APPLIED GEOMETRIC ALGEBRA

László Tisza
Massachusetts Institute of Technology

LibreTexts™

Applied Geometric Algebra

The LibreTexts libraries are Powered by NICE CXOne and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

This text was compiled on 04/15/2025

TABLE OF CONTENTS

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing****.*

CHAPTER OVERVIEW

1: Introduction

1.1: Introduction

Mathematical physics operates with a combination of the three major branches of mathematics: geometry, algebra and infinitesimal analysis. The interplay of these elements has undergone a considerable change since the turn of the century.

In classical physics, analysis, in particular differential equations, plays a central role. This formalism is supplemented most harmoniously by Gibbsian vector algebra and calculus to account for the. spatial, geometric properties of particles and fields.

Few theorists have labored as patiently as Gibbs to establish the simplest possible formalism to meet a particular need. His success can be assessed by the fact that — almost a century later — his calculus, in the original notation, is in universal use. However, once full advantage has been taken of all simplifications permitted in the classical theory, there did not remain the reserve capacity to deal with quantum mechanics and relativity. The gap in the classical algebraic framework was supplemented by Minkowski tensors and Hilbert vectors, matrix algebras, spinors, Lie groups and a host of other constructs.

Unfortunately, the advantages derived from the increased power and sophistication of the new algebraic equipment are marred by side effects. There is a proliferation of overlapping techniques with too little standardization.

Also, while the algebras maintain a vaguely geometrical character, the “spaces” referred to are mathematical abstractions with but scant appeal to ordinary spatial intuition.

These features are natural consequences of rapid growth which can be remedied by consolidation and streamlining; the problem is to adapt the Gibbsian principle of economy to more demanding conditions.

This course is a progress report on such a project. Our emphasis on formalism does not mean neglect of conceptual problems. In fact, the most rewarding aspect of our consolidation is the resulting conceptual clarity.

The central idea of the present approach is that group theory provides us with a flexible and comprehensive framework to describe and classify fundamental physical processes. It is hardly necessary to argue that this method is indeed important in modern physics, yet its potentialities are still far from exhausted. This may stem from the tendency to resort to group theory only for difficult problems when other methods fail, or become too cumbersome. As a result, the discussions tend to be complicated and abstract.

The distinctive feature of the present method is to start with elementary problems, but treat them from an advanced view point, and to reconcile intuitive interpretation with a smooth transition to deeper problems. By focusing on group theory from the outset, we can make full use of its unifying power.

As mentioned above, this is a report on work in progress. Although we shall confine ourselves to problems that fall within the scope of the “consolidated” theory, we shall be in a position to discuss some of the conceptual problems of quantum mechanics.

CHAPTER OVERVIEW

2: Algebraic Preliminaries

2.1: Groups

When group theory was introduced into the formalism of quantum mechanics in the late 1920's to solve abstruse spectroscopic problems, it was considered to be the hardest and the most unwelcome branch of mathematical physics. Since that time group theory has been simplified and popularized and it is widely practiced in many branches of physics, although this practice is still limited mostly to difficult problems where other methods fail.

In contrast, I wish to emphasize that group theory has also simple aspects which prove to be eminently useful for the systematic presentation of the material of this course.

Postponing for a while the precise definition,- we state somewhat loosely that we call a set of elements a group if it is closed with respect to a single binary operation usually called multiplication. This multiplication is, in general not to be taken in the common sense of the word, and need not Mbe commutative. It is, however, associative and invertible.

The most common interpretation of such an operation is a transformation. Say, the translations and rotations of Euclidean space; the transformations that maintain the symmetry of an object such as a cube or a sphere. The transformations that connect the findings of different inertial observers with each other.

With some training we recognize groups anywhere we look. Thus we can consider the group of displacement of a rigid body, and also any particular subset of these displacements' that arise in the course of a particular motion.

We shall see indeed, that group theory provides a terminology that is invaluable for the precise and intuitive discussion of the most elementary and fundamental principles of physics. As to the discussion of specific problems we shall concentrate on those that can be adequately handled by stretching the elementary methods, and we shall not invoke advanced group theoretical results. Therefore we turn now to a brief outline of the principal definitions and theorems that we shall need in the sequel.

Let us consider a set of elements $A, B, C, \ldots$ and a binary operation that is traditionally called "multiplication". We refer to this set as a group $\mathcal{G}$ if the following requirements are satisfied.

1. For any ordered pair, A, B there is a product AB = C. The set is closed with respect to multiplication.
2. The associative law holds: (AB)C = A(BC).
3. There is a unit element $E \in \mathcal{G}$ such that $EA = AE = A$ for all $A \in \mathcal{G}$. .
4. For each element A there is an inverse A^{-1} with $A^{-1}A = AA^{-1} = E$.

The multiplication need not be commutative. If it is, the group is called **Abelian**.

The number of elements in $\mathcal{G}$ is called the **order of the group**. This may be finite or infinite, denumerable or continuous.

If a subset of $\mathcal{G}$ satisfies the group postulates, it is called a **subgroup**.

2.1.1 Criterion for Subgroups

If a subset of the elements of a group of finite order $\mathcal{G}$ is closed under multiplication, then it is a subgroup of $\mathcal{G}$.

Prove that the group postulates are satisfied. Discuss the case of groups of infinite order.

In order to explain the use of these concepts we list a few examples of sets chosen from various branches of mathematics of interest in physics, for which the group postulates are valid.

Examples

1. The set of integers (positive, negative and zero) is an Abelian group of infinite order where the common addition plays the role of multiplication. Zero serves as the unit and the inverse of a is $-a$.

2. The set of permutations of n objects, called also the **symmetric group** $S(n)$, is of order n!. It is non-Abelian for $n > 2$.

3. The infinite set of $n \times n$ matrices with non-vanishing determinants. The operation is matrix multiplication; it is in general non-commutative.

4. The set of covering operations of a symmetrical object such as a rectangular prism (fourgroup), a regular triangle, tetrahedron, a cube or a sphere, to mention only a few important cases. Expressing the symmetry of an object, they are called symmetry groups. Multiplication of two elements means that the corresponding operations are carried out in a definite sequence. Except for the first case, these groups are non-Abelian.

The concrete definitions given above specify the multiplication rule for each group. For finite groups the results are conveniently represented in multiplication tables, from which one extracts the entire group structure. One recognizes for instance that some of the groups of covering operations listed under (4) are subgroups of others.

It is easy to prove the rearrangement theorem: In the multiplication table each column or row contains each element once and only once. This theorem is very helpful in setting up multiplication tables. (Helps to spot errors!)

2.1.2 Cyclic Groups

For an arbitrary element A of a finite $\mathcal{G}$ form the sequence: $A, A^2, A^3 \dots$ let the numbers of distinct elements in the sequence be p. It is easy to show that $A^p = E$. The sequence

$$A, A^2, \dots, A^p = E \tag{2.1.1}$$

is called the period of $A; p$ is the order of A. The period is an Abelian group, a subgroup of $\mathcal{G}$. It may be identical to it, in which case $\mathcal{G}$ is called a cyclic group.

Corollary: Since periods are subgroups, the order of each element is a divisor of the order of the group.

2.1.3 Cosets

Let $\mathcal{H}$ be a subgroup of $\mathcal{G}$ with elements $E, H_2, \dots H_h$; the set of elements

$$EA, H_2A, \dots, H_hA \tag{2.1.2}$$

is called a right coset $\mathcal{H}_A$ provided A is not in $\mathcal{H}$. It is easily shown that $\mathcal{G}$ can be decomposed as

$$\mathcal{G} = \mathcal{H}_E + \mathcal{H}_{A2} + \mathcal{H}_{Ah} \tag{2.1.3}$$

into distinct cosets, each of which contains h elements. Hence the order g of the group is

$$g = hk \quad \text{and} \quad h = g/k \tag{2.1.4}$$

Thus we got the important result that the order of a subgroup is a divisor of the order of the group. Note that the cosets are not subgroups except for $\mathcal{H}_E = \mathcal{H}$ which alone contains the unit element. Similar results hold for left cosets.

2.1.4 Conjugate Elements and Classes

The element XAX^{-1} 1 is said to be an element conjugate to A. The relation of being conjugate is reflexive, symmetric and transitive. Therefore the elements conjugate to each other form a class.

A single element A determines the entire class:

$$EAE^{-1} = A, A_2AA_2^{-1}, \dots, A_nAA_n^{-1} \tag{2.1.5}$$

Here all elements occur at least once, possibly more than once. The elements of the group can be divided into classes, and every element appears in one and only one class.

In the case of groups of covering operations of symmetrical objects, elements of the same class correspond to rotations by the same angle around different axes that transform into each other by symmetry operations.

E.g. the three mirror planes of the regular triangle are in the same class and so are the four rotations by $2\pi/3$ in a tetrahedron, or the eight rotations by $\pm 2\pi/3$ in a cube.

It happens that the elements of two groups defined in different conceptual terms are in one-one relation to each other and obey the same multiplication rules. A case in point is the permutation group $\mathcal{S}(3)$ and the symmetry group of the regular triangle. Such groups are called isomorphic. Recognizing isomorphisms may lead to new insights and to practical economies in the study of individual groups.

It is confirmed in the above examples that the term “multiplication” is not to be taken in a literal sense. What is usually meant is the performance of operations in a specified sequence, a situation that arises in many practical and theoretical contexts.

The operations in question are often transformations in ordinary space, or in some abstract space (say, the configuration space of an object of interest). In order to describe these transformations in a quantitative fashion, it is important to develop an algebraic formalism dealing with vector spaces.

However, before turning to the algebraic developments in Section 2.3, we consider first a purely geometric discussion of the rotation group in ordinary three-dimensional space.

2.2: The geometry of the three-dimensional rotation group. The Rodrigues-Hamilton theorem

There are three types of transformations that map the Euclidean space onto itself: translations, rotations and inversions. The standard notation for the proper rotation group is $\mathcal{O}^+$, or $\mathcal{SO}(3)$ short for “simple orthogonal group in three dimensions”. “Simple” means that the determinant of the transformation is $+1$, we have proper rotations with the exclusion of the inversion of the coordinates:

$$\begin{aligned} x &\rightarrow -x \\ y &\rightarrow -y \\ z &\rightarrow -z \end{aligned} \tag{2.2.1}$$

In contrast to the group of translations, $\mathcal{SO}(3)$ is non-Abelian, and its theory, beginning with the adequate choice of parameters is quite complicated. Nevertheless, its theory was developed to a remarkable degree during the 18th century by Euler.

Within classical mechanics the problem of rotation is not considered to be of fundamental importance. The Hamiltonian formalism is expressed usually in terms of point masses, which do not rotate. There is a built-in bias in favor of translational motion.

The situation is different in quantum mechanics where rotation plays a paramount role. We have good reasons to give early attention to the rotation group, although at this point we have to confine ourselves to a purely geometrical discussion that will be put later into an algebraic form.

According to a well known theorem of Euler, an arbitrary displacement of a rigid body with a single fixed point can be conceived as a rotation around a fixed axis which can be specified in terms of the angle of rotation ϕ, and the unit vector $\hat{u}$ along the direction of the rotational axis. Conventionally the sense of rotation is determined by the right hand rule. Symbolically we may write $R = \{\hat{u}, \phi\}$.

The first step toward describing the group structure is to provide a rule for the composition of rotations with due regard for the noncommuting character of this operation. The gist of the argument is contained in an old theorem by Rodrigues-Hamilton.

Our presentation follows that of C. L. K. Whitney [Whi68]. Consider the products

$$R_3 = R_2 R_1 \tag{2.2.2}$$

$$R_3' = R_1 R_2 \tag{2.2.3}$$

where R_3 is the composite rotation in which R_1 is followed by R_2.

Figure 2.1 represents the unit sphere and is constructed as follows: the endpoints of the vectors $\hat{u}_1$ and $\hat{u}_2$ determine a great circle, the smaller arc of which forms the base of mirror-image triangles having angles $\phi_1/2$ and $\phi_2/2$ as indicated. The endpoint of the vector $\hat{u}_1'$ is located by rotating $\hat{u}_1$ by angle ϕ_2 about $\hat{u}_2$. Our claim, that the other quantities appearing on the figure are legitimately labeled $\phi_3/2, \hat{u}_3, \hat{u}_3'$ is substantiated easily. Following the sequence of operations indicated in 2.2.3, we see that the vector called $\hat{u}_3$ is first rotated by angle ϕ_1, about $\hat{u}_1$, which takes in into $\hat{u}_3'$. Then it is rotated by angle ϕ_2 about $\hat{u}_2$, which takes it back to $\hat{u}_3$. Since it is invariant, it is indeed the axis of the combined rotation. Furthermore, we see that the first rotation leaves $\hat{u}_1$ invariant and the second rotation, that about $\hat{u}_2$ carries it into $\hat{u}_1'$, the position it would reach if simply rotated about $\hat{u}_3$, by the angle called ϕ_3. Thus that angle is indeed the angle of the combined rotation. Note that a symmetrical argument shows that $\hat{u}_3'$ and ϕ_3 are the axis and angle of the rotation $P_3' = R_1 R_2$.

Equation 2.2.3 can be expressed also as

$$R_3^{-1} R_2 R_1 = 1 \tag{2.2.4}$$

which is interpreted as follows: rotation about $\hat{u}_1$, by ϕ_1, followed by rotation about $\hat{u}_2$, by ϕ_2 followed by rotation about $\hat{u}_3$, by minus ϕ_3 produces no change. This statement is the Rodrigues-Hamilton theorem.

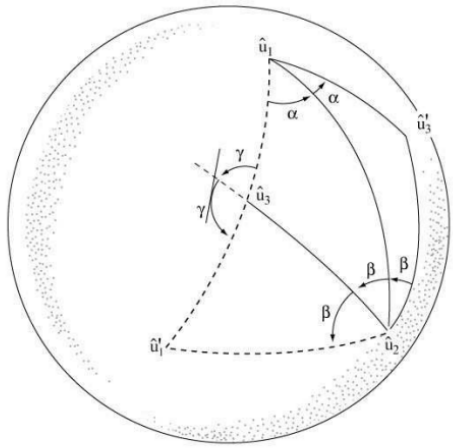

Figure 2.1: Composition of the Rotations of the Sphere $\alpha = \phi_1/2, \beta = \phi_2/2, \gamma = \phi_3/2$.

2.3: The n-dimensional vector space V(n)

The manipulation of directed quantities, such as velocities, accelerations, forces and the like is of considerable importance in classical mechanics and electrodynamics. The need to simplify the rather complex operations led to the development of an abstraction: the concept of a vector.

The precise meaning of this concept is implicit in the rules governing its manipulations. These rules fall into three main categories: they pertain to

1. the addition of vectors,

2. the multiplication of vectors by numbers (scalars),

3. the multiplication of vectors by vectors (inner product and vector product.

While the subtle problems involved in 3 will be taken up in the next chapter, we proceed here to show that rules falling under 1 and 2 find their precise expression in the abstract theory of finite dimensional vector spaces.

The rules related to the addition of vectors can be concisely expressed as follows: vectors are elements of a set V that forms an Abelian group under the operation of addition, briefly an additive group.

The inverse of a vector is its negative, the zero vector plays the role of unity.

The numbers, or "scalars" mentioned under (ii) are usually taken to be the real or the complex numbers. For many considerations involving vector spaces there is no need to specify which of these alternatives is chosen. In fact all we need is that the scalars form a field. More explicitly, they are elements of a set which is closed with respect to two binary operations: addition and multiplication which satisfy the common commutative, associative and distributive laws; the operations are all invertible provided they do not involve division by zero.

A vector space V (F) over a field F is formally defined as a set of elements forming an additive group that can be multiplied by the elements of the field F.

In particular, we shall consider real and complex vector fields V (R) and V (C) respectively.

I note in passing that the use of the field concept opens the way for a much greater variety of interpretations, but this is of no interest in the present context. In contrast, the fact that we have been considering "vector" as an undefined concept will enable us to propose in the sequel interpretations that go beyond the classical one as directed quantities. Thus the above defintion is consistent with the interpretation of a vector as a pair of numbers indicating the amounts of two chemical species present in a mixture, or alternatively, as a point in phase space spanned by the coordinates and momenta of a system of mass points.

We shall now summarize a number of standard results of the theory of vector spaces.

Suppose we have a set of non-zero vectors $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ in V which satisfy the relation

$$\sum_k a_k \vec{x}_k = 0 \tag{2.3.1}$$

where the scalars $a_k \in F$, and not all of them vanish. In this case the vectors are said to be linearly dependent. If, in contrast, the relation 2.3.1 implies that all $a_k = 0$, then we say that the vectors are linearly independent.

In the former, case there is at least one vector of the. set that.can be written as a linear combination of the rest:

$$\vec{x}_m = \sum_1^{m-1} b_k \vec{x}_k \tag{2.3.2}$$

Definition 2.1. A (linear) basis in a vector space V is a set $E = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ of linearly independent vectors such that every vector in V is a linear combination of the $\vec{e}_n$. The basis is said to span or generate the space.

A vector space is finite dimensional if it has a finite basis. It is a fundamental theorem of linear algebra that the number of elements in any basis in a finite dimensional space is the same as in any other basis. This number n is the basis independent dimension of V ; we include it into the designation of the vector space: $V(n, F)$

Given a particular basis we can express any $\vec{x} \in V$ as a linear combination

$$\vec{x}=\sum_{1}^{n} x^k \vec{e}_k \tag{2.3.3}$$

where the coordinates x^k are uniquely determined by E. The $x^k \vec{e}_k (k=l,2,\ldots,n)$ are called the components of $\vec{x}$. The use of superscripts is to suggest a contrast between the transformation properties of coordinates and basis to be derived shortly.

Using bases, called also coordinate systems, or frames is convenient for handling vectors — thus addition performed by adding coordinates. However, the choice of a particular basis introduces an element of arbitrariness into the formalism and this calls for countermeasures.

Suppose we introduce a new basis by means of a nonsingular linear transformation:

$$\vec{e}\,'_i=\sum_k S_i^k \vec{e}_k \tag{2.3.4}$$

where the matrix of the transformation has a nonvanishing determinant

$$|S_i^k|\neq 0 \tag{2.3.5}$$

ensuring that the $\vec{e}\,'_i$ form a linearly independent set, i.e., an acceptable basis. Within the context of the linear theory this is the most general transformation we have to consider.

We ensure the equivalence of the different bases by requiring that

$$\vec{x}=\sum x^k \vec{e}_k=\sum x^{i\prime} \vec{e}\,'_i \tag{2.3.6}$$

Inserting Equation 2.3.4 into Equation 2.3.6 we get

$$\begin{aligned}\vec{x}&=\sum x^{i'}\left(\sum S_i^k \vec{e}_k\right)\\&=\sum\left(\sum x^{i\prime} S_i^k\right)\vec{e}_k\end{aligned}$$

and hence in conjunction with Equation 2.3.5

$$x^k=\sum S_i^k x^{i'} \tag{2.3.7}$$

Note the characteristic “turning around” of the indices as we pass from Equation 2.3.4 to Equation 2.3.7 with a simultaneous interchange of the roles of the old and the new frame. The underlying reason can be better appreciated if the foregoing calculation is carried out in symbolic form.

