a} = \vec{\alpha} + i\vec{\beta} \quad (3.4.115)$$

Hence  $\vec{\alpha}^2 - \vec{\beta}^2 = 0$  and  $\alpha \cdot \beta = 0$ . Since the real and the imaginary parts of  $a$  are perpendicular, we can rotate these directions by a unitary similarity transformation into the x- and y-directions respectively. The transformed matrix is

$$\frac{\alpha}{2}(\sigma_1 + i\sigma_2) = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix} \quad (3.4.116)$$

with a positive. A further similarity transformation with

$$S = \begin{pmatrix} \alpha^{-1/2} & 0 \\ 0 & \alpha^{1/2} \end{pmatrix} \quad (3.4.117)$$

transforms Equation 3.4.116 into the canonical form given in Table 3.2.

As we have seen in Section 3.4.3 all unimodular matrices induce Lorentz transformations in Minkowski, or four-momentum space. According to the results summarized in Table 3.2, the mappings induced by axial matrices can be brought by similarity transformations into so-called Lorentz four-screws consisting of a circular and hyperbolic rotation around the same axis, or in other words: a rotation around an axis, and a boost along the same axis.

What about the Lorentz transformation induced by a nonaxial matrix? The nature of these transformations is very different from the common case, and constitutes an unusual limiting situation. It is justified to call it an exceptional Lorentz transformation. The special status of these transformations was recognized by Wigner in his fundamental paper on the representations of the Lorentz group.

The present approach is much more elementary than Wigner's, both from the point of view of mathematical technique, and also the purpose in mind. Wigner uses the standard algebraic technique of elementary divisors to establish the canonical Jordan form of matrices. We use, instead a specialized technique adapted to the very simple situation in the Pauli algebra. More important, Wigner was concerned with the problem of representations of the inhomogeneous Lorentz group, whereas we consider the much simpler problem of the group structure itself, mainly in view of application to the electromagnetic theory.

The intuitive meaning of the exceptional transformations is best recognized from the polar form of the generating matrix. This can be carried out by direct application of the method discussed at the end of the last section. It is more instructive, however, to express the solution in terms of (circular and hyperbolic) trigonometry.

We ask for the conditions the polar factors have to satisfy in order that the relation

$$1 + \hat{a} \cdot \vec{\sigma} = H\left(\hat{h}, \frac{\mu}{2}\right) U\left(\hat{u}, \frac{\phi}{2}\right) \quad (3.4.118)$$

should hold with  $\mu \neq 0$ ,  $\phi \neq 0$ . Since all matrices are unimodular, it is sufficient to consider the equality of the traces:

$$\frac{1}{2} \text{Tr } A = \cosh\left(\frac{\mu}{2}\right) \cos\left(\frac{\phi}{2}\right) - i \sinh\left(\frac{\mu}{2}\right) \sin\left(\frac{\phi}{2}\right) \hat{h} \cdot \hat{u} = 1 \quad (3.4.119)$$

This condition is satisfied if and only if

$$\hat{h} \cdot \hat{u} = 0 \quad (3.4.120)$$

and

$$\cosh\left(\frac{\mu}{2}\right) \cos\left(\frac{\phi}{2}\right) = 1 \quad (3.4.121)$$

The axes of circular and hyperbolic rotation are thus perpendicular, to each other and the angles of these rotations are related in a unique fashion: half of the circular angle is the so-called Gudermannian function of half of the hyperbolic angle

$$\frac{\phi}{2} = \text{gd}\left(\frac{\mu}{2}\right) \quad (3.4.122)$$

However, if  $\mu$  and  $\phi$  are infinitesimal, we get

$$\left(1 + \frac{\mu^2}{2} + \dots\right) \left(1 + \frac{\phi^2}{2} + \dots\right) = 1, \text{ i.e.} \quad (3.4.123)$$

$$\mu^2 - \phi^2 = 0 \quad (3.4.124)$$

We note finally that products of exceptional matrices need not be exceptional, hence exceptional Lorentz transformations do not form a group.