Let us write the coordinates and the basis vectors as $n\times 1$ column matrices

$$X=\begin{pmatrix}x^1\\ \vdots\\ x^k\end{pmatrix}\quad E=\begin{pmatrix}\vec{e}_1\\ \vdots\\ \vec{e}_k\end{pmatrix} \tag{2.3.8}$$

Equation 2.3.6 appears then as a matrix product

$$\vec{x}=X^T E=X^T S^{-1} S E=X'^T E' \tag{2.3.9}$$

where the superscript stands for “transpose.”

We ensure consistency by setting

$$E'=SE \tag{2.3.10}$$

$$X'^T=X^T S^{-1} \tag{2.3.11}$$

$$X'=S^{-1T} X \tag{2.3.12}$$

Thus we arrive in a lucid fashion at the results contained in Equations 2.3.4 and 2.3.7. We see that the “objective” or “invariant” representations of vectors are based on the procedure of transforming bases and coordinates in what is called a contragredient way.

The vector $\vec{x}$ itself is sometimes called a contravariant vector, to be distinguished by its transformation properties from covariant vectors to be introduced later.

There is a further point to be noted in connection with the factorization of a vector into basis and coordinates.

The vectors we will be dealing with have usually a dimention such as length, velocity, momentum, force and the like. It is important, in such cases, that the dimension be absorbed in the basis vectors $\vec{e}_k$. In contrast, the coordinates x^k are elements of the field F, the products of which are still in F, they are simply numbers. It is not surprising that the multiplication of vectors with other vectors constitutes a subtle problem. Vector spaces in which there is provision for such an operation are called algebras; they deserve a careful examination.

It should be finally pointed out that there are interesting cases in which vectors have a dimensionless character. They can be built up from the elements of the field F, which are arranged as n-tuples, or as $m \times n$ matrices.

The $n \times n$ case is particularly interesting, because matrix multiplication makes these vector spaces into algebra in the sense just defined.

2.4: How to multiply vectors? Heuristic considerations

In evaluating the various methods of multiplying vectors with vectors, we start with a critical analysis of the procedure of elementary vector calculus based on the joint use of the inner or scalar product and the vector product.

The first of these is readily generalized to $V(n, R)$, and we refer to the literature for further detail. In contrast, the vector product is tied to three dimensions, and in order to generalize it, we have to recognize that it is commonly used in two contexts, to perform entirely different functions.

First to act as a rotation operator, to provide the increment $\delta\vec{a}$ of a vector $\vec{a}$ owing to a rotation by an angle $\delta\theta$ around an axis $\hat{n}$:

$$\delta\vec{a} = \delta\theta\hat{n} \times \vec{a} \tag{2.4.1}$$

Here $\delta\theta\hat{n}$ is a dimensionless operator that transforms a vector into another vector in the same space.

Second, to provide an “area”, the dimension of which is the product of the dimension of the factors. In addition to the geometrical case, we have also such constructs as the angular momentum

$$\vec{L} = \vec{r} \times \vec{p} \tag{2.4.2}$$

The product is here “exterior” to tha original vector space.

There is an interesting story behind this double role of the vector product. Gibbs’ vector algebra arose out of the attempt of reconciling and simplifying two ingenious, but complicated geometric algebras which were advanced almost simultaneously in the 1840’s. Sir William Rowan Hamilton’s theory of quaternions is adapted to problems of rotation in three- and four-dimensional spaces, whereas Hermann Grassman’s Ausdehnungslehre (Theory of Extensions) deals with volumes in spaces of an arbitrary number of dimensions. The dichotomy corresponds to that of Equations 2.4.1 and 2.4.2.

The complementary character of the two calculi was not recognized at the time, and the adherents of the two methods were in fierce competition. Gibbs found his way out of the difficulty by removing all complicated and controversial elements from both calculi and by reducing them to their common core. The result is our well known elementary vector calculus with its dual-purpose vector product which seemed adequate for three-dimensional/space.

Ironically, the Gibbsian calculus became widely accepted at a time when the merit of Hamilton’s four-dimensional rotations was being vindicated in the context of the Einstein-Minkowski fourdimensional world.

Although it is possible to adapt quaternions to deal with the Lorentz group, it is more practical to use instead the algebra of complex two-by-two matrices, the so-called Pauli algebra, and the complex vectors (spinors) on which these matrices operate. These methods are descendents of quaternion algebra, but they are more general, and more in line with quantum mechanical techniques. We shall turn to their development in the next Chapter.

In recent years, also some of Grassmann’s ideas have been revived and the exterior calculus is now a standard technique of differential geometry (differential forms, calculus of manifolds). These matters are relevant to the geometry of phase space, and we shall discuss them later on.

2.5: A Short Survey of Linear Groups

The linear vector space $V(n, F)$ provides us with the opportunity to define a number of linear groups which we shall use in the sequel.

We start with the group of nonsingular linear transformations defined by Equations 2.3.4 and 2.3.5 of Section 2.3 and designated as $\mathcal{GL}(n, R)$, for "general linear group over the field F." If the matrices are required to have unit determinants, they are called unimodular, and the group is $\mathcal{SL}(n, F)$ for simple linear group.

Let us consider now the group $\mathcal{GL}(n, R)$ over the real field, and assume that an inner product is defined:

$$x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = X^T Y \tag{2.5.1}$$

Transformations which leave this form invariant are called orthogonal. By using Equations 2.3.10 and 2.3.12 of Sectionsec:vec-space, we see that they satisfy the condition

$$O^T O = \mathcal{I} \tag{2.5.2}$$

where $\mathcal{I}$ is the unit matrix. The corresponding group is called $\mathcal{O}(n)$.

It follows from 2.5.2 that the determinant of O is $\det O = |O| = \pm 1$. The matrices with positive determinant form a subgroup $\mathcal{SO}(n)$.

The orthogonal groups have an important geometrical meaning, they leave the so-called metric properties, lengths and angles invariant. The group $SO(n)$ corresponds to pure rotations, these operations can be continuously connected with the identity. In contrast, transformations with negative determinants involve the inversion, and hence mirrorings and improper rotations. The set of matrices with $|O| = -1$, does not form a group, since it does not contain the unit element.

The geometrical interpretation of $\mathcal{GL}(n, R)$ is not explained as easily. Instead of metric Euclidean geometry, we arrive at the less familiar affine geometry, the practical applications of which are not so direct. We shall return to these questions in Chapter VII. However, in the next section we shall show that the geometrical interpretation of the group of unimodular transformations $\mathcal{SL}(n, R)$ is to leave volume invariant.

We turn now to an extension of the concept of metric geometry. We note first that instead of requiring the invariance of the expression 2.5.1, we could have selected an arbitrary positive definite quadratic form in order to establish a metric. However, a proper choice of basis in $\mathcal{V}(n, R)$ leads us back to Equation 2.5.1.

If the invariant quadratic form is indefinite, it reduces to the canonical form

$$x_1^2 + x_2^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_{k+l}^2 \tag{2.5.3}$$

The corresponding group of invariance is pseudo-orthogonal denoted as $O(k, l)$.

In this category the Lorentz group $SO(3, 1)$ is of fundamental physical interest. At this point we accept this as a fact, and a sufficient incentive for us to examine the mathematical structure of $SO(3, 1)$ in Section 3. However, subsequently, in Section 4, we shall review the physical principles which are responsible for the prominent role of this group. The nature of the mathematical study can be succinctly explained as follows.

The general $n \times n$ matrix over the real field contains n^2 independent parameters. The condition 2.5.2 cuts down this number to $n(n-l)/2$. For $n = 3$ the number of parameters is cut down from nine to three, for $n = 4$ from sixteen to six. The parameter count is the same for $SO(3, 1)$ as for $SO(4)$. One of the practical problems involved in the applications of these groups is to avoid dealing with the redundant variables, and to choose such independent parameters that can be easily identified with geometrically and physically relevant quantities. This is the problem discussed in Section 3. We note that $SO(3)$ is a subgroup of the Lorentz group, and the two groups are best handled within the same framework.

It will turn out that the proper parametrization can be best attained in terms of auxiliary vector spaces defined over the complex field. Therefore we conclude our list of groups by adding the unitary groups.

Let us consider the group $\mathcal{GL}(n, C)$ and impose an invariant Hermitian form

$$\sum a_{ik} x_i x_k^* \tag{2.5.4}$$

that can be brought to the canonical form

$$x_1x_1^* + x_2x_2^* + \ldots + x_nx_n^* = X^\dagger X \tag{2.5.5}$$

where $X^\dagger = X^{*T}$ is the Hermitian adjoint of X and the star stands for the conjugate complex. Expression 2.5.5 is invariant under transformations by matrices that satisfy the condition

$$U^\dagger U = \mathcal{I} \tag{2.5.6}$$

These matrices are called unitary, they form the unitary group $\mathcal{U}(n)$. Their determinants have the absolute value one. If the determinant is equal to one, the unitary matrices are also, unimodular, we have the simple unitary group $\mathcal{SU}(n)$.

2.6: The unimodular group SL(n, R) and the invariance of volume

It is well known that the volume of a parallelepiped spanned by linearly independent vectors is given by the determinant of the vector components. It is evident therefore that a transformation with a unimodular matrix leaves this expression for the volume invariant.

Yet the situation has some subtle aspects which call for a closer examination. Although the calculation of volume and area is among the standard procedures of geometry, this is usually carried out in metric spaces, in which length and angle have their well known Euclidean meaning. However, this is a too restrictive assumption, and the determinantal formula can be justified also within affine geometry without using metric concepts.

Since we shall repeatedly encounter such situations, we briefly explain the underlying idea for the case of areas in a two-dimensional vector space $V(2, R)$.

We advance two postulates:

1. Area is an additive quantity: the area of a figure is equal to the sum of the areas of its parts.

2. Translationally congruent figures have equal areas.

(The point is that Euclidean congruence’involves also rotational congruence, which is not available to us because of the absence of metric.) We proceed now in successive steps as shown in Figure 2.2.

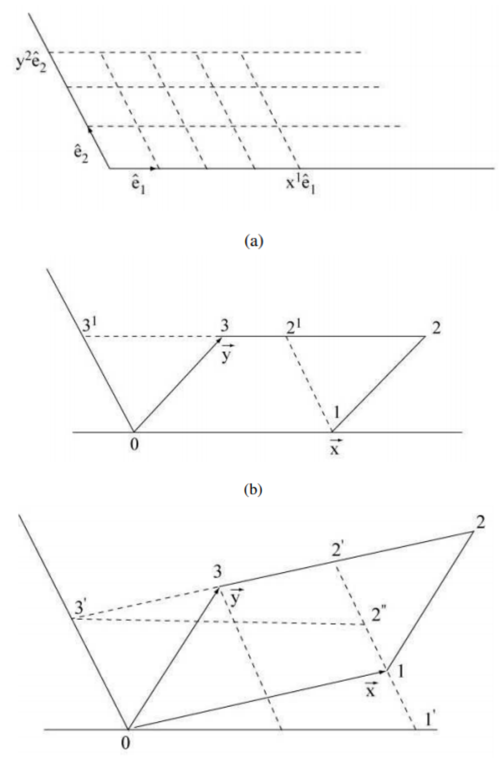

Figure 2.2: Translational congruence and equal area.

Consider at first the vectors

$$\vec{x} = x^1\vec{e}_1 \\ \vec{y} = y^2\vec{e}_2 \tag{2.6.1}$$

where the coordinates are integers (Figure 2.2a). The area relative to the unit cell is obtained through simple counting as x^1y^2. The same result can be justified for any real values for the coordinates by subdivision and a limiting process.

We are permitted to write this result in determinantal form:

$$x^1y^2 = \begin{vmatrix} x^1 & 0 \\ 0 & y^2 \end{vmatrix} \tag{2.6.2}$$

If the vectors

$$\begin{aligned} \vec{x} &= x^1\vec{e}_1 + x^2\vec{e}_2 \\ \vec{y} &= y^1\vec{e}_1 + y^2\vec{e}_2 \end{aligned} \tag{2.6.3}$$

do not coincide with the coordinate axes, the coincidence can be achieved in no more than two steps (Figures 2.2b and 2.2c) using the translational congruence of the parallelograms (0123) (012'3') (012"3').

By an elementary geometrical argument one concludes from here that the area spanned by $\vec{x}$ and $\vec{y}$ is equal to the area spanned by $\hat{e}_1$ and $\hat{e}_2$ multiplied by the determinant

$$\begin{vmatrix} x^1 & x^2 \\ y^1 & y^2 \end{vmatrix} \tag{2.6.4}$$

This result can be justified also in a more elegant way: The geometrical operations in figures b and c consist of adding the multiple of the vector $\vec{y}$ to the vector $\vec{x}$, or adding the multiple of the second row of the determinant to the first row, and we know that such operations leave the value of the determinant unchanged.

The connection between determinant and area can be generalized to three and more dimensions, although the direct geometric argument would become increasingly cumbersome.

This defect will be remedied most effectively in terms of the Grassmann algebra that will be developed in Chapter VII.

2.7: On “alias” and “alibi”. The Object Group

It is fitting to conclude this review of algebraic preliminaries by formulating a rule that is to guide us in connecting the group theoretical concepts with physical principles, One of the concerns of physicists is to observe, identify and classify particles. Pursuing this objective we should be able to tell whether we observe the same object when encountered under different conditions in different states. Thus the identity of an object is implicitly given by the set of states in which we recognize it to be the same. It is plausible to consider the transformations which connect these states with each other, and to assume that they form a group. Accordingly, a precise way of identifying an object is to specify an associated object group.

The concept of object group is extremely general, as it should be, in view of the vast range of situations it is meant to cover. It is useful to consider specific situations in more detail.

First, the same object may be observed by different inertial observers whose findings are connected by the transformations of the inertial group, to be called also the passive kinematic group. Second, the space-time evolution of the object in a fixed frame of reference can be seen as generated by an active kinematic group. Finally, if the object is specified in phase space, we speak of the dynamic group.

The fact that linear transformations in a vector space can be given a passive and an active interpretation, is well known. In the mathematical literature these are sometimes designated by the colorful terms “alias” and “alibi,” respectively. The first means that the transformation of the basis leads to new “names” for the same geometrical, or physical objects. The second is a mapping by which the object is transformed to another ”location” with respect to the same frame.

The important groups of invariance are to be classified as passive groups. Without in any way minimizing their importance, we shall give much attention also to the active groups. This will enable us to handle, within a unified group-theoretical framework, situations commonly described in terms of equations of motion, and also the so-called “preparations of systems” so important in quantum mechanics.

It is the systematic joint use of “alibi” and “alias” that characterizes the following argument.

CHAPTER OVERVIEW

3: The Lorentz Group and the Pauli Algebra

3.1: Introduction

Twentieth century physics is dominated by the development of relativity and quantum mechanics, disciplines centered around the universal constants c and h respectively. Historically, the emergence of these constants revealed a socalled breakdown of classical concepts.

From the point of view of our present knowledge, it would be evidently desirable to avoid such breakdowns and formulate only principles which are correct according to our present knowledge. Unfortunately, no one succeeded thus far to suggest a ”correct” postulational basis which would be complete enough for the wide ranging practical needs of physics.

The purpose of this course is to explore a program in which we forego, or rather postpone, the requirement of completeness, and consider at first only simple situations. These are described in terms of concepts which form the basis for the development of a precise mathematical formalism with empirically verified physical implications. The continued alternation of conceptual analysis with formal developments gradually extends and deepens the range of situations covered, without affecting consistency and empirical validity.

According to the central idea of quantum mechanics all particles have undulatory properties, and electromagnetic radiation has corpuscular aspects. In the quantitative development of this idea we have to make a choice, whether to start with the classical wave concept and build in the corpscular aspects, or else start with the classical concept of the point particle, endowed with a constant and invariant mass, and modify these properties by means of the wave concept. Eventually, the resulting theory should be independent of the path chosen, but the details of the construction process are different.

The first alternative is apparent in Einstein’s photon hypothesis, which is closely related with his special theory of relativity.

In contrast, the wealth of nonrelativistic problems within atomic, molecular and nuclear physics favored the second approach which is exploited in the Bohr-Heisenberg quantum mechanics.

The course of the present developments is set by the decision of following up the Einsteinian departure.

3.2: The Corpuscular Aspects of Light

As a first step in carrying out the problem just stated, we start with a precise, even though schematic, formulation of wave kinematics.

We consider first a spherical wave front

$$r_o^2 - \vec{r}^2 = 0 \tag{3.2.1}$$

where

$$r_o = ct \tag{3.2.2}$$

and t is the time elapsed since emission to the instantaneous wave front.

In order to describe propagation in a definite direction, say along the unit vector $\hat{k}$ we introduce an appropriately chosen tangent plane corresponding to a monochromatic plane wave.

$$k_o r_o - \vec{k} \cdot \vec{r} = 0 \tag{3.2.3}$$

with

$$k_o = \omega / c \tag{3.2.4}$$

$$\vec{k} = \frac{2\pi}{\lambda} \hat{k} \tag{3.2.5}$$

and

$$k_o^2 - \vec{k}^2 = 0 \tag{3.2.6}$$

where the symbols have their conventional meanings.

Next we postulate that radiation has a granular character, as it is expressed already in Definition 1 of Newton's Optics! However, in a more quantitative sense we state the standard quantum condition according to which a quantum of light pulse with the wave vector $\left(k_o, \vec{k}\right)$ is associated with a four-momentum.

$$(p_o, \vec{p}) = \hbar \left(k_o, \vec{k}\right) \tag{3.2.7}$$

$$p_o^2 - \vec{p}^2 = 0 \tag{3.2.8}$$

with

$$p_o = \frac{E}{c} \tag{3.2.9}$$

where E is the energy of the light quantum, or photon.

The proper coordination of the two descriptions involving spherical and plane waves, presents problems to which we shall return later. At this point it is sufficient to note that individual photons have directional properties described by a wave vector,and a spherical wave can be considered as an assembly of photons emitted isotropically from a small source.

As the next step in our procedure we argue that the photon as a particle should be associated with an object group, as introduced in Sec. 1.7. Assuming with Einstein that light velocity is unaffected by an inertial transformation, the passive kinematic group that leaves Eq's (3.2.1) - (3.2.3) invariant is the Lorentz group.

There are few if any principles in physics which are as thoroughly justified by their implications as the principle of Lorentz invariance. Our objective is to develop these implications in a systematic fashion.

In the early days of relativity the consequences of Lorentz invariance involved mostly effects of the order of $(v/c)^2$, a quantity that is small for the velocities attainable at that time. The justification is much more dramatic at present when we can refer to the operation of high energy accelerators operating near light velocity.

Yet this is not all. Lorentz invariance has many consequences which are valid even in nonrelativistic physics, but classically they require a number of independent postulates for their justification. In such cases the Lorentz group is used to achieve conceptual economy.

In-view of this far-reaching a posteriori verification of the constancy of light velocity, we need not be unduly concerned with its a priori justification. It is nevertheless a puzzling question, and it has given rise to much speculation: what was Einstein's motivation in advancing this postulate?

Einstein himself gives the following account of a paradox on which he hit at the age of sixteen:

"If I pursue a beam of light with the velocity c (velocity of light in vacuum), I should observe such a beam of light' as a spatially oscillatory electromagnetic field at rest. However, there seems to be no such thing, whether on the basis of experience or according to Maxwell's equations."

The statement could be actually even sharpened: on overtaking a travelling wave, the resulting phenomenon would simply come to rest, rather than turn into a standing wave.

However it may be, if the velocity of propagation were at all affected by the motion of the observer, it could be "transformed away." Should we accept such a radical change from an inertial transformation? At least in hindsight, we know that the answer is indeed no.

Note that the essential point in the above argument is that a light quantum cannot be transformed to rest. This absence of a preferred rest system with respect to the photon does not exclude the existence of a preferred frame defined from other considerations. Thus it has been recently suggested that a preferred frame be defined by the requirement that the $3K$ radiation be isotropic in it [Wei72].

Since Einstein and his contemporaries emphasized the absence of any preferred frame of reference, one might have wondered whether the aforementioned radiation, or some other cosmologically defined frame,might cause difficulties in the theory of relativity.

Our formulation, based on weaker assumptions, shows that such concern is unwarranted.

We observe finally, that we have considered thus far primarily wave kinematics, with no reference to the electrodynamic interpretation of light. This is only a tactical move. We propose to derive classical electrodynamics (CED) within our scheme, rather than suppose its validity.

Problems of angular momentum and polarization are also left for later inclusion.

However, we are ready to widen our context by being more explicit with respect to the properties of the four-momentum.

Eq (3.2.5) provides us with a definition of the four-momentum, but only for the case of the photon, that is for a particle with zero rest mass and the velocity c.

This relation is easily generalized to massive particles that can be brought to rest. We make use of the fact that the Lorentz transformation leaves the left-hand side of Eq (3.2.5) invariant, whether or not it vanishes. Therefore we set

$$p_o^2 - \vec{p} \cdot \vec{p} = m^2 c^2 \tag{3.2.10}$$

and define the mass m of a particle as the invariant "length" of the four-momentum according to the Minkowski metric (with $c = 1$).

We can now formulate the postulate: The four-momentum is conserved. This principle includes the conservation of energy and that of the three momentum components. It is to be applied to the interaction between the photon and a massive particle and also to collision processes in general.

In order to make use of the conservation law, we need explicit expressions for the velocity dependence of the four-momentum components. These shall be obtained from the study of the Lorentz group.

3.3: On Circular and Hyperbolic Rotations

We propose to develop a unified formalism for dealing with the Lorentz group $\mathcal{SO}(3,1)$ and its subgroup $\mathcal{SO}(3)$. This program can be divided into two stages. First, consider a Lorentz transformation as a hyperbolic rotation, and exploit the analogies between circular and hyperbolic trigonometric functions, and also of the corresponding exponentials. This simple idea is developed in this section in terms of the subgroups $\mathcal{SO}(2)$ and $\mathcal{SO}(1,1)$. The rest of this chapter is devoted to the generalization of these results to three spatial dimensions in terms of a matrix formalism.

Let us consider a two-component vector in the Euclidean plane:

$$\vec{x}=x_1\hat{e}_1+x_2\hat{e}_2 \tag{3.3.1}$$

We are interested in the transformations that leave $x_1^2+x_2^2$ invariant. Let us write

$$x_1^2+x_2^2=(x_1+ix_2)(x_1-ix_2) \tag{3.3.2}$$

and set

$$(x_1+ix_2)'=a(x_1+ix_2) \tag{3.3.3}$$

$$(x_1-ix_2)'=a^*(x_1-ix_2) \tag{3.3.4}$$

where the star means conjugate complex. For invariance we have

$$aa^*=1 \tag{3.3.5}$$

or

$$a=\exp(-i\phi),\quad a^*=\exp(i\phi) \tag{3.3.6}$$

From these formulas we easily recover the elementary trigonometric expressions. Table 3.1 summarizes the presentations of rotational transformations in terms of exponentials, trigonometric functions and algebraic irrationalities involving the slope of the axes. There is little to recommend the use of the latter, however it completes the parallel with Lorentz transformations where this parametrization is favored by tradition.

We emphasize the advantages of the exponential function, mainly because it lends itself to iteration, which is apparent from the well known formula of de Moivre:

$$\exp(\text{ in }\phi)=\cos(n\phi)+i\sin(n\phi)=(\cos(\phi)+i\sin(\phi))^n \tag{3.3.7}$$

The same Table contains also the parametrization of the Lorentz group in one spatial variable. The analogy between $SO(2)$ and $SO(1,1)$ is far reaching and the Table is selfexplanatory. Yet there are a number of additional points which are worth making.

The invariance of

$$x_0^2-x_3^2=(x_0+x_3)(x_0-x_3) \tag{3.3.8}$$

is ensured by

$$(x_0+x_3)'=a(x_0+x_3) \tag{3.3.9}$$

$$(x_0-x_3)'=a^{-1}(x_0-x_3) \tag{3.3.10}$$

for an arbitrary a. By setting $a=\exp(-\mu)$ in the Table we tacitly exclude negative values. Admitting a negative value for this parameter would imply the interchange of future and past. The Lorentz transformations which leave the direction of time invariant, are called orthochronic. Until further notice these are the only ones we shall consider.

The meaning of the parameter μ is apparent from the well known relation

$$\tanh\mu=\frac{v}{c}=\beta \tag{3.3.11}$$

where v is the velocity of the primed system Σ' measured in Σ. Being a (non-Euclidean) measure of a velocity, μ is sometimes called rapidity, or velocity parameter.

Rotation	Lorentz Transformation
$(x_1 - ix_2)' = e^{i\phi}(x_1 - ix_2)$ $(x_1 + ix_2)' = e^{-i\phi}(x_1 + ix_2)$	$(x_0 + x_3)' = e^{-\mu}(x_0 + x_3)$ $(x_0 - x_3)' = e^{\mu}(x_0 - x_3)$
$x_1' = x_1 \cos\phi + x_2 \sin\phi$ $x_2' = -x_1 \sin\phi + x_2 \cos\phi$	$x_3' = x_3 \cosh\mu - x_0 \sinh\mu$ $x_0' = -x_3 \sinh\mu + x_0 \cosh\mu$
$x_1' = \frac{x_1 + \kappa x_2}{\sqrt{1+\kappa^2}}$ $x_2' = \frac{\kappa x_1 + x_2}{\sqrt{1+\kappa^2}}$ $\kappa = \tan\phi = \left(\frac{x_2}{x_1}\right)_{x_2'=0}$ $= -\left(\frac{x_2'}{x_1'}\right)_{x_2=0}$	$x_3' = \frac{x_3 - \beta x_0}{\sqrt{1-\beta^2}}$ $x_0' = \frac{-\beta x_3 + x_0}{\sqrt{1-\beta^2}}$ $\beta = \tanh\mu = \left(\frac{x_3}{x_0}\right)_{x_3'=0}$ $= \left(\frac{x_3'}{x_0'}\right)_{x_3=0}, \quad x_0 = ct$
$\frac{X_2}{x_1} = \tan\theta$ $\theta' = \theta - \phi$	$\frac{X_3}{x_0} = \tanh\nu$ $\nu' = \nu - \mu$

Table 3.1: Summary of the rotational transformations. (The signs of the angles correspond to the passive interpretation.)

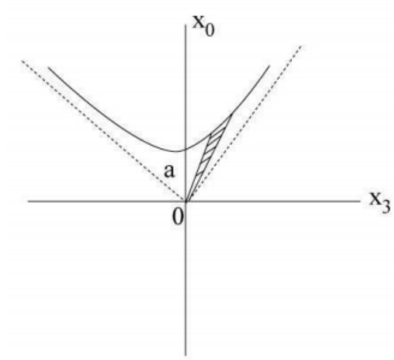

Figure 3.1: Area in (x_0, x_3) -plane.

We shall refer to μ also as hyperbolic angle. The formal analogy with the circular angle ϕ is evident from the Table. We deepen this parallel by means of the observation that μ can be interpreted as an area in the (x_0, x_3) plane (see Figure 3.1).

Consider a hyperbola with the equation

$$\left(\frac{x_0}{a}\right)^2 - \left(\frac{x_3}{b}\right)^2 = 1 \tag{3.3.12}$$

$$x_0 = a\cosh\mu \quad x_3 = b\sinh\mu \tag{3.3.13}$$

The shaded triangular area (shown in Figure 3.1) is according to Equation 2.6.2 of Section 2.6:

$$\frac{1}{2}\begin{vmatrix} x_3 + dx_3 & x_3 \\ x_0 + dx_0 & x_0 \end{vmatrix} = \frac{1}{2}(x_0 dx_3 - x_3 dx_0) = \tag{3.3.14}$$

$$\frac{ab}{2}\left(\cosh\mu^2 - \sinh\mu^2\right) d = \frac{ab}{2} d\mu \tag{3.3.15}$$

We could proceed similarly for the circular angle ϕ and define it in terms of the area of a circular sector, rather than an arc. However, only the area can be generalized for the hyperbola.

Although the formulas in Table 3.1 apply also to the wave vector and the four momentum.and can be used in each case also according to the active interpretation, the various situations have their individual features, some of which will now be surveyed.

Consider at first a plane wave the direction of propagation of which makes an angle θ with the direction x_3 of the Lorentz transformation. We write the phase, Equation 3.3.11 of Section 3.2, as

$$\frac{1}{2}\left[\left(k_0+k_3\right)\left(x_0-x_3\right)+\left(k_0-k_3\right)\left(x_0+x_3\right)\right]-k_1x_1-k_2x_2 \tag{3.3.16}$$

This expression is invariant if $\left(k_0 \pm k_3\right)$ transforms by the same factor $\exp(\pm\mu)$ as $\left(x_0 \pm x_3\right)$.

Thus we have

$$k_3' = k_3\cosh\mu - k_0\sinh\mu \tag{3.3.17}$$

$$k_0' = -k_3\sinh\mu + k_0\cosh\mu \tag{3.3.18}$$

Since $\left(k_0, \vec{k}\right)$ is a null-vector, i.e., k has vanishing length, we set

$$k_3 = k_0\cos\theta, \quad k_3' = k_0'\cos\theta' \tag{3.3.19}$$

and we obtain for the aberration and the Doppler effect:

$$\cos\theta' = \frac{\cos\theta\cosh\mu - \sinh\mu}{\cosh\mu - \cos\theta\sinh\mu} = \frac{\cos\theta - \beta}{1-\beta\cos\theta} \tag{3.3.20}$$

and

$$\frac{k_0'}{k_0} = \frac{\omega_0'}{\omega_0} = \cosh\mu - \cos\theta\sinh\mu \tag{3.3.21}$$

For $\cos\theta = 1$ we have

$$\frac{\omega_0'}{\omega_0} = \exp(-\mu) = \sqrt{\frac{1-\beta}{1+\beta}} \tag{3.3.22}$$

Thus the hyperbolic angle is directly connected with the frequency scaling in the Doppler effect.

Next, we turn to the transformation of the four-momentum of a massive particle. The new feature is that such a particle can be brought to rest. Let us say the particle is at rest in the frame Σ' (rest frame), that moves with the velocity $v_3 = c\tanh^{-1}\mu$ in the frame Σ (lab frame). Thus v_3 can be identified as the particle velocity along x_3.

Solving for the momentum in Σ':

$$p_3 = p_3'\cosh\mu + p_0'\sinh\mu \tag{3.3.23}$$

$$p_0 = p_3'\sinh\mu + p_0'\cosh\mu \tag{3.3.24}$$

with $p_3' = 0, p_0' = mc$, we have

$$p_3 = mc\sinh\mu = \frac{mc\beta}{\sqrt{1-\beta^2}} \tag{3.3.25}$$

$$p_0 = mc\cosh\mu = \frac{mc}{\sqrt{1-\beta^2}} = \frac{E}{c} \tag{3.3.26}$$

$$\gamma = \cosh\mu, \quad \gamma\beta = \sinh\mu \tag{3.3.27}$$

Thus we have solved the problem posed at the end of Section 3.2.

The point in the preceding argument is that we achieve the transition from a state of rest of a particle to a state of motion, by the kinematic means of inertial transformation. Evidently, the same effect can be achieved by means of acceleration due to a force, and consider this “boost” as an active Lorentz transformation. Let us assume that the particle carries the charge e and is exposed to a constant electric intensity E. We get from Equantion 3.3.25 for small velocities:

$$\frac{dp_3}{dt} = mc\cosh\mu\frac{d\mu}{dt} \simeq mc\frac{d\mu}{dt} \tag{3.3.28}$$

and this agrees with the classical equation of motion if

$$E=\frac{mc}{e}\frac{d\mu}{dt} \tag{3.3.29}$$

Thus the electric intensity is proportional to the hyperbolic angular velocity.

In close analogy, a circular motion can be produced by a magnetic field:

$$B=-\frac{mc}{e}\frac{d\phi}{dt}=-\frac{mc}{e}\omega \tag{3.3.30}$$

This is the well known cyclotron relation.

The foregoing results are noteworthy for a number of reasons. They suggest a close connection between electrodynamics and the Lorentz group and indicate how the group theoretical method provides us with results usually obtained by equations of motion.

All this brings us a step closer to our program of establishing much of physics within a group theoretical framework, starting in particular with the Lorentz group. However, in order to carry out this program we have to generalize our technique to three spatial dimensions. For this we have the choice between two methods.

The first is to represent a four-vector as a 4×1 column matrix and operate on it by 4×4 matrices involving 16 real parameters among which there are ten relations (see Section 2.5).

The second approach is to map four-vectors on Hermitian 2×2 matrices

$$P=\begin{pmatrix} p_0+p_3 & p_1-ip_2 \\ p_1+ip_2 & p_0-p_3 \end{pmatrix} \tag{3.3.31}$$

and represent Lorentz transformations as

$$P'=VPV^{\dagger} \tag{3.3.32}$$

where V and $V^{\dagger}$ are Hermitian adjoint unimodular matrices depending .just on the needed six parameters.

We choose the second alternative and we shall show that the mathematical parameters have the desired simple physical interpretations. In particular we shall arrive at generalizations of the de Moivre relation, Equation 3.3.7.

The balance of this chapter is devoted to the mathematical theory of the 2×2 matrices with physical applications to electrodynamics following in Section 4.

3.4: The Pauli Algebra

3.4.1 Introduction

Let us consider the set of all 2×2 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 $\mathcal{V}\left(n^2, C\right)$. 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 A_2 of all 2×2 complex matrices

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

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

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{3.4.2}$$

$$e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{3.4.3}$$

$$e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{3.4.4}$$

$$e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.4.5}$$

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

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.4.6}$$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.4.7}$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{3.4.8}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.4.9}$$

The three Pauli matrices satisfy the well known multiplication rules

$$\sigma_j^2 = 1 \quad j = 1, 2, 3 \tag{3.4.10}$$

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

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

$$\sigma_\mu^\dagger = \sigma_\mu \quad \mu = 0, 1, 2, 3 \tag{3.4.12}$$

(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 σ_μ denoted by a_μ. We shall refer to the numbers a_μ 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

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

where Tr stands for trace. In detail,

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

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

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

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

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_μ, 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,

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

$$= \begin{pmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3 \end{pmatrix} \tag{3.4.19}$$

We associate with .each matrix the half trace and the determinant

$$\frac{1}{2} \mathrm{Tr}\, A = a_0 \tag{3.4.20}$$

$$|A| = a_0^2 - \vec{a}^2 \tag{3.4.21}$$

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:

$$A \rightarrow |A|^{-1/2} A \tag{3.4.22}$$

The case of singular matrices

$$|A| = a_0^2 - \vec{a}^2 = 0 \tag{3.4.23}$$

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:

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

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

Evidently, for any two matrices A and B

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

The matrices A and B commute, if and only if

$$\vec{a} \times \vec{b} = 0 \tag{3.4.26}$$

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:

$$A \rightarrow A = a_0 I + \vec{a} \cdot \vec{\sigma} \tag{3.4.27}$$

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

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

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

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

$$(AB)^{\dagger} = B^{\dagger} A^{\dagger} \tag{3.4.31}$$

$$(\tilde{AB}) = \tilde{B}\tilde{A} \tag{3.4.32}$$

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

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 $\sim, t, -$, together with the identity operator, form a group (the four-group, “Vierergruppe”). 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.

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

For invertible matrices

$$A^{-1} = \frac{\tilde{A}}{|A|} \tag{3.4.35}$$

For unimodular matrices we have the useful rule:

$$A^{-1} = \tilde{A} \tag{3.4.36}$$

A Hermitian marrix $A = A^{\dagger}$ 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:

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

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

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

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

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

If U is also unimodular, then

$$u_0^* = u_0 = \text{ real} \tag{3.4.40}$$

$$\vec{u}^* = \vec{u} = \text{ imaginary} \tag{3.4.41}$$

and

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

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

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

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

with

$$|\xi_0|^2 + |\xi_1|^2 = 1 \tag{3.4.45}$$

where ξ_0, ξ_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

$$\vec{n} \times \vec{n}^* = 0 \tag{3.4.46}$$

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

$$N = n_0 1 + n \hat{n} \cdot \vec{\sigma} \tag{3.4.47}$$

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

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

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

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

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 $\hat{u}=\hat{n}$

$$N=HU=UH \tag{3.4.50}$$

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:

$$z=r\exp(-i\phi/2) \tag{3.4.51}$$

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

$$A=HU=UH' \tag{3.4.52}$$

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:

$$a_0a_0^*+\vec{a}\cdot\vec{a}^*>0 \tag{3.4.53}$$

$$|A|\left|A^\dagger\right|>0 \tag{3.4.54}$$

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

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

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

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

with

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

That U is indeed unitary is easily verified:

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

and these expressions are equal by Equation 3.4.55.

From Equation 3.4.56 we get

$$A=U\left(U^{-1}HU\right) \tag{3.4.58}$$

and

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

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

$$A=HU=H_1U_1 \tag{3.4.60}$$

then

$$AA^{\dagger}=H^2=H_1^2 \tag{3.4.61}$$

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

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

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 invertable.