In spite of their special character, the exceptional matrices have interesting physical applications, both in connection with the electromagnetic field as discussed in Section 4, and also for the construction of representations of the inhomogeneous Lorentz group [Pae69, Wig39].

We conclude by noting that the canonical forms of Table 3.2 lend themselves to express the powers  $A_0^k$  in simple form.

For the axial singular matrix we have

$$A_0^2 = A \quad (3.4.125)$$

These projection matrices are called idempotent. The nonaxial singular matrices are nilpotent:

$$A_0^2 = 0 \quad (3.4.126)$$

The exceptional matrices (unimodular nonaxial) are raised to any power  $k$  (even non-real) by the formula

$$A^k = 1^k (1 + k\vec{a} \cdot \vec{\sigma}) \quad (3.4.127)$$

$$= 1^k \exp(k\vec{a} \cdot \vec{\sigma}) \quad (3.4.128)$$

For integer  $k$ , the factor  $1^k$  becomes unity. The axial unimodular case is handled by formulas that are generalizations of the well known de Moivre formulas:

$$A^k = 1^k \exp\left(k\frac{\kappa}{2} + kl2\pi i\right) \quad (3.4.129)$$

where  $l$  is an integer. For integer  $k$ , Equation 3.4.129 reduces to

$$A^k = \exp\left(k\left(\frac{\kappa}{2}\right)\vec{a} \cdot \vec{\sigma}\right) \quad (3.4.130)$$

In connection with these formulae, we note that for positive  $A$  ( $\phi = 0$  and a real), there is a unique positive  $m^{\text{th}}$  root of  $A$ :

$$A = \exp\left\{\left(\frac{\mu}{2}\right)\hat{a} \cdot \vec{\sigma}\right\} \quad (3.4.131)$$

$$A^{1/m} = \exp\left\{\left(\frac{\mu}{2m}\right)\hat{a} \cdot \vec{\sigma}\right\} \quad (3.4.132)$$

The foregoing results are summarized in Table 3.2.

	Unimodular $ A  = 1$	Singular $ A  = 0$
<p>Axial</p> <p><math>\vec{a} = a\hat{a}</math></p>	<p><math>A = \exp(\frac{\kappa}{2}\vec{a} \cdot \vec{\sigma})</math></p> <p><math>A_0 = \begin{pmatrix} \exp(\frac{\kappa}{2}) &amp; 0 \\ 0 &amp; \exp(\frac{-\kappa}{2}) \end{pmatrix}</math></p> <p><math>\kappa = \mu - i\phi</math></p> <p><math>-\infty &lt; \mu &lt; \infty</math></p> <p><math>0 \leq \phi &lt; 4\pi</math></p>	<p><math>A = \frac{1}{2}(1 \pm \vec{a} \cdot \vec{\sigma})</math></p> <p><math>A_0 = \begin{pmatrix} 1 &amp; 0 \\ 0 &amp; 0 \end{pmatrix}, \quad \begin{pmatrix} 0 &amp; 0 \\ 0 &amp; 1 \end{pmatrix}</math></p>
<p>Non-axial</p> <p><math>\vec{a}^2 = 0</math></p> <p><math>\vec{a} = \frac{1}{2}(\hat{e}_1 \pm \hat{e}_2)</math></p>	<p><math>A = 1 + \vec{a} \cdot \vec{\sigma} = \exp(\vec{a} \cdot \vec{\sigma})</math></p> <p><math>A_0 = \begin{pmatrix} 1 &amp; 1 \\ 0 &amp; 1 \end{pmatrix}</math></p>	<p><math>A = \vec{a} \cdot \vec{\sigma}</math></p> <p><math>A_0 = \begin{pmatrix} 0 &amp; 1 \\ 0 &amp; 0 \end{pmatrix}, \quad \begin{pmatrix} 1 &amp; 0 \\ 0 &amp; 0 \end{pmatrix}</math></p>

Table 3.2: Canonical Forms for the Similarity classes of  $A_2$

This page titled [3.4: The Pauli Algebra](#) is shared under a [CC BY-NC-SA](#) license and was authored, remixed, and/or curated by [László Tisza](#) (MIT OpenCourseWare) .