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 $\mathcal{V}(\in,\mathcal{C})$. By using the familiar bra-ket formalism we write

$$A|\xi\rangle=|\xi'\rangle \tag{3.4.63}$$

$$A^{\dagger}\langle\xi|=\langle\xi'| \tag{3.4.64}$$

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 × 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:

$$A=a_0 1+\vec{a}\cdot\vec{\sigma} \tag{3.4.65}$$

$$A=\{a_0,\vec{a}\} \tag{3.4.66}$$

We have

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

or more generally

$$a_0b_0-\vec{a}\cdot\vec{b}=\frac{1}{2}\operatorname{Tr}(A\bar{B}) \tag{3.4.68}$$

A Lorentz transformation is defined as a linear transformation

$$\{a_0,\vec{a}\}=\mathcal{L}\left\{a_0',\vec{a}'\right\} \tag{3.4.69}$$

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 $4x4$ 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: $R=\{r,\vec{r}\}$

$$K=\left\{k_0,\vec{k}\right\}, \quad P=\{p_0,\vec{p}\} \tag{3.4.70}$$

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”

$$A' = VAW \tag{3.4.71}$$

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

$$A' = VAV^\dagger \tag{3.4.72}$$

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

$$\overline{A'} = \bar{V}\bar{A}\tilde{V} \tag{3.4.73}$$

and the product of two four-vectors:

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

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 ϕ differ by 2π)) 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

$$A = V^{-1}A'V^{\dagger-1} \equiv \tilde{V}A'\bar{V} \tag{3.4.74}$$

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

$$A' = UAU^\dagger \equiv UAU^{-1} \tag{3.4.75}$$

with

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

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

Let us decompose $\vec{a}$:

$$\vec{a} = \vec{a}_\parallel + \vec{a}_\perp \tag{3.4.76}$$

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

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 ϕ : $R\{\hat{u},\phi\}$.

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

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

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 3×3 matrices are reduced to the three independent ones of $U\left(\hat{u}, \frac{\phi}{2}\right)$. Moreover we have the simple result

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

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 $V = H$

$$A' = HAH \tag{3.4.80}$$

with

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

$$h_1^2 + h_2^2 + h_3^2 = 1, \quad -\infty < \mu < \infty \tag{3.4.82}$$

We decompose $\vec{a}$ as

$$\vec{a} = a\hat{h} + \vec{a}_\perp \tag{3.4.83}$$

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

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

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

Hence

$$a_0' = \cosh\mu a_0 + \sinh\mu a \tag{3.4.86}$$

$$a' = \sinh\mu a_0 + \cosh\mu a_0 \tag{3.4.87}$$

$$\vec{a}_\perp' = \vec{a}_\perp \tag{3.4.88}$$

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

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

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

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

According to Equation 3.4.59 of Section 3.4.2, H and $H^{\prime}$ are connected by a similarity transformation, which does not affect the angle μ 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.

$$\left[N, N^{\prime}\right]=2 i \sinh \frac{\kappa}{2} \sinh \frac{\kappa}{2}\left(\hat{n} \times \hat{n}^{\prime}\right) \cdot \vec{\sigma} \tag{3.4.91}$$

where $\kappa=\mu-i \phi, \kappa^{\prime}=\mu^{\prime}-i \phi^{\prime}$

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

$$U\left(\hat{u}_{k}, \frac{d \phi}{2}\right)=1-\frac{i}{2} d \phi \sigma_{k}=1+d \phi I_{k} \tag{3.4.92}$$

$$I_{k}=-\frac{i}{2} \sigma_{k} \tag{3.4.93}$$

$$H\left(\hat{h}_{k}, \frac{d \mu}{2}\right)=1+\frac{d \mu}{2} \sigma_{k}=1+d \mu L_{k} \tag{3.4.94}$$

$$L_{k}=\frac{1}{2} \sigma_{k} \tag{3.4.95}$$

The commutation relations are

$$\left[I_{1}, I_{2}\right]=I_{3} \tag{3.4.96}$$

$$\left[L_{1}, L_{2}\right]=-I_{3} \tag{3.4.97}$$

$$\left[L_{1}, I_{2}\right]=L_{3} \tag{3.4.98}$$

and cyclic permutations.

It is a well known result of the Lie-Cartan theory of continuous group that these infinitesimalgenerators 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

$$A_{2}=S A_{1} S^{-1} \tag{3.4.99}$$

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:

$$A_1|\xi\rangle_1 = |\xi'\rangle_1 \tag{3.4.100}$$

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

$$|\xi\rangle_2 = S|\xi\rangle_1, \quad |\xi'\rangle_2 = S|\xi'\rangle_1 \tag{3.4.101}$$

Inserting into Equation 3.4.100 we obtain

$$A_1 S^{-1}|\xi\rangle_2 = S^{-1}|\xi\rangle_2 \tag{3.4.102}$$

and hence

$$A_2|\xi\rangle_2 = |\xi\rangle_2 \tag{3.4.103}$$

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 × 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 Σ_1:

$$P_1' = A_1 P_1 A_1^\dagger \tag{3.4.104}$$

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:

$$P_2 = S P_1 S^\dagger \tag{3.4.105}$$

$$P_2 = S P_1' S^\dagger \tag{3.4.106}$$

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

$$S^{-1} P_2' \tilde{S}^\dagger = A_1 S^{-1} P_2 \tilde{S}^\dagger A_1^\dagger S \tag{3.4.107}$$

or

$$P_2' = A_2 P_2 A_1^\dagger \tag{3.4.108}$$

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 $A = a_0 l + \vec{a} \cdot \vec{\sigma}$ 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 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

$$\hat{v}_2 \cdot \vec{\sigma} = S\hat{v}_1 \cdot \vec{\sigma} S^1 \tag{3.4.109}$$

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 π about an axis which bisects the angle between $\hat{v}_1,$ and $\hat{v}_2$:

$$S = -i\hat{s} \cdot \vec{\sigma} \tag{3.4.110}$$

where

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

Even if $\hat{v}_1,$ and $\hat{v}_2$ are not real, it is easily verified that S as given formally by Equations 3.4.110and 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

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

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

This construction fails only if

$$\hat{v}_1 \cdot \hat{v}_2 + 1 = 0 \tag{3.4.113}$$

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

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

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 diagonizable, but normal matrices and only normal matrices are diagonizable 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:

$$\vec{a} = \vec{\alpha} + i\vec{\beta} \tag{3.4.115}$$

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

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

with a positive. A further similarity transformation with

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

transforms Equation 3.4.116into 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

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

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

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

This condition is satisfied if and only if

$$\hat{h} \cdot \hat{u} = 0 \tag{3.4.120}$$

and

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

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

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

However, if μ and ϕ are infinitesimal, we get

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

$$\mu^2-\phi^2=0 \tag{3.4.124}$$

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 inhomogeous 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

$$A_0^2=A \tag{3.4.125}$$

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

$$A_0^2=0 \tag{3.4.126}$$

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

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

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

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:

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

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

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

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

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

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

The foregoing results are summarized in Table 3.2.

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

Table 3.2: Canonical Forms for the Simlarity classes of A_2

CHAPTER OVERVIEW

4: Pauli Algebra and Electrodynamics

4.1: Lorentz transformation and Lorentz force

The main importance of the Pauli algebra is to provide us with a stepping stone for the theory of spinor spaces to which we turn in Section 5. Yet it is useful to stop at this point to show that the formalism already developed provides us with an efficient framework for limited, yet important aspects of classical electrodyanmics (CED).

We have seen on page 26 that the effect on electric field on a test charge, a "boost," can be considered as an active Lorentz transformation, whereby the field is proportional to the "hyperbolic angular velocity $\dot{\mu}$.

This is in close analogy with the well known relation between the magnetic field and the cyclotron frequency, i.e., a "circular angular velocity" $\dot{\phi}$. These results had been obtained under very special conditions. The Pauli algebra is well suited to state them in much greater generality.

The close connection between the algebra of the Lorentz group and that of the electromagnetic field, is well known. However, instead of developing the two algebras separately and noting the isomorphism of the results, we utilize the mathematical properties of the Lorentz group developed in Section 3 and translate them into the language of electrodynamics. The definition of the electromagnetic field implied by this procedure is, of course, hypothetical, and we turn to experience to ascertain its scope and limits. The proper understanding of the limitation of this conception is particularly important, as it serves to identify the direction for the deepening of the theory. The standard operational definition of the electromagnetic field involves the use of a test charge. Accordingly we assume the existence of particles that can act in such a capacity. The particle is to carry a charge e, a constant rest mass m, and the effect of the field acting during the time dt is to manifest itself in a change of the 4-momentum only, without involving any change in internal structure.

This means that the field has a sufficiently low frequency in the rest frame of the particle so as not to affect its internal structure. This is in harmony with the temporary exclusion of radiative interaction stated already.

Let the test charge be exposed to an electromagnetic field during a small time dt.. We propose to describe the resulting change of the four-momentum $P \to P' = P + dP$ as an infinitesimal Lorentz transformation. In this preliminary form the statement would seem to be trivial, since it is valid for any force that does not affect the intrinsic structure, say a combination of gravitational and frictional forces. In order to characterize specifically the Lorentz force, we have to add that the characterization of the field is independent of the fourmomentum of the test charge, moreover it is independent of the frame of reference of the observer. These conditions can be expressed formally in the following.

Postulate 4.1.1

The effect of the Lorentz force on a particle (test charge) is represented as the transformation of the four-momentum space of the particle unto itself, and the transformations are elements of the active Lorentz group. Moreover, matrix representations in different Lorentz frames are connected by similarity transformations. (See Section 3.4.4.)

We now proceed to show that this postulate implies the known properties of the Lorentz force.

First, we show that an infinitesimal Lorentz transformation indeed reduces to the Lorentz force provided we establish a "dictionary" between the parameters of the transformation and the electromagnetic field (see below Equation 4.1.12). Consider a pure Lorentz transformation along $\hat{h}$.

$$p' = p\cosh\mu + p_0\sinh\mu \tag{4.1.1}$$

$$p'_0 = p\sinh\mu + p_0\cosh\mu \tag{4.1.2}$$

where $\vec{p} = p\hat{h} + \vec{p}$ with $\vec{p}\cdot\vec{h} = 0$. For infinitesimal transformations $\mu \to d\mu$:

$$p' - p = p_0 d\mu \tag{4.1.3}$$

$$p'_0 - p_0 = p d\mu \tag{4.1.4}$$

or

$$\dot{\vec{p}} = p_0\dot{\mu}\hat{h} \tag{4.1.5}$$

$$\dot{p}_0 = \vec{p}\cdot\hat{h}\dot{\mu} \tag{4.1.6}$$

By making use of

$$\vec{p} = mc\sinh\mu = \gamma m\vec{v} \tag{4.1.7}$$

$p_0 = mc\cosh\mu = \gamma mc$

we obtain

$$\dot{\vec{p}} = pmc\hat{h}\dot{\mu} \tag{4.1.8}$$

$\dot{p}_0 = pm\frac{\vec{v}}{c}\cdot\hat{h}\dot{\mu}$

Turning to rotation we have from Equation A.3.8 of Appendix A (See note on page 51).

$$\vec{p}_\perp = \vec{p}_\perp\cos\phi + \hat{u}\times\vec{p}_\perp\sin\phi \tag{4.1.9}$$

For an infinitesimal rotation $\phi \simeq d\phi$, and by using Equation 4.1.7 we obtain, since $\vec{p}_\parallel = \vec{p}_\parallel$ and $\vec{p}_\parallel \times \hat{u} = 0$

$$\vec{p}' - \vec{p} = -\vec{p}\times\hat{u}d\phi = -\gamma m\vec{v}\times\hat{u}d\phi \tag{4.1.10}$$

or

$$\dot{\vec{p}} = -\gamma m\vec{v}\times\hat{u}d\phi \tag{4.1.11}$$

With the definitions of 3.3.28 and 3.3.29 of page 26 written vectorially:

$\vec{E} = \frac{\gamma mc}{e}\dot{\mu}\hat{h}$

$$\vec{B} = \frac{-\gamma mc}{e}\dot{\phi}\hat{u} \tag{4.1.12}$$

Equations 4.1.8 and 4.1.9 reduce to the Lorentz force equations.

Let us consider now an infinitesimal Lorentz transformation generated by

$$V = 1 + \frac{\mu}{2}\hat{h}\cdot\vec{\sigma} - \frac{i\phi}{2}\hat{u}\cdot\vec{\sigma} \tag{4.1.13}$$

$$= 1 + \frac{edt}{2\gamma mc}(\vec{E} + i\vec{B})\cdot\vec{\sigma} \tag{4.1.14}$$

$$= 1 + \frac{edt}{2\gamma mc}F \tag{4.1.15}$$

with

$$\vec{f} = \vec{E} + i\vec{B}, \quad F = \vec{f}\cdot\vec{\sigma} \tag{4.1.16}$$

It is apparent from Equations 4.1.15 and 4.1.16 that the transformation properties of V and F are identical. Since the transformation of V has been obtained already in Section 3.4.4, we can write down at once that of the field $\vec{f}$

Let us express the passive Lorentz transformation of the four-momentum P from the inertial frame Σ to Σ' as

$$P' = SPS^\dagger \tag{4.1.17}$$

where S is unimodular. The field matrix transforms by a similarity transformation:

$$F' = SFS^{-1} \tag{4.1.18}$$

with the complex reflections (contragradient entities) transforming as

$$\bar{P}' = \bar{S}\bar{P}\bar{S}^{-1} \tag{4.1.19}$$

$$\bar{F}' = \bar{S}\bar{F}\bar{S}^\dagger \tag{4.1.20}$$

For

$$S=H=\exp\left(-\frac{\mu}{2}\hat{h}\cdot\vec{\sigma}\right) \tag{4.1.21}$$

we obtain the passive Lorentz transformation for a frame Σ_2 moving with respect to Σ_1, with-the velocity

$$\vec{v}=c\hat{h}\tanh\mu \tag{4.1.22}$$

$$F'=HFH^{-1} \tag{4.1.23}$$

We extract from here the standard expressions by using the familiar decomposition

$$\vec{f}=\vec{f}_{\|}+\vec{f}_{\perp}=(\vec{f}\cdot\hat{h})\hat{h} \tag{4.1.24}$$

we get

$$\vec{f}'_{\|}=\vec{f}_{\|} \tag{4.1.25}$$

$$\overrightarrow{f_{\perp}}\cdot\vec{\sigma}=H\left(\overrightarrow{f_{\perp}}\cdot\vec{\sigma}\right)H^{-1}=H^{2}\left(\overrightarrow{f_{\perp}}\cdot\vec{\sigma}\right) \tag{4.1.26}$$

$$=(\cosh\mu-\sinh\mu\hat{h}\cdot\vec{\sigma})\vec{f}_{\perp}\cdot\vec{\sigma} \tag{4.1.27}$$

Hence

$$\vec{f}'_{\perp}=\vec{f}_{\perp}\cosh\mu+i\sinh\mu\hat{h}\times\overrightarrow{f_{\perp}} \tag{4.1.28}$$

$$=\cosh\mu\left(\vec{f}_{\perp}+i\tanh\mu\hat{h}\times\vec{f}_{\perp}\right) \tag{4.1.29}$$

$$=\gamma\left(\overrightarrow{f_{\perp}}+i\frac{\vec{v}}{c}\times\overrightarrow{f_{\perp}}\right) \tag{4.1.30}$$

where we used Equation 4.1.22. Inserting from Equation 4.1.16 we get

$\vec{E}'_{\perp}=\gamma\left(\vec{E}_{\perp}+\frac{\vec{v}}{c}\times\vec{B}_{\perp}\right)$

$$\vec{B}'_{\perp}=\gamma\left(\vec{B}_{\perp}-\frac{\vec{v}}{c}\times\vec{E}_{\perp}\right) \tag{4.1.31}$$

$\vec{E}'_{\|}=\vec{E}_{\|},\quad \vec{B}'_{\|}=\vec{B}_{\|}$

It is interesting to compare the two compact forms 4.1.18 and 4.1.31. Whereas the latter may be the most convenient for solving specific problems, the former will be the best stepping stone for the deepening of the theory. The only Lorentz invariant of the field is the determinant, which we write for convenience with the negative sign:

$$-|F|=\frac{1}{2}TrF^{2}=\vec{f}^{2}=\vec{E}^{2}-\vec{B}^{2}+2i\vec{E}\cdot\vec{B}=g^{2}e^{2}\Psi \tag{4.1.32}$$

Hence we obtain the well know invariants

$$I_1=\vec{E}^{2}-\vec{B}^{2}=g^{2}\cos 2\psi \tag{4.1.33}$$

$I_2=2\vec{E}\cdot\vec{B}=g^2\sin 2\psi$

We distinguish two cases

1. $\overrightarrow{f^2}\neq 0$
2. $\overrightarrow{f^2}=0$

These cases can be associated with the similarity classes of Table 3.2. In the case (i) F is unimodular axial, for (ii) it is nonaxial singular. (Since F is traceless, the two other entries in the table do not apply.) We first dispose of case (ii). A field having this Lorentz invariant property is called a null-field. The F matrix generates an exceptional Lorentz transformation (Section 3.4.4). In this field configuration $\vec{E}$ and $\vec{B}$ are perpendicular and are of equal size. This is a relativistically invariant property that is characteriestic of plane waves to be discussed in Section 4.2.

In the "normal" case (i) it is possible to find a canonical Lorentz frame, in which the electric and the magnetic fields are along the same line, they are parallel, or antiparallel. The Lorentz screw corresponds to a Maxwell wrench. It is specified by a unit vector $\hat{s}n$ and the values of the fields in the canonical frame E_{can} and B_{can}. The wrench may degenerate can can with $E_{can}=0$, or $B_{can}=0$. The canonical frame is not unique, since a Lorentz transformation along $\hat{n}$ leaves the canonical fields invariant.

We can evaluate the invariant eqn:iii-8-18ab in the canonical frame and obtain

$$E_{\text{can}}^{2}-B_{\text{can}}^{2}=I_{1}=g^{2}\cos 2\psi \tag{4.1.34}$$

$$2E_{\text{can}}B_{\text{can}}=I_{2}=g^{2}\sin 2\psi \tag{4.1.35}$$

One obtains from here

$$E_{can}=g\cos\psi \tag{4.1.36}$$

$$B_{can}=g\sin\psi \tag{4.1.37}$$

The invariant character of the field is determined by the ratio

$$\frac{B_{can}}{E_{can}}=\tan\psi \tag{4.1.38}$$

that has been called its pitch by Synge (op. cit. p. 333) who discussed the problem of canonical frames of the electromagnetic field with the standard tensorial method.

The definition of pitch in problem #8 is the reciprocal to the one here given and should be changed to agree with Eq. (4.1.20)

4.2: The Free Maxwell Field

Our approach to CED thus far is unusual inasmuch as we have effectively defined, classified and transformed the electromagnetic field at a small region of space-time without having used the Maxwell equations. This is, of course, an indication of the effectiveness of our definition of the field in terms of active Lorentz transformations.

In order to arrive at the Maxwell equations we invoke the standard principle of relativistic invariance, involving the passive interpretation of the Lorentz group.

Postulate 4.2.1

The electromagnetic field satisfies a first order differential equation in the space-time coordinates that is covariant under Lorentz transformations.

We consider the four-dimensional del operator $\{\partial_0, \nabla\}$ with $\partial_0 = \partial/\partial(ct) = \partial/\partial r_0$ as a fourvector, with its matrix equivalent

$$D = \partial_0 - \nabla \cdot \vec{\sigma} \tag{4.2.1}$$

The rationale for the minus sign is as follows. Let D operate on a function representing a plane wave:

$$\psi = \exp -i(\omega t - \vec{k} \cdot \vec{r}) = \exp -i\left(k_0 r_0 - \vec{k} \cdot \vec{r}\right) \tag{4.2.2}$$

we have

$$iD\psi = \left(k_0 1 + \vec{k} \cdot \vec{\sigma}\right)\psi = K\psi \tag{4.2.3}$$

Thus D has the same transformation properties as K:

$$D' = SDS^{\dagger} \tag{4.2.4}$$

while the complex reflection

$$\bar{D} = \partial_0 1 + \nabla \cdot \vec{\sigma} \tag{4.2.5}$$

transforms as $\bar{K}$, i.e.

$$\bar{D}' = \bar{S}\bar{D}S^{-1} \tag{4.2.6}$$

By using the transformation rules 4.1.18, 4.1.20 of the last section we see that $\bar{D}F$ transforms as a four-vector J:

$$\left(\bar{S}\bar{D}S^{-1}\right)\left(SFS^{-1}\right) = \bar{S}\bar{J}S^{-1} \tag{4.2.7}$$

Thus

$$\bar{D}F = \bar{J} \tag{4.2.8}$$

is a differential equation satisfying the conditions of Postulate 2. Setting tentatively

$$J = \rho 1 + \frac{\vec{j}}{c} \cdot \vec{\sigma} \tag{4.2.9}$$

with $\rho, \vec{j}$ the densities of charge and current, 4.2.8 is indeed a compact form of the Maxwell equations.

This is easily verified by sorting out the terms with the factors $(1, \sigma_k)$ and by separating the real and imaginary parts.

By operating on the Equation 4.2.9 with D and taking the trace we obtain

$$D\bar{D}F = \left(\partial_0^2 - \nabla^2\right)F = D\bar{J} \tag{4.2.10}$$

and

$$\frac{1}{2} Tr D\bar{J} = \partial_0 \rho + \frac{1}{c}\vec{\nabla} \cdot \vec{j} = 0 \tag{4.2.11}$$

These are standard results which are easily provided by the formalism. However, we do not have an explicit expression for J that would be satisfactory for a theory of radiative interaction.

Therefore, in accordance with our program stated in Section 4.1 we set $J=0$ and examine only the free field that obeys the homogeneous equations $\bar{D}F=0$

$$\left(\partial_0^2-\nabla^2\right)F=0 \tag{4.2.12}$$

The simplest elementary solution of 4.2.12 are monochromatic plane waves from which more complicated solutions can be built up. Hence we consider

$$F(\vec{r},t)=F_+(\vec{k},\omega)\exp\{i(\omega t-\vec{k}\cdot\vec{r})\}+F_-(\vec{k},\omega)\exp\{i(\omega t-\vec{k}\cdot\vec{r})\} \tag{4.2.13}$$

where $F_\pm$ are matrices independent of space-time. Inserting into Equation 4.2.12 yields the condition

$$\omega^2-c^2k^2=0 \tag{4.2.14}$$

Introducing the notation

$$\theta=k_0r_0-\vec{k}\cdot\vec{r} \tag{4.2.15}$$

we write Equation 4.2.13 as

$$F(\vec{r},t)=F_+\exp(-i\theta)+F_-\exp(i\theta) \tag{4.2.16}$$

Inserting into 4.2.12 we have

$$KF_\pm=K\left(\vec{E}_\pm+i\vec{B}_\pm\right)\cdot\vec{\sigma}=0 \tag{4.2.17}$$

From here we get explicitly

$$\vec{k}\cdot\left(\vec{E}_\pm+i\vec{B}_\pm\right)=0 \tag{4.2.18}$$

$$\vec{E}_\pm+i\vec{B}_\pm=i\hat{k}x\left(\vec{E}_\pm+i\vec{B}_\pm\right) \tag{4.2.19}$$

and we infer the well known properties of plane waves: $\vec{E}$ and $\vec{B}$ are of equal magnitude, and $\vec{E},\vec{B},\vec{k}$ from a right-handed Cartesian triad. We note that this constellation corresponds to the field of the type (ii) with $\vec{f}^{\,2}=0$ mentioned on page 51.

Since the classical $\vec{E},\vec{B}$ are real, we have also the relations

$$\vec{E}_-=\vec{E}_+^*,\quad \vec{B}_-=\vec{B}_+^* \tag{4.2.20}$$

Consider now the case in which

$$\vec{f}_-=\vec{E}_-+i\vec{B}_-=0 \tag{4.2.21}$$

which, in view of 4.2.20 implies

$$\vec{E}_+=i\vec{B}_+ \tag{4.2.22}$$

Thus at a fixed point and direction in space the electric field lags the magnetic field by a phase $\pi/2$, and $\vec{f}_+=\left(\vec{E}_++i\vec{B}_+\right)$ is the amplitude of a circularly polarized wave of positive helicity, i.e., the rotation of the electric and magnetic vectors and the wave vector $\vec{k}$ form a right screw or the linear and angular momentum point in the same direction $+\vec{k}$. In the traditional optical terminology this is called a left circularly polarized wave. However, following current practice, we shall refer to positive helicity as right polarization R. The negative helicity state is represented by $\vec{f}_-=\left(\vec{E}_-+i\vec{B}_-\right)$.

Actually, we have the alternative of associating $\overrightarrow{f_-^*}=\left(\vec{E}_--i\vec{B}_-\right)$ with R and $\overrightarrow{f_+^*}=\left(\vec{E}_+-i\vec{B}_+\right)$ with L. However, we prefer the first choice and will use the added freedom in the formalism to describe the absorption and the emission process at a later stage.

Meanwhile, we turn to the discussion of polarization which has a number of interesting aspects, particularly if carried out in the context of spinor algebra.

CHAPTER OVERVIEW

5: Spinor Calculus

5.1: From Triads and Euler Angles to Spinors - A Heuristic Introduction

As mentioned already in Section 3.4.3, it is an obvious idea to enrich the Pauli algebra formalism by introducing the complex vector space V(2, C) on which the matrices operate. The two-component complex vectors are traditionally called spinors. We wish to show that they give rise to a wide range of applications. In fact we shall introduce the spinor concept as a natural answer to a problem that arises in the context of rotational motion.

In Section 3 we have considered rotations as operations performed on a vector space. Whereas this approach enabled us to give a group-theoretical definition of the magnetic field, a vector is not an appropriate construct to account for the rotation of an orientable object. The simplest mathematical model suitable for this purpose is a Cartesian (orthogonal) three-frame, briefly, a triad. The problem is to consider two triads with coinciding origins, and the rotation of the object frame is described with respect to the space frame. The triads are represented in terms of their respective unit vectors: the space frame as $\Sigma_s\left(\hat{x}_1, \hat{x}_2, \hat{x}_3\right)$ and the object frame as $\Sigma_c\left(\hat{e}_1, \hat{e}_2, \hat{e}_3\right)$ Here c stands for "corpus," since o for "object" seems ambiguous. We choose the frames to be right-handed.

These orientable objects are not pointlike, and their parametrization offers novel problems. In this sense we may refer to triads as "higher objects," by contrast to points which are "lower objects." The thought that comes most easily to mind is to consider the nine direction cosines $\hat{e}_i \cdot \hat{x}_k$ but this is impractical, because of the six relations connecting these parameters. This difficulty is removed by the three independent Eulerian angles, a most ingenious set of constructs, which leave us nevertheless with another problem: these parameters do not have good algebraic properties; their connection with the ordinary Euclidean vector space is provided by rather cumbersome relations. This final difficulty is solved by the spinor concept.

The theory of the rotation of triads has been usually considered in the context of rigid body mechanics According to the traditional definition a rigid body is "a collection of point particles keeping rigid distances." Such a system does not lend itself to useful relativistic generalization. Nor is this definition easily reconciled with the Heisenberg principle of uncertainty.

Since the present discussion aims at applications to relativity and quantum mechanics, we hasten to point out that we consider a triad as a precise mathematical model to deal with objects that are orientable in space. Although we shall briefly consider the rigid body rotation in Section 5.2, the concept of rigidity in the sense defined above is not essential in our argument. We turn now to a heuristic argument that leads us in a natural fashion from triad rotation to the spinor concept. According to Euler's theorem any displacement of a rigid body fixed at a point O is equivalent to a rotation around an axis through O. (See [Whi64], page 2.)

This theorem provides the justification to describe the orientational configuration of Σ_c in terms of the unitary matrix in $\mathrm{su}(2)$ that produces the configuration in question from the standard position in which the two frames coincide. Denoting the unitary unimodular matrices corresponding to two configuration by V_1, V_2 a transition between them is conveyed by an operator U

$$V_2 = UV_1 \tag{5.1.1}$$

Let

$$\begin{aligned} V &= \cos\frac{\phi}{2} 1 - i \sin\frac{\phi}{2} \hat{v} \cdot \vec{\sigma} \\ &= q_0 1 - i\vec{q} \cdot \vec{\sigma} \end{aligned} \tag{5.1.2}$$

Here $q_0, \vec{q}$ are the so-called quaternion components, since the $(-i\sigma_k)$ obey the commutation rules of the quaternion units $e_k : e_1 e_2 = -e_2 e_1 = e_3$. We have

$$|V| = q_0^2 + \vec{q}^2 = q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1 \tag{5.1.3}$$

The Equations 5.1.1 - 5.1.3 can be given an elegant geometrical interpretation: $q_0, \vec{q}$ are considered as the coordinates of a point on the three-dimensional unit hypersphere in four-dimensional space $\mathcal{V}(4, R)$ Thus the rotation of the triad is mapped on the rotation of this hypersphere. The operation leaves 5.1.3 invariant.

The formalism is that of elliptic geometry, a counterpart to the hyperbolic geometry in Minkowski space.

This geometry implies a "metric": the "distance" of two displacements $V_1,$ and V_2 is defined as

$$\frac{1}{2}\mathrm{Tr}\left(V_2 \tilde{V}_1\right) = \cos\frac{\phi_1}{2}\cos\frac{\phi_2}{2} + \sin\frac{\phi_1}{2}\sin\frac{\phi_2}{2}\hat{v}_1 \cdot \hat{v}_2 \tag{5.1.4}$$

$$=\cos \frac{\phi}{2}=q_{10} q_{20}+\vec{q}_1 \cdot \vec{q}_2 \tag{5.1.5}$$

where φ is the angle of rotation carrying V_1, into V_2. Note the analogy with the hyperbolic formula 3.4.67 in Section 3.4.3.

We have here an example for an interesting principle of geometry: a "higher object" in a lower space can be often represented as a "lower object," i.e., a point in a higher space. The "higher object" is a triad in ordinary space $\mathcal{V}(3, R)$. It is represented as a point in the higher space $\mathcal{V}(4, R)$.

We shall see that this principle is instrumental in the intuitive interpretation of quantum mechanics. The points in the abstract spaces of this theory are to be associated with complex objects in ordinary space.

Although the representation of the rotation operator U and the rotating object V in terms of the same kind of parametrization can be considered a source of mathematical elegance, it also has a shortcoming. Rotating objects may exhibit a preferred intrinsic orientation, such as a figure axis, or the electron spin, for which there is no counterpart in Equations 5.1.1 and 5.1.3.

This situation is remedied by the following artifice. Let the figure axis point along the unit vector $\hat{e}_3$ that coincides in the standard position with $\hat{x}_3$ Instead of generating the object matrix V in terms of single rotation, we consider the following standard sequence to be read from right to left, (see Figure 5.1):

$$U\left(\hat{x}_3, \frac{\alpha}{2}\right) U\left(\hat{x}_2, \frac{\beta}{2}\right) U\left(\hat{x}_3, \frac{\gamma}{2}\right)=V(\alpha, \beta, \gamma) \tag{5.1.6}$$

Here α, β, γ are the well known Euler angles, and the sequence of rotations is one of the variants traditionally used for their definition.

The notation calls for explanation. We shall continue to use, as we did in Section 3, $U(\hat{u}, \phi / 2)$ for the 2 × 2 unitary matrix parametrized in terms of axis angle variables. We shall call this also a uniaxial parametrization, to be distinguished from the biaxial parametrization of the unitary V matrices in which both the spatial direction $\hat{x}_3,$ and the figure axis $\hat{e}_3$ play a preferred role.

In Equation 5.1.6 the rotations are defined along axes specified in the space frame Σ_s. However, in the course of each operation the axis is fixed in both frames. Thus it is merely a matter of another name (an alias I) to describe the operation (4) in Σ_c. We have then for the same unitary matrix

$$V(\alpha, \beta, \gamma)=U\left(\hat{e}_3, \frac{\gamma}{2}\right) U\left(\hat{e}_2, \frac{\beta}{2}\right) U\left(\hat{e}_3, \frac{\alpha}{2}\right) \tag{5.1.7}$$

Note the inversion of the sequence of operations involving the rotations a and y. This relation is to be interpreted in the kinematic sense: the body frame moves from the initial orientation of coincidence with Σ_s into the final position.

The equivalence of 5.1.6 and 5.1.7 can be recognized by geometrical intuition, or also by explicit transformations between Σ_s and Σ_c (See [Got66], p 268).

In the literature one often considers the sequence

$$U\left(\hat{x}_3^{\prime \prime}, \frac{\gamma}{2}\right) U\left(\hat{x}_2^{\prime}, \frac{\beta}{2}\right) U\left(\hat{x}_3, \frac{\alpha}{2}\right) \tag{5.1.8}$$

where $\hat{x}_2^{\prime},$ and $\hat{x}_3^{\prime \prime}$ are axis positions after the first and the second step respectively. This procedure seems to have the awkward property that the different rotations are performed in different spaces. On closer inspection, however, one notices that Equation 5.1.8 differs only in notation from Equation 5.1.7. In the usual static interpretation $\hat{e}_1, \hat{e}_2, \hat{e}_3$ is used only for the final configuration, and $\hat{x}_2^{\prime}, \hat{x}_3^{\prime \prime}$ are introduced as auxiliary axes. If, in contrast, one looks at the object frame kinematically, one realizes that at the instant of the particular rotations the following axes coincide:

$$\hat{x}_3=\hat{e}_3, \quad \hat{x}_2^{\prime}=\hat{e}_2, \quad \hat{x}_3^{\prime \prime}=\hat{e}_3, \tag{5.1.9}$$

We now write Equation 5.1.6 explicitly as

$$
\begin{aligned}
V(\alpha,\beta,\gamma) &= U\left(\hat{x}_3,\frac{\alpha}{2}\right)U\left(\hat{x}_2,\frac{\beta}{2}\right)U\left(\hat{x}_3,\frac{\gamma}{2}\right)\\
&=\begin{pmatrix} e^{-i\alpha/2} & 0\\ 0 & e^{i\alpha/2}\end{pmatrix}\begin{pmatrix}\cos(\beta/2) & -\sin(\beta/2)\\ \sin(\beta/2) & \cos(\beta/2)\end{pmatrix}\begin{pmatrix} e^{-i\gamma/2} & 0\\ 0 & e^{i\gamma/2}\end{pmatrix}\\
&=\begin{pmatrix} e^{-i\alpha/2}\cos(\beta/2)e^{-i\gamma/2} & -e^{-i\alpha/2}\sin(\beta/2)e^{i\gamma/2}\\ e^{i\alpha/2}\sin(\beta/2)e^{-i\gamma/2} & e^{i\alpha/2}\cos(\beta/2)e^{i\gamma/2}\end{pmatrix}\\
&=\begin{pmatrix}\xi_0 & -\xi_1^*\\ \xi_1 & \xi_0^*\end{pmatrix}
\end{aligned}
\tag{5.1.10}
$$

with

$$
\begin{aligned}
\xi_0 &= e^{-i\alpha/2}\cos(\beta/2)e^{-i\gamma/2}\\
\xi_1 &= e^{i\alpha/2}\sin(\beta/2)e^{-i\gamma/2}
\end{aligned}
\tag{5.1.11}
$$

The four matrix elements appearing in this relation are the so-called Cayley-Klein parameters. (See Equation 3.4.43 in Section 3.4.2.)

It is a general property of the matrices of the algebra $\mathcal{A}_2$, that they can be represented either in terms of components or in terms of matrix elements. We have arrived at the conclusion that the representation of a unitary matrix in terms of elements is suitable for the parametrization of orientational configuration, while the rotation operator is represented in terms of components (axisangle variables).

There is one more step left to express this result most efficiently. We introduce the two-component complex vectors (spinors) of $\mathcal{V}(2,C)$ already mentioned at the beginning of the chapter. In particular, we define two conjugate column vectors, or ket spinors:

$$
|\xi\rangle=\begin{pmatrix}\xi_0\\ \xi_1\end{pmatrix},\quad |\bar{\xi}\rangle=\begin{pmatrix}-\xi_1^*\\ \xi_0^*\end{pmatrix}
\tag{5.1.12}
$$

and write the unitary V matrix symbolically as

$$
V=(\langle\xi||\mid\bar{\xi}\rangle)
\tag{5.1.13}
$$

We define the corresponding bra vectors by splitting the Hermitian conjugate V horizontally into row vectors:

$$
V^{\dagger}=\begin{pmatrix}\xi_0^* & \xi_1^*\\ -\xi_1 & \xi_0\end{pmatrix}=\begin{pmatrix}\langle\xi|\\ \langle\bar{\xi}|\end{pmatrix}
\tag{5.1.14}
$$

or

$$
\langle\xi|=\left(\xi_0^*,\xi_1^*\right);\quad \langle\bar{\xi}|=(-\xi_1,\xi_0)
\tag{5.1.15}
$$

The condition of unitarity of V can be expressed as

$$
V^{\dagger}V=\begin{pmatrix}\langle\xi|\\ \langle\bar{\xi}|\end{pmatrix}(|\xi\rangle,|\bar{\xi}\rangle)
\tag{5.1.16}
$$

$$
=\begin{pmatrix}\langle\xi\mid\xi\rangle & \langle\xi\mid\bar{\xi}\rangle\\ \langle\bar{\xi}\mid\xi\rangle & \langle\bar{\xi}\mid\bar{\xi}\rangle\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}
\tag{5.1.17}
$$

yielding at once the conditions of orthonormality

$$
\begin{aligned}
\langle\xi\mid\xi\rangle=\langle\bar{\xi}\mid\bar{\xi}\rangle=1\\
\langle\xi\mid\bar{\xi}\rangle=\langle\bar{\xi}\mid\xi\rangle=0
\end{aligned}
\tag{5.1.18}
$$

These can be, of course, verified by direct calculation. The orthogonal spinors are also called conjugate spinors. We see from these relations that our definition of spin conjugation is, indeed, a sensible one. However, the meaning of this concept is richer than the analogy with the ortho-normality relation in the real domain might suggest.

First of all we express spin conjugation in terms of a matrix operation. The relation is nonlinear, as it involves the operation of complex conjugation $\mathcal{K}$.

We have

$$|\bar{\xi}\rangle = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \mathcal{K}|\xi\rangle = -i\sigma_2\mathcal{K}|\xi\rangle \tag{5.1.19}$$

and

$$\langle\bar{\xi}| = \mathcal{K}\langle\xi| \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \mathcal{K}\langle\xi|i\sigma_2 \tag{5.1.20}$$

We obtain from here

$$|\bar{\bar{\xi}}\rangle = -|\xi\rangle, \quad \langle\bar{\bar{\xi}}| = -\langle\xi| \tag{5.1.21}$$

The bar notation for spin conjugation suggests a connection with the complex reflection of the Pauli algebra. We shall see that such a connection indeed exists. However, we have to remember that, in contrast to Equation 5.1.21, complex reflection is involutive, i.e., its iteration is the identit $\bar{\bar{A}} = A$.

The emergence of the negative sign in Equation 5.1.21 is a well known property of the spin function, however we have to defer the discussion of this intriguing fact for later.

We shall occasionally refer to spinors normalized according to Equation 5.1.18 as unitary spinors, in order to distinguish them from relativistic spinors normalized as $\langle\xi \mid \xi\rangle = k_0$ where k_0 is the 0-th component of a four-vector.

Let us take a closer look at the connection between spinors and triads. In our heuristic procedure we started with an object triad specified by three orthonormal unit vectors $\hat{e}_1, \hat{e}_2, \hat{e}_3$ and arrived at an equivalent specification in terms of an associated spinor $|\xi\rangle$. Our task is now to start from the spinor and establish the corresponding triad in terms of its unit vectors. This is achieved by means of quadratic expressions.

We consider the so-called outer products

$$\begin{aligned} |\xi\rangle\langle\xi| &= \begin{pmatrix} \xi_0 \\ \xi_1 \end{pmatrix} (\xi_0^*, \xi_1^*) \\ &= \begin{pmatrix} \xi_0\xi_0^* & \xi_0\xi_1^* \\ \xi_1\xi_0^* & \xi_1\xi_1^* \end{pmatrix} \end{aligned} \tag{5.1.22}$$

and

$$\begin{aligned} |\xi\rangle\langle\bar{\xi}| &= \begin{pmatrix} \xi_0 \\ \xi_1 \end{pmatrix} (-\xi_1, \xi_0) \\ &= \begin{pmatrix} -\xi_0\xi_1 & \xi_0^2 \\ -\xi_1^2 & \xi_0\xi_1 \end{pmatrix} \end{aligned} \tag{5.1.23}$$

which can be considered as products of a 2×1 and 1×2 matrix.

In order to establish the connection with the unit vectors $\hat{e}_k$, we consider first the unit configuration in which the triads coincide: $\alpha = \beta = \gamma = 0$ i.e.

$$\xi_0 = 1, \xi_1 = 0 \quad \text{or} \quad |\xi\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{5.1.24}$$

with

$$|\bar{\xi}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{5.1.25}$$

Denoting these spinors briefly as $|1\rangle$ and $|\bar{1}\rangle$ respectively, we obtain from 5.1.22 and 5.1.23

$$|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(1 + \sigma_3) = \frac{1}{2}(1 + \hat{x}_3 \cdot \vec{\sigma}) \tag{5.1.26}$$

$$|1\rangle\langle\bar{1}|=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}=\frac{1}{2}(\sigma_1+i\sigma_2)=\frac{1}{2}(\hat{x}_1+i\hat{x}_2)\cdot\vec{\sigma} \tag{5.1.27}$$

Let V be the unitary matrix that carries the object frame from the unit position into $\Sigma_c\left(\hat{e}_1,\hat{e}_2,\hat{e}_3\right)$ Since $V^{\dagger}=V^{-1}$ and $\bar{V}=V$, we have

$$V|1\rangle=|\xi\rangle \quad V|\bar{1}\rangle=|\bar{\xi}\rangle \tag{5.1.28}$$

$$\langle 1|V^{-1}=\langle\xi| \quad \langle\bar{1}|V^{-1}=\langle\bar{\xi}| \tag{5.1.29}$$

By operating on 5.1.26 and 5.1.27 from left and right by V and V^{-1} respectively, we obtain

$$|\xi\rangle\langle\xi|=\frac{1}{2}(1+\hat{e}_3\cdot\vec{\sigma}) \tag{5.1.30}$$

$$|\xi\rangle\langle\bar{\xi}|=\frac{1}{2}(\hat{e}_1+i\hat{e}_2)\cdot\vec{\sigma} \tag{5.1.31}$$

and hence, by using Equation 3.4.13 of Section 3.4.2,

$$\hat{e}_1=\mathrm{Tr}(|\xi\rangle\langle\xi|\vec{\sigma})=\langle\xi|\vec{\sigma}|\xi\rangle \tag{5.1.32}$$

$$\hat{e}_1+i\hat{e}_2\equiv\hat{e}_+=\mathrm{Tr}(|\xi\rangle\langle\bar{\xi}|\vec{\sigma})=\langle\bar{\xi}|\vec{\sigma}|\xi\rangle \tag{5.1.33}$$

We have used here the rule:

$$\mathrm{Tr}(|\xi\rangle\langle\eta|)=\langle\eta\mid\xi\rangle \tag{5.1.34}$$

Equations 5.1.32 and 5.1.33 constitute a most compact expression for the relation between a spinor and its associated triad. One can extract from here the values of the direction cosines

$$\hat{e}_j\cdot\hat{x}_k\equiv e_{jk} \quad j,k=1,2,3 \tag{5.1.35}$$

$$\begin{aligned}\hat{e}_{31}&=\langle\xi|\sigma_1|\xi\rangle=\xi_0\xi_1^*+\xi_0^*\xi_1=\Re\left(\xi_0^*\xi_1\right)\\ \hat{e}_{32}&=\langle\xi|\sigma_2|\xi\rangle=i\left(\xi_0\xi_1^*-\xi_0^*\xi_1\right)=\Im\left(\xi_0^*\xi_1\right)\\ \hat{e}_{33}&=\langle\xi|\sigma_3|\xi\rangle=\xi_0\xi_0^*+\xi_1^*\xi_1 \quad (c)\end{aligned} \tag{5.1.36}$$

$$\begin{aligned}\hat{e}_{11}+i\hat{e}_{21}&=\langle\bar{\xi}|\sigma_1|\xi\rangle=\xi_0^2-\xi_1^2\\ \hat{e}_{12}+i\hat{e}_{22}&=\langle\bar{\xi}|\sigma_2|\xi\rangle=i\left(\xi_0^2+\xi_1^2\right)\\ \hat{e}_{13}+i\hat{e}_{23}&=\langle\bar{\xi}|\sigma_3|\xi\rangle=-2\xi_0\xi_1 \quad (c)\end{aligned} \tag{5.1.37}$$

By using Equation 5.1.11 we obtain these quantities in terms of Euler angles

$$\begin{aligned}e_{31}&=\sin\beta\cos\alpha\\ e_{32}&=\sin\beta\sin\alpha\\ e_{33}&=\cos\beta\end{aligned} \tag{5.1.38}$$

$$\begin{aligned}e_{11}&=\cos\gamma\cos\beta\cos\alpha-\sin\gamma\sin\alpha & e_{21}&=-\sin\gamma\cos\beta\cos\alpha-\cos\gamma\sin\alpha\\ e_{12}&=\cos\gamma\cos\beta\sin\alpha+\sin\gamma\cos\alpha & e_{22}&=-\sin\gamma\cos\beta\sin\alpha+\cos\gamma\cos\alpha\\ e_{13}&=-\cos\gamma\sin\beta & e_{23}&=\sin\gamma\sin\beta\end{aligned} \tag{5.1.39}$$

The relation between vectors and spinors displayed in Equations 5.1.36 can be established also by means of a stereographic projection. This method yields quicker results than the present lengthy build-up which in turn has a wider scope. Instead of rotating vector spaces, we operate on triads and thus obtain also Equation 5.1.37. To my knowledge, this relation has not appeared in the literature.

The Equations 5.1.36 and 5.1.37 solve the parametrization problem stated at the beginning of this chapter. The nine interrelated direction cosines e_{jk} are expressed by the three independent spinor parameters.

This is the counterpart of the parametrization problem concerning the nine parameters of the $\mathcal{SO}(3)$ matrices (see page 13), a problem that has been solved by the $\mathcal{SU}(2)$ representation $\mathcal{SO}(3)$ with the unitary matrices $U(\hat{u},\phi/2)$.

It is noteworthy that the decisive step was taken in both cases by Euler who introduced the "Euler angles" α, β, γ and also the axis-angle parameters $\hat{u}, \phi$ for the rotation operators.

Euler's results come to fruition in the version of spinor calculus in which spinors representing orientational states are parametrized in terms of Euler angles and the unitary operators in terms of $\hat{u}, \phi$.

We propose to demonstrate the ease by which this formalism lends itself to algebraic operations. This stems particularly from the constructs 5.1.30 and 5.1.31 in which we recognize the singular matrices of Table 3.2 (page 46).

We define more fully

$$\begin{aligned} |\xi\rangle\langle\xi| &= \frac{1}{2}(1+\hat{e}_3\cdot\vec{\sigma}) \equiv E_3 \\ |\bar{\xi}\rangle\langle\bar{\xi}| &= \frac{1}{2}(1-\hat{e}_3\cdot\vec{\sigma}) \equiv \bar{E}_3 \\ |\xi\rangle\langle\bar{\xi}| &= \frac{1}{2}(\hat{e}_1+i\hat{e}_2)\cdot\vec{\sigma} \equiv E_+ \\ |\bar{\xi}\rangle\langle\xi| &= \frac{1}{2}(\hat{e}_1-i\hat{e}_2)\cdot\vec{\sigma} \equiv E_- = -\bar{E}_+ \end{aligned} \tag{5.1.40}$$

Here $E_3, \bar{E}_3$ are idempotent projection operators and E_+, E nilpotent step operators. Since $E_3 + bar E_3 = 1$, we have

$$|\eta\rangle = |\xi\rangle\langle\xi \mid \eta\rangle + |\bar{\xi}\rangle\langle\bar{\xi} \mid \eta\rangle \tag{5.1.41}$$

$$= |\xi\rangle a_0 + |\bar{\xi}\rangle a_1 \tag{5.1.42}$$

with

$$a_0 = \langle\xi \mid \eta\rangle, \quad a_1 = \langle\bar{\xi} \mid \eta\rangle \tag{5.1.43}$$

Also

$$E_+|\bar{\xi}\rangle = |\xi\rangle \quad E_-|\xi\rangle = |\bar{\xi}\rangle \tag{5.1.44}$$

$$E_+|\xi\rangle = 0 \quad E_-|\bar{\xi}\rangle = 0 \tag{5.1.45}$$

We see from Equations 5.1.40 that the transition $|\xi\rangle \rightarrow |\bar{\xi}\rangle$ corresponds to an inversion of the figure axis with a simultaneous inversion of the figure axis with a simultaneous inversion of the γ-rotation around the axis. Therefore the transformation corresponds to a transition from a right to a left frame with a simultaneous change from counterclockwise to clockwise as the positive sense of rotation. Thus we should look at the transition from 5.1.40 as $E_+ \rightarrow \bar{E}_+$ or

$$\frac{1}{2}(\hat{e}_1+i\hat{e}_2)\cdot\vec{\sigma} \rightarrow \frac{1}{2}[-\hat{e}_1 - i(-\hat{e}_2)]\cdot\vec{\sigma} \tag{5.1.46}$$

All this is apparent also if we represent the transition $|\xi\rangle \rightarrow |\bar{\xi}\rangle$ in terms of Euler angles as

$$\alpha \rightarrow \pi + \alpha \tag{5.1.47}$$

$$\beta \rightarrow \pi - \beta \tag{5.1.48}$$

$$\gamma \rightarrow \pi - \gamma \tag{5.1.49}$$

We note also the following relations for later use:

$$E_- E_+ = \bar{E}_3, \quad E_+ E_- = E_3 \tag{5.1.50}$$

In addition to the short symbols $|\xi\rangle, |\bar{\xi}\rangle$ for spinors and their conjugates, we shall use also more explicit notations depending on the context:

$$|\alpha, \beta, \gamma\rangle = |\pi+\alpha, \pi-\beta, \pi-\gamma\rangle \tag{5.1.51}$$

$$|\hat{k}, \gamma\rangle, |\overline{\hat{k}, \gamma}\rangle = |-\hat{k}, \pi-\gamma\rangle \tag{5.1.52}$$

$$|\hat{k}\rangle|\bar{\hat{k}}\rangle = |-\hat{k}\rangle \tag{5.1.53}$$

Here $\hat{k}$ is the unit vector denoted by $\hat{e}_3$, in Equation 5.1.30. Its association with the spinor is evident from the following eigenvalue problem.

By using Equations 5.1.40 and 5.1.18 we obtain

$$\frac{1}{2}(1+\hat{k}\cdot\vec{\sigma})|\hat{k}\rangle = |\hat{k}\rangle\langle\hat{k}\mid\hat{k}\rangle = |\hat{k}\rangle \tag{5.1.54}$$

$$\frac{1}{2}(1-\hat{k}\cdot\vec{\sigma})|\bar{\hat{k}}\rangle = |\bar{\hat{k}}\rangle\langle\bar{\hat{k}}\mid\bar{\hat{k}}\rangle = |\bar{\hat{k}}\rangle \tag{5.1.55}$$

Hence

$$\hat{k}\cdot\vec{\sigma}|\hat{\rangle}k = |\hat{k}\rangle \tag{5.1.56}$$

$$\hat{k}\cdot\vec{\sigma}|\hat{\rangle}k = |\hat{k}\rangle \tag{5.1.57}$$

Thus $|k\rangle$ and $|\bar{\hat{k}}\rangle$ are eigenvectors of the Hermitian operator $\hat{k}\cdot\vec{\sigma}$ with the eigenvalues +1 and −1 respectively. This is a well known result, although usually obtained by a somewhat longer computation.

By using the explicit expression for $U(\hat{k},\phi/2)$ we obtain from 5.1.56 and 5.1.57

$$U(\hat{k},\phi/2)|\hat{k},\gamma\rangle = \exp(-i\phi/2)|\hat{k},\gamma\rangle = |\hat{k},\gamma+\phi\rangle \tag{5.1.58}$$

$$U(\hat{k},\phi/2)|\hat{k},\hat{\gamma}\rangle = \exp(i\phi/2)|\hat{k},\hat{\gamma}\rangle = |\overline{\hat{k},\gamma+\phi\rangle} \tag{5.1.59}$$

There is also the unitary diagonal matrix

$$U\left(\hat{x}_3,\phi/2\right) = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix} \tag{5.1.60}$$

the effect of which is easily described:

$$U\left(\hat{x}_3,\phi/2\right)|\alpha,\beta,\gamma\rangle = |\alpha+\phi,\beta,\gamma\rangle \tag{5.1.61}$$

These relations bring out the "biaxial" character of spinors: both $\hat{x}_3$, and $\hat{k}$ play a distinguished role. The same is true of a unitary matrix parametrized in terms of Euler angles: $V(\alpha,\beta,\gamma)$ Cayley-Klein parameters. This is to be contrasted with the uniaxial form $U(\hat{u},\phi/2)$.

Our discussion in this chapter has been thus far purely geometrical although active transformations of geometrical objects can be given a kinematic interpretation. We go now one step further and introduce the conept of time. By setting $\phi=\omega t$ with a constant ω in the unitary rotation operator we obtain the description of rotation processes:

$$\begin{aligned} U\left(\hat{k},\tfrac{\omega t}{2}\right)\left|\hat{k},\tfrac{\gamma}{2}\right\rangle = \exp(-i\omega t/2)\left|\hat{k},\tfrac{\gamma}{2}\right\rangle = \left|\hat{k},\tfrac{\gamma+\omega t}{2}\right\rangle \\ U\left(\hat{k},\tfrac{\omega t}{2}\right)|\hat{k},\gamma\rangle = \exp(i\omega t/2)|\hat{k},\gamma\rangle = \left|\hat{k},\tfrac{\gamma+\omega t}{2}\right\rangle \end{aligned} \tag{5.1.62}$$

These rotations are stationary, because U operates on its eigenspinors. There are various ways to represent the evolution of arbitrary spinors as well. We have

$$\begin{aligned} U\left(\hat{k},\frac{\omega t}{2}\right)|\eta\rangle &= \exp\left(-i\frac{\omega t}{2}\hat{k}\cdot\vec{\sigma}\right)|\eta\rangle \\ \langle\eta|U^{-1}\left(\hat{k},\frac{\omega t}{2}\right) &= \langle\eta|\exp\left(i\frac{\omega t}{2}\hat{k}\cdot\vec{\sigma}\right) \end{aligned} \tag{5.1.63}$$

Or, in differential form

$$\begin{aligned} i|\dot{\eta}\rangle &= \frac{\omega}{2}\hat{k}\cdot\vec{\sigma}|\eta\rangle \\ -i\langle\dot{\eta}| &= \langle\eta|\frac{\omega}{2}\hat{k}\cdot\vec{\sigma} \end{aligned} \tag{5.1.64}$$

The state functions solving these differential equations are obtained explicitly by using Equations 5.1.42, 5.1.43, 5.1.62, and 5.1.63:

$$|\eta(t)\rangle = U\left(\hat{k}, \frac{\omega t}{2}\right)|\eta(0)\rangle = \exp(-i\omega t/2)|\hat{k}\rangle a_0 + \exp(i\omega t/2)|\bar{\hat{k}}\rangle a_1 \tag{5.1.65}$$

and similarly for $\langle\eta(t)|$

By introducing the symbol H for the Hermitian operator $H = (\omega t/2)\hat{k}\cdot\vec{\sigma}$ in 5.1.64 we obtain

$$i|\dot{\eta}\rangle = H|\eta\rangle \tag{5.1.66}$$

$$-i\langle\dot{\eta}| = \langle\eta|H \tag{5.1.67}$$

These equations are reminiscent of the Schrödinger ¨ equation. Also it would be easy to derive from here a Heisenberg type operator equation.

It must be apparent to those familiar with quantum mechanics, that our entire spinor formalism has a markedly quantum mechanical flavor. What all this means is that the orientability of objects is of prime importance in quantum mechanics and the concept of the triad provides us with a more direct path to quantization, or to some aspects of it, than the traditional point mass approach.

In order to make use of this opportunity, we have to apply our spinor formalism to physical systems.

Our use of the concept of time in Equations 5.1.62–5.1.66 is quite formal. We merely selected a one-parameter subgroup of the rotation group to describe possible types of stationary rotation.

We have to turn to experiment, or to an experimentally established dynamical theory, to decide whether such motions actually occur in nature. We shall examine this question in connection with the rigid body rotation in the next section.

However, our main objective is the discussion of polarized light. Here the connection between classical and quantum theories is very close and the quantization procedure is particularly clear in terms of the spinor formalism.

The fact that the same formalism can be adjusted both to rigid body motion and to a wave phenomenon is interesting by itself. We know that the particle-wave duality is among the central themes in quantum mechanics. The contrast between these objects is very pronounced if we confine ourselves to point particles and to scalar waves. It is remarkable how this contrast is toned down within the context of rotational problems.

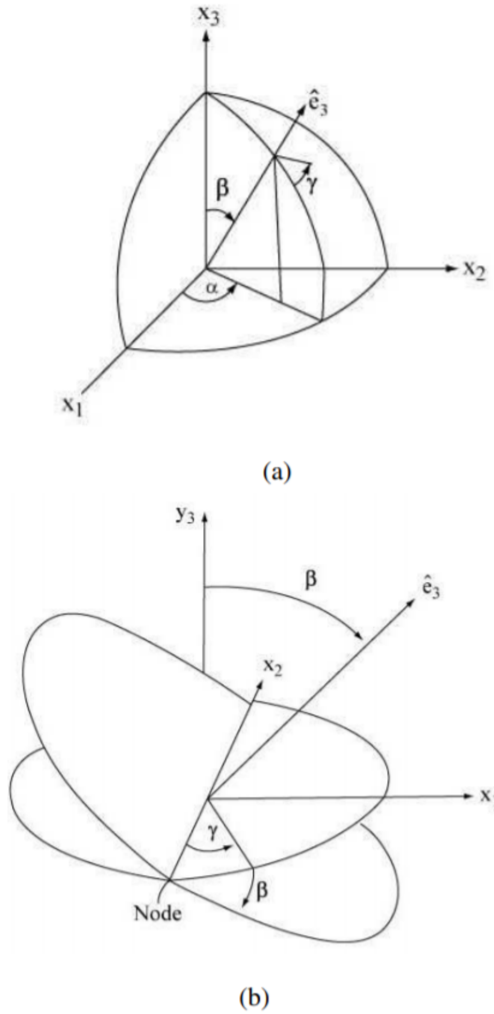

Figure 5.1: Euler Angles: (a) Static Picture. (b) Kinematics Display.

5.2: Rigid Body Rotation

In Equations 5.1.62-5.1.65 we have introduced the concept of time formally as a parameter to specify some simple types of motion which have a stationary character.

We examine now the usefulness of these results by considering the inertial motions of a rigid body fixed at one of its points, the so-called gyroscope.

We may sum up the relevant experimental facts as follows: there are objects of a sufficiently high symmetry (the spherical top) that indeed display a stationary, inertial rotation around any of their axes. In the general case (asymmetric top) such a stationary rotation is possible only around three principal directions marked out in the body triad.

The point of greatest interest for us, however, is the fact that there are also modes of motion that can be considered stationary in a weaker sense of the word.

We mean the so-called precession. We shall consider here only the regular precession of the symmetric top, or gyroscope, that can be visualized in terms of the well known geometrical construction developed by Poinsot in 1853. The motion is produced by letting a circular cone fixed in the body triad Σ_c roll over a circular cone fixed in the space triad Σ_s (see Figure 5.2).

The noteworthy point is that the biaxial nature of spinors renders them well suited to provide an algebraic counterpart to this geometric picture.

In order to prove this point we have to make use of the theorem that angular velocities around different axes can be added according to the rules of vectorial addition. This theorem is a simple corollary of our formalism.

Let us consider the composition of infinitesimal rotations with $\delta\phi = \omega\delta t << 1$:

$$U_2\left(\hat{u}_2, \frac{\omega_2\delta t}{2}\right) U_1\left(\hat{u}_1, \frac{\omega_1\delta t}{2}\right) \simeq \left(1 - \frac{\omega_2\delta t}{2}\hat{u}_2\cdot\vec{\sigma}\right)\left(1 - \frac{\omega_1\delta t}{2}\hat{u}_1\cdot\vec{\sigma}\right) \tag{5.2.1}$$

$$\simeq 1 - \frac{\delta t}{2}(\omega_2\hat{u}_2 + \omega_1\hat{u}_1)\cdot\vec{\sigma} \tag{5.2.2}$$

We define the angular velocity vectors

$$\vec{\omega} = \omega\hat{u} \tag{5.2.3}$$

and notice from Equation 5.2.1 that

$$\vec{\omega}_1 + \vec{\omega}_2 = \vec{\omega} \tag{5.2.4}$$

Consequently we obtain for the situation presented in Figure 5.2:

$$\vec{\omega} = \dot{\gamma}\hat{e}_3 + \dot{\alpha}\hat{x}_3 \tag{5.2.5}$$

$$\omega^2 = \dot{\alpha}^2 + \dot{\gamma}^2 + 2\dot{\alpha}\dot{\gamma}\cos\beta \tag{5.2.6}$$

$$\beta = \hat{x}_3\cdot\hat{e}_3 \tag{5.2.7}$$

We can describe the precession in spinorial terms as follows. We describe the gyroscope configuration in terms of the unitary matrix 5.1.10 and operate on it from right and left with two unitary operators:

$$\begin{aligned} V(t) =& U\left(\hat{x}_3, \frac{\dot{\alpha}t}{2}\right) V(0) U\left(\hat{e}_3, \frac{\dot{\gamma}t}{2}\right) \\ =& \begin{pmatrix} e^{-i\dot{\alpha}t/2} & 0 \\ 0 & e^{i\dot{\alpha}t/2} \end{pmatrix} \begin{pmatrix} e^{-i\alpha(0)/2}\cos(\beta/2)e^{-i\gamma(0)/2} & -e^{-i\alpha(0)/2}\sin(\beta/2)e^{i\gamma(0)/2} \\ e^{i\alpha(0)/2}\sin(\beta/2)e^{-i\gamma(0)/2} & e^{i\alpha(0)/2}\cos(\beta/2)e^{i\gamma(0)/2} \end{pmatrix} \\ &\times \begin{pmatrix} e^{-i\dot{\gamma}t/2} & 0 \\ 0 & e^{i\dot{\gamma}t/2} \end{pmatrix} \end{aligned} \tag{5.2.8}$$

Thus

$$\alpha = \alpha(0) + \dot{\alpha}t, \quad \gamma = \gamma(0) + \dot{\gamma}t \tag{5.2.9}$$

This relation displays graphically the biaxial character of the V matrix. Thus premultiplication corresponds to rotation in Σ_s and postmultiplication to that in Σ_c.

Note that the situations represented in Figure 5.2 (a) and (b) are called progressive and retrograde precessions respectively.

The rotational axis $\vec{\omega}$ is instantaneously at rest in both frames. The vector components can be expressed as follows:

$$\begin{array}{c|c} \Im\Sigma_s & \Im\Sigma_c \\ \hline \omega_1 = \dot{\gamma}\sin\beta\cos\alpha & \omega_1 = \dot{\alpha}\sin\beta\cos\gamma \\ \omega_2 = \dot{\gamma}\sin\beta\sin\alpha & \omega_2 = \dot{\alpha}\sin\beta\sin\gamma \\ \omega_3 = \dot{\gamma}\cos\beta + \dot{\alpha} & \omega_3 = \dot{\alpha}\cos\beta + \dot{\gamma} \end{array} \tag{5.2.10}$$

These expressions can be derived formally from Equation 5.2.8. The left column of Equation 5.2.10 follows from the application of the left-operator on a ket spinor and the right column of Equation 5.2.10 from a right operation on a bra spinor.

Another way of arriving at these results is as follows: Expressions in the left column of Equation 5.2.10 are evident from the vector addition rule given in 5.2.4. Expressions in the right column of Equation 5.2.10 do not follow so easily from geometrical intuition. However, we can invoke the kinematic relativity between the two triads. A rotation of Σ_c with respect to Σ_s can be thought of also as the reverse rotation of Σ_s in Σ_c. Thus V_c is equivalent to

$$V_s^{-1} = V_s^{\dagger} = V_s(-\alpha, -\beta, -\gamma) \tag{5.2.11}$$

and we arrive from left to the right columns in Equation 5.2.10 by the following substitution:

$$\alpha \rightarrow -\gamma \tag{5.2.12}$$

$$\beta \rightarrow -\beta \tag{5.2.13}$$

$$\gamma \rightarrow -\alpha \tag{5.2.14}$$

$$t \rightarrow -t \tag{5.2.15}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \tag{5.2.16}$$

Up to this point the discussion has been only descriptive, kinematic. We have to turn to dynamics to answer the deeper questions as to the factors that determine the nature of the precession in any particular instance.

We invoke the kinematic relation Equation 5.1.62:

$$\exp\left(-i\frac{\omega t}{2}\hat{k}\cdot\vec{\sigma}\right)\left|\hat{k}, \frac{\gamma}{2}\right\rangle = \exp\left(-i\frac{\omega t}{2}\right)\left|\hat{k}, \frac{\gamma}{2}\right\rangle \tag{5.2.17}$$

This expression offers a way of generalization to dynamics. If this relation indeed describes a stationary process, then the generators $\frac{i}{2}\sigma_j$ of the unitary operator are constants of that motion. Later we shall pursue this idea to establish the concept of angular momentum and its quantizations. However, at this preliminary stage we are merely looking for an elementary illustration of the formalism, and we draw on the standard results of rigid body dynamics.

The dynamic law consists of three propositions. First, we have in Σ_s

$$\frac{d\vec{L}_s}{dt} = \vec{N} \tag{5.2.18}$$

where N is the external torque.

In Σ_c we have a constitutive relation connecting angular velocity and angular momentum. We assume that the object triad is along the principal axes of inertia:

$$\begin{aligned} L_{c1} &= I_1\omega_1 \\ L_{c2} &= I_2\omega_2 \\ L_{c3} &= I_3\omega_3 \end{aligned} \tag{5.2.19}$$

Finally, the angular momentum components in Σ_s and Σ_c are connected by the relation

$$\frac{d\vec{L}_s}{dt} = \frac{d\vec{L}_c}{dt} + \vec{\omega} \times \vec{L}_c \tag{5.2.20}$$

Equations 5.2.18–5.2.20 imply the Euler equations. Dynamically the precession may stem either from an external torque, or from the anisotropy of the moment of inertia (or both).

We shall assume $\vec{N} = 0$ and $I_1 = I_2 \neq I_3$. The Euler equations simplified accordingly yield for the precession as viewed in Σ_c:

$$\begin{aligned} \dot{\omega}_{c1} + i\dot{\omega}_{c2} &= -i\left(\omega_{c1} + i\omega_{c2}\right)\omega_3\delta \\ I_3\dot{\omega}_3 &= 0 \end{aligned} \tag{5.2.21}$$

with

$$\delta = 1 - \frac{I_3}{I_1} \tag{5.2.22}$$

From Equation 5.2.10, right hand column, row (a) and (b), we obtain

$$\dot{\omega}_1 + i\dot{\omega}_2 = -i\dot{\gamma}\left(\omega_1 + i\omega_2\right) \tag{5.2.23}$$

and by comparison with 5.2.21 we have

$$\dot{\gamma} = \omega_3\delta = \omega_3\left(1 - \frac{I_3}{I_1}\right) \tag{5.2.24}$$

We obtain from 5.2.24, 5.2.22 and 5.2.10 (the right hand column, row c):

$$I_3\omega_3 = I_1\dot{\alpha}\cos\beta \tag{5.2.25}$$

and

$$\frac{\dot{\gamma}}{\dot{\alpha}\cos\beta} = \frac{I_1}{I_3} - 1 \tag{5.2.26}$$

Thus the nature of inertial precession is determined by the inertial anisotropy 5.2.22. In particular, let $\cos\beta > 0$, then

$$I_1 > I_3 \rightarrow \frac{\dot{\gamma}}{\dot{\alpha}} > 0 \text{ see Figure 5.2 -a} \tag{5.2.27}$$

$$I_1 < I_3 \rightarrow \frac{\dot{\gamma}}{\dot{\alpha}} < 0 \text{ see Figure } 5.2 - \text{b} \tag{5.2.28}$$

Finally, from Equation 5.2.25 $L_{c3} = I_3\omega_3 = I_1\dot{\alpha}\cos\beta$ and the total angular momentum squared is

$$L^2 = I_1^2\dot{\alpha}^2 \tag{5.2.29}$$

Note that $L_{c3} = I_1\dot{\alpha}\cos\beta$ is the projection of the total angular momentum on to the figure axis. The precession γ in Σ_c comes about if $I_3 \neq I_1$. For further detail we refer to [KS65].

We note also that Euler's theory has been translated into the modern language of Lie Groups by V. Arnold ([Arn66] pp 319-361). However in this work stationary motions are required to have fixed rotational axes.

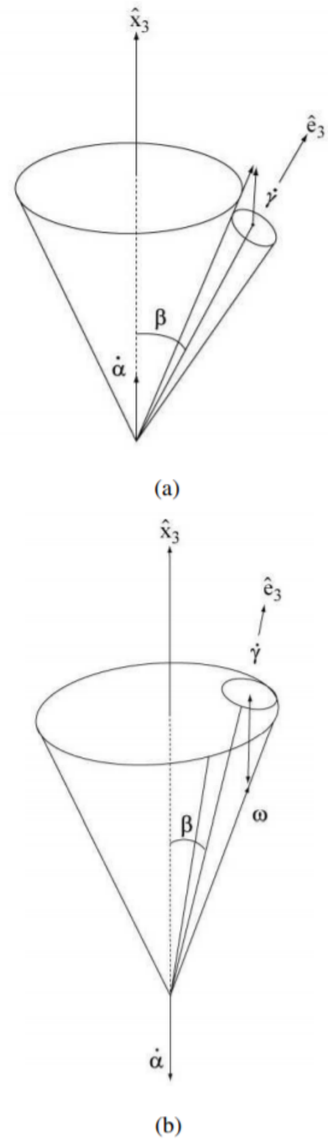

Figure 5.2: Progressive (a) and Retrograde (b) Precession.

5.3: Polarized Light

Polarization optics provides a most appropriate field of application for the Pauli algebra and the spinor formalism. Historically, of course, it went the other way around, and various aspects of the formalism had been advanced by many authors, often through independent discovery in response to a practical need.

In the present discussion we forego the historical approach and by using the mathematical formalism already developed, we arrive at the consolidation and streamlining of much disconnected material.

Another factor which simplifies our argument is that we do not attempt to describe polarization in all the complexity of a real situation, but concentrate first on a simple mathematical model, the two-dimensional isotropic, harmonic oscillator. This is, of course, the standard method of the elementary theory, however, by translating this description into the spinorial formalism, we set the stage for generalizations. A potential generalization would be to establish the connection with the statistical theory of coherence. However, at the present stage we shall be more concerned with applications to quantum mechanics.

Let us consider a monochromatic, polarized plane wave propagating in the z direction and write for the x and y components of the electric field

$$\begin{aligned} E_x &= p_1 \cos(\omega t + \phi_1) = p_1 \cos\tau \\ E_y &= p_2 \cos(\omega t + \phi_2) = p_2 \cos(\tau - \phi) \end{aligned} \tag{5.3.1}$$

where

$$\phi = \phi_1 - \phi_2, \quad p_1, p_2 \geq 0 \tag{5.3.2}$$

Let us define new parameters:

$$\begin{aligned} p_1 &= p\cos\frac{\theta}{2} \\ &0 \leq \theta \leq \pi \\ p_2 &= p\sin\frac{\theta}{2} \end{aligned} \tag{5.3.3}$$

It is convenient to express the information contained in Equations 5.3.1–5.3.3 in terms of the spinor

$$|\hat{k}\rangle = p \begin{pmatrix} e^{-i\phi/2} & \cos(\theta/2) \\ e^{i\phi/2} & \sin(\theta/2) \end{pmatrix} e^{-i\psi/2} \tag{5.3.4}$$

Here $\psi = \omega t + \phi_1$ 1 represents the common phase of the two components which does not affect the state of polarization. However, the presence of this third angle is in line with our definition of spinor in Equations 5.3.10 and 5.3.11 in Section 5.1. It will prove to be of significance in the problem of beam splitting and composition. By normalizing the intensity and setting $p = 1$, the spinor 5.3.4 conforms to our unitary normalization of Section 5.1.

By using Equations 5.3.30, 5.3.36 and 5.3.38 of Section 5.1 we obtain

$$|\hat{k}\rangle\langle\hat{k}| = \frac{1}{2}(1 + \hat{k}\cdot\vec{\sigma}) \tag{5.3.5}$$

$$\begin{aligned} k_1 &= \left\langle \hat{k}\,|\sigma_1|\,\hat{k} \right\rangle = \sin\theta\cos\phi = 2p_1p_2\cos\phi \\ k_2 &= \left\langle \hat{k}\,|\sigma_2|\,\hat{k} \right\rangle = \sin\theta\sin\phi = 2p_1p_2\sin\phi \\ k_3 &= \left\langle \hat{k}\,|\sigma_3|\,\hat{k} \right\rangle = \cos\theta = p_2^2 - p_1^2 \end{aligned} \tag{5.3.6}$$

In such a fashion the spinor 5.3.5, and hence each state of polarization is mapped on the surface of the unit sphere, the so-called Poincare sphere.

We see that the unit vector $(1,0,0)(\theta = \pi/2, \phi = 0)$ corresponds to linear polarization along $\frac{1}{\sqrt{2}}(\hat{x} + \hat{y})$, or $|45^\circ\rangle$ $(0,1,0)(\theta = \pi/2, \phi = \pi/2)$ corresponds to right circularly polarized light $|R\rangle$, and $(0,0,1)$, and $(0,0,1)$ or $\theta = 0$ to linear polarization in the $\hat{x}$ direction: $|\hat{x}\rangle$. . (See Figures 5.3 and 5.3.)

There is an alternative, and even more favored method of parametrizing the Poincaré sphere, in which the preferred axis for the definition of spherical coordinates corresponds to light of positive helicity $|R\rangle$. This choice implies a new set of spherical angles, say α, β to replace ϕ, θ. Their relation is displayed geometrically in Figures 5.3 and 5.4. The corresponding algebraic treatment is summed up as follows.

We relabel the Cartesian axes in the "Poincare space" as

$$\begin{aligned} k_3 &= s_1 = \sin\beta\cos\alpha \\ k_1 &= s_2 = \sin\beta\sin\alpha \\ k_2 &= s_3 = \cos\beta \end{aligned} \tag{5.3.7}$$

The vector $\hat{s}$ is associated with the unitary spinor

$$|\hat{s}\rangle = \begin{pmatrix} \exp(-i\alpha/2) & \cos(\beta/2) \\ \exp(i\alpha/2) & \sin(\beta/2) \end{pmatrix} \tag{5.3.8}$$

and

$$|\hat{s}\rangle\langle\hat{s}| = \frac{1}{2}(1 + \hat{s}\cdot\sigma) \tag{5.3.9}$$

The advantage of this choice is that the angles α, β have a simple meaning. We claim that

$$\begin{aligned} a_1 &= a\cos\left(\tfrac{1}{2}\left(\tfrac{\pi}{2} - \beta\right)\right) \\ a_2 &= a\sin\left(\tfrac{1}{2}\left(\tfrac{\pi}{2} - \beta\right)\right) \end{aligned} \tag{5.3.10}$$

where a_1, a_2 are the half major and minor axes of the ellipse traced by the $\vec{E}$ vector; we associate a positive and negative a_2 with an ellipse circled in the positive and negative sense respectively. Moreover, the angle α is twice the angle of inclination of the major axis against the x axis (Figure 5.4-d). The angle γ refers to the overall phase in complete analogy to ψ.

The proof of these statements are found in Born and Wolf (see pp. 24-32 of [BW64], the later editions are almost unchanged). A somewhat simplified derivation follows.

First we prove that Equations 5.3.1 and 5.3.2 provide indeed a parametric representation of an ellipse. The elimination of τ from the two equations 5.3.1 yields

$$\left(\frac{E_1}{p_1\sin\phi}\right)^2 - \left(\frac{2E_1E_2\cos\phi}{p_1p_2\sin^2\phi}\right) + \left(\frac{E_2}{p_2\sin\phi}\right)^2 = 1 \tag{5.3.11}$$

This is an equation of the form

$$\sum_{i=1}^{2} a_{ik}x_ix_k = 1 \tag{5.3.12}$$

with the a_{ik} real, symmetric, and $a_{11} > 0$,

$$a_{11}a_{22} - a_{12}^2 > 0 \tag{5.3.13}$$

The axes of the ellipse are derived from the eigenvalue problem:

$$\begin{aligned} (a_{11} - \lambda)x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 &= 0 \end{aligned} \tag{5.3.14}$$

Hence

$$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}^2 = 0 \tag{5.3.15}$$

with

$$\lambda_1 = \frac{1}{a_1^2}, \quad \lambda_2 = \frac{1}{a_2^2} \tag{5.3.16}$$

where a_1, a_2 are the half major and half minor axes respectively. We have, by inserting for the a_{ik}, from Equation 5.3.11

$$\lambda_1+\lambda_1=a_{11}+a_{22}=\left(\frac{1}{p_1^2}+\frac{1}{p_2^2}\right)\frac{1}{\sin^2\phi} \tag{5.3.17}$$

$$\lambda_1+\lambda_1=a_{11}a_{22}-a_{12}^2=\frac{1}{p_1^2p_2^2\sin^2\phi} \tag{5.3.18}$$

From these equations we have

$$a_1^2a_2^2=p_1^2p_2^2\sin^2\phi \tag{5.3.19}$$

$$a_1^2+a_2^2=p_1^2+p_2^2 \tag{5.3.20}$$

From Equation 5.3.19 we have $a_1a_2=\pm p_1p_2\sin\phi$. We require

$$a_1a_2=p_1p_2\sin\phi \tag{5.3.21}$$

and let $a_2<0$ for $\sin\phi<0$

We introduce now the auxiliary angle 3 as defined in Equation 5.3.10. With such an assignment $\beta=0,\pi$ correspond indeed to right and left circularly polarized light $|R\rangle, |\bar{R}\rangle$ respectively. More over $a_1\geq|a_2|$. Hence a_1 is the half major axis.

From Equations 5.3.3, 5.3.19 and 5.3.10 we obtain

$$\cos\beta=\sin\theta\sin\phi \tag{5.3.22}$$

We complete the parametrization of ellipticity by introducing $\alpha/2$ for the angle between the major axis and the $\hat{x}$ direction (Figure 5.4-c).

From Equation 5.3.14 we have

$$\tan\frac{\alpha}{2}=\frac{x_2}{x_1}=\frac{\lambda-a_{11}}{a_{12}}=\frac{a_{12}}{\lambda-a_{22}} \tag{5.3.23}$$

and

$$\begin{aligned}\tan\alpha&=\frac{\tan\frac{\alpha}{2}+\tan\frac{\alpha}{2}}{1-\tan^2\frac{\alpha}{2}}=\frac{\frac{\lambda-a_{11}}{a_{12}}+\frac{a_{12}}{\lambda-a_{22}}}{\frac{\lambda-a_{11}}{\lambda-a_{22}}}\\&=\frac{(\lambda-a_{11})(\lambda-a_{22})\frac{1}{a_{12}}+a_{12}}{a_{11}-a_{22}}=\frac{2a_{12}}{a_{11}-a_{22}}\\&=\frac{2p_1p_2\cos\phi}{p_1^2-p_2^2}=\frac{\sin\theta\cos\phi}{\cos\theta}=\tan\theta\cos\phi\end{aligned} \tag{5.3.24}$$

It is apparent from Equation 5.3.22 that the axis s_3 can be indeed identified with k_2. Moreover Equation 5.3.24 yields

$$\frac{s_2}{s_1}=\frac{k_1}{k_3} \tag{5.3.25}$$

Since, $k_1^2+k_2^2+k_3^2=1=s_1^2=s_2^2=s_3^2$ we arrive at the rest of the identification suggested in Equation 5.3.7.

We shall refer to the formalism based on the parametrizations $\hat{k}(\phi,\theta,\psi)$ and $\hat{s}(\alpha,\beta,\gamma)$ as the $\hat{k}$ scheme and the $\hat{s}$-scheme respectively. Since either of the two pairs of angles ϕ,θ and α,β provide a satisfactory description of the polarization state, it is worthwhile to deal with both schemes.

The role of the “third angle” ψ or γ respectively, is more subtle. It is well known that a spinor can be visualized as a vector and an angle, a “flagpole” and a “flag” in Penrose’s terminology. However, the angle represents a phase, and as such has notoriously ambivalent properties. While a single phase is usually unimportant, phase relations are often most significant. Although one can solve particular problems in polarization optics in terms of the Poincaré sphere without an explicit use of the third angle, for us these problems are merely stepping stones for deeper problems and we prefer to present them as instances of the general formalism. No matter if this seems to be a somewhat heavy gun for the purpose.

In proceeding this way we have to ignore some fine distinctions; thus we assign $|\xi\rangle$ and $-|\xi\rangle 7$ to the same state of polarization. We consider it an advantage that the formalism has the reserve capacity to be used later to such problems as the electron spin.

We demonstrate the usefulness of the spinor formalism by translating one of its simple propositions into, what might be called the fundamental theorem of polarization optics.

Consider two pairs of conjugate spinors $|\xi\rangle, |\bar{\xi}\rangle$ and $|\xi'\rangle, |\bar{\xi}'\rangle$

Theorem 3. There is a uniquely determined unimodular unitary matrix U such that

$$\begin{gathered} |\xi'\rangle = U|\xi\rangle \\ |\bar{\xi}'\rangle = U|\bar{\xi}\rangle \end{gathered} \tag{5.3.26}$$

Proof. By using Equation 5.3.27of Section 5.1 we consider the unitary matrices associated with the spinor pairs:

$$V = (|\xi\rangle, |\bar{\xi}\rangle) \tag{5.3.27}$$

$$V' = \left(|\xi'\rangle, |\bar{\xi}'\rangle\right) \tag{5.3.28}$$

The matrix $U = V'V^{-1}$ has the desired properties, since $UV = V'$

Let the monoaxial parametrization of U be $U(\hat{u}, \chi/2)$. By using Equations 5.1.58 and 5.1.59 of Section 5.1 we see that U has two eigenspinors:

$$U|\hat{u}\rangle = \exp(-i\chi/2)|\hat{u}\rangle \tag{5.3.29}$$

$$U|\bar{\hat{u}}\rangle = \exp(i\chi/2)|\bar{\hat{u}}\rangle \tag{5.3.30}$$

Hence U produces a phase shift between the conjugate states $|\hat{u}\rangle$ and $|\bar{\hat{u}}\rangle$; moreover it rotates their linear combinations:

$$|\xi\rangle = a_0|\hat{u}\rangle + a_1|\bar{\hat{u}}\rangle \tag{5.3.31}$$

where

$$|a_0|^2 + |a_1|^2 = 1 \tag{5.3.32}$$

These results translate into polarization optics as follows. An arbitrary, fully polarized beam can be transformed into another beam of the same kind by a phase shifter, the axis $\hat{u}$ of which is uniquely determined in terms of the spinor representation of the given beams. Since the result is a rotation of the Poincaré sphere, the axis of the phase shifter can be determined also geometrically.

To counteract the complete generality of the Poincaré construction, let us consider the special cases

$$U = U\left(\hat{k}_3, \frac{\Delta\phi}{2}\right) \tag{5.3.33}$$

$$U = U\left(\hat{s}_3, \frac{\Delta\alpha}{2}\right) \tag{5.3.34}$$

The phase shifter, Equation 5.3.33, is called a linear retarder, it establishes a phase lag between one state of linear polarization and its antipodal state. For $\Delta\phi = \pi/2$ we have a quarter waveplate that transforms elliptic into linear polarization or vice versa.

The phase shifter, Equation 5.3.34, produces a phase lag between right and left circularly polarized beams. (A circularly biregringent crystal, say quartz is cut perpendicularly to the optic axis: spiral staircase effect.)

Since a linearly polarized beam is the linear composition $|R\rangle$ and $|L\rangle$ the phase lag manifests itself in a rotation of the plane of polarization, hence a rotation around $\hat{s}_3$. The device is called a rotator.

Thus rotations of the Poincare´ sphere may produce either a change of shape, or a change of orientation in ordinary space.

We may add that by combining two quarter waveplates with one general rotator we can realize an arbitrary phase shifter $U(\hat{u}, \chi/2)$.

Our main theorem on the representation of the transformation of fully polarized beams is evidently the counterpart of Euler's theorem on the displacements of the gyroscope mentioned on page 56.

Although we have a formal identity, in the sense that we have in both cases the rotation of a triad, there is a great deal of difference in the physical interpretation. The rotation takes place now in an abstract space, we may call it the Poincare´ space. Also it makes a great deal of difference that the angular velocities of the rotating object are now replaced by the time rates of change of the phase difference between pairs of conjugate polarizations. On going from rigid bodies to polarized waves (degenerate vibrations) we do not have to modify the formalism, but the new interpretation opens up new opportunities. The concept of phase difference evokes the idea of coherent superposition as contrasted with incoherent composition. These matters have no analog in the case of rigid rotation, and we now turn to the examination of the new features.

Let us consider a polarized beam represented in the $\hat{s}(\alpha, \beta)$ scheme by the spinor $|\hat{s}\rangle$ where

$$S=|\hat{s}\rangle\langle\hat{s}|=\frac{1}{2}(1+\hat{s}\cdot\vec{\sigma}) \tag{5.3.35}$$

or alternatively

$$S=\begin{pmatrix} s_0 s_0^* & s_0 s_1^* \\ s_1 s_0^* & s_0 s_1^* \end{pmatrix} \tag{5.3.36}$$

S is called the density matrix or coherency matrix associated with a polarized beam (a pure state in quantum mechanics). As we have seen already, it is idempotent and the determinant $|S|=0$

We analyze this beam with an instrument $U(u, \Delta\psi/2)$ where $\hat{u}\neq\hat{s}$ where $\hat{u}\neq\hat{s}$, and obtain

$$|\hat{s}\rangle=a_0|\hat{u}\rangle+a_1|\bar{\hat{u}}\rangle \tag{5.3.37}$$

with

$$a_0=\langle\hat{u}\mid\hat{s}\rangle \quad a_1=\langle\bar{\hat{u}}\mid\hat{s}\rangle \tag{5.3.38}$$

$$|a_0|^2+|a_1|^2=1 \tag{5.3.39}$$

From Equations 5.1.58 and 5.1.58 we have

$$\begin{aligned} |\xi'\rangle &= U\left(\hat{u}, \frac{\Delta\psi}{2}\right)|\hat{s}\rangle+ \\ &= a_0 \exp\left(-i\frac{\Delta\psi}{2}\right)|\hat{u}\rangle a_1 \exp\left(i\frac{\Delta\psi}{2}\right)|\bar{\hat{u}}\rangle \\ &= a_0\left|\hat{u}, \frac{\psi+\Delta\psi}{2}\right\rangle+a_1\left|\hat{u}, \frac{\psi+\Delta\psi}{2}\right\rangle \end{aligned} \tag{5.3.40}$$

Let us now assume that the instrument U is doubled up with a reverse instrument that reunites the two beams that have been separated in the first step. This reunification may happen after certain manipulations have been performed on the separated beams. Such a device, the so-called analyzing loop has been used more for the conceptual analysis of the quantum mechanical formalism than for the practical purposes of polarization optics.

Depending on the nature of the manipulations we have a number of distinct situations which we proceed to disentangle on hand of the following formulas.

We obtain from Equation 5.3.40

$$\frac{1}{2}S'=|\hat{s}'\rangle\left\langle\hat{s}'\right|=|a_0|^2\mid\hat{u}\rangle\left\langle\hat{u}\right|+|a_1|^2\mid\bar{\hat{u}}\rangle\left\langle\bar{\hat{u}}\right|+a_0 a_1^*\exp(-i\Delta\psi)|\hat{u}\rangle\langle\bar{\hat{u}}| \tag{5.3.41}$$

Here S' is idempotent and of determinant zero just as S is, since $|\hat{s}'\rangle$ arises out of $|\hat{s}\rangle$ by means of a unitary operation.

Let us consider now a different case in which the phase difference between the two partial beams has been randomized. In fact, take first the extreme case in which the interference terms vanish:

$$\left\langle a_0 a_1^*\exp(-i\Delta\psi)\right\rangle_{av}=\left\langle a_1 a_0^*\exp(i\Delta\psi)\right\rangle_{av}=0 \tag{5.3.42}$$

We obtain from Equations 5.3.41, 5.3.42and 5.3.39

$$S'=1+\left(|a_0|^2-|a_1|^2\right)\hat{u}\cdot\vec{\sigma} \tag{5.3.43}$$

We write S' as

$$S' = 1 + s'\hat{u} \cdot \vec{\sigma} \tag{5.3.44}$$

where $0 \leq s' < 1$ and

$$0 < |S| = 1 + s'^2 \leq 1 \tag{5.3.45}$$

We have now a generazized form of the density matrix associated with a partially polarized or even natural light (if $s' = 0$). In quantum mechanics we speak of a mixture of states.

It is usual in optics to change the normalization and set for a partially polarized beam

$$S = s_0 + s\hat{s} \cdot \vec{\sigma} \tag{5.3.46}$$

where s_0 is the total intensity and s the intensity of the polarized component. We have for the determinant

$$0 \leq |S| = s_0^2 - s^2 \leq s_0^2(31) \tag{5.3.47}$$

which is zero for polarized light and positive otherwise.

In addition to conserving or destroying phase relations, one may operate directly on the intensity as well. If one of the components of the analyser, say $|\hat{u}\rangle$ or $|\bar{\hat{u}}\rangle$ is blocked off, the instrument acts as a perfect polarizer.

Formally, we can let the projection operator $\frac{1}{2}(1 \pm \hat{u} \cdot \vec{\sigma})$ act on the density matrix of the beam, which may be polarized fully, partially, or not at all. Nonpolarized or natural light can be considered as a statistical ensemble of polarized light beams uniformly distributed over the Poincare´ sphere. (See problem #15.)

An imperfect polarizer (such as a sheet of polaroid) exhibits an unequal absorption of two conjugate linear polarizations. It can be represented as a Hermitian operator acting on S.

We have seen above that the inchoherent composition of the two beams of an analyzer is accounted, for by the addition of the density matrix.

Conversely, every partially polarized beam can be constructed in such a fashion. (See problem #13.)

Yet we may wish to add incoherently an arbitrary set of partially polarized beams, and this is always accomplished by adding their density matrices. The question arises then: Could we not operate phenomenologically in terms of the density matrices alone?

The matter was considered already by Stokes (1852) who introduced a column vector with the four components I, M, C, S corresponding to our $s_0, \vec{s}$. A general instrument is represented by a real 4×4 matrix. Note that the “instrument” might be also a molecule producing a change of polarization on scattering.

The $4x4$ matrices are commonly called Mueller matrices. This formalism is usually mentioned along with the Jones calculus of 2×2 complex matrices. This was developed by R. Clark Jones of the Polaroid Co. and his collaborators in a long series of papers in Journal of the American Optical Society in the 1940’s (quoted e.g., by Shurcliff, and C. Whitney). This is basically a two-component spinor theory to deal with instruments which modify the polarization without depolarization or loss of intensity. It was developed in close contact with experiment without reliance on an existing mathematical formalism.

Mueller liked to emphasize the purely phenomenological character of his formalism. The four Stokes parameters of a beam can be determined from measurements by four filters. However a difficulty of this phenomenological approach is that not every 4×4 matrix corresponds to a physically realizable instrument or scattering object. This means that a so-called passive instrument must neither increase total intensity nor create phase correlations. The situation is simpler in the 2×2 matrix formulation in which the redundant parameters have been eliminated.

However, we do not enter into such details, since polarization optics is not our primary concern. In fact, the two-valued ness of the full spinor formalism brings about a certain complication which is justified by the fact that our main interest is in the applications to quantum mechanics. We shall compare the different types of applications which are available at this juncture in Section 5.5.

Meanwhile in the next section we show that the concept of unitary spinor can be generalized to relativistic situations. This is indispensable if the formalism is to be applied also to the propagation rather than only the polarization of light.

5.4: Relativistic triads and spinors. A preliminary discussion

We have arrived at the concept of unitary spinors by searching for the proper parametrization of a Euclidean triad. We shall arrive at relativistic spinors by parametrising the relativistic triad. This is not a standard term, but it seems appropriate to so designate the configuration $\vec{E}, \vec{B}, \vec{k}$ (electric and magnetic fields, and the wave vector) in a monochromatic electromagnetic plane wave in vacuum.

The propagation of light is a dynamic problem and we are not ready to discuss it within the geometric-kinematic context of this chapter.

The purpose of this section is only to show that the formalism of unitary spinors developed thus far can be extended to relativistic situations with only a few indispensable adjustments.

It is a remarkable fact that the mutual orthogonality of the above mentioned vectors is a Lorentz invariant property. However, we have to abandon the unitary normalization since the length of the vectors is affected by inertial transformations.

Accordingly, we set up the relativistic analog of the Equations 5.1.40. We consider first

$$|\xi\rangle\langle\xi| = \frac{1}{2}\left(k_0 + \vec{k}\cdot\vec{\sigma}\right) = \frac{1}{2}K \tag{5.4.1}$$

$$|\bar{\xi}\rangle\langle\bar{\xi}| = \frac{1}{2}\left(k_0 - \vec{k}\cdot\vec{\sigma}\right) = \frac{1}{2}\bar{K} \tag{5.4.2}$$

with the unitary normalization changed to

$$\langle\xi \mid \xi\rangle = \langle\bar{\xi} \mid \bar{\xi}\rangle = k_0 \tag{5.4.3}$$

The Lorentz transformation properties of the spinors follow from that of K:

$$|\xi'\rangle = V|\xi\rangle \tag{5.4.4}$$

$$\left|\bar{\xi}'\right\rangle = \bar{V}|\bar{\xi}\rangle \tag{5.4.5}$$

$$\langle\xi'| = \langle\xi|V^\dagger \tag{5.4.6}$$

$$\left\langle\bar{\xi}'\right| = \langle\bar{\xi}|V^{-1} \tag{5.4.7}$$

If $V = U$ is unitary, we have $\bar{U} = U, U^\dagger = U^{-1}$

Let us define a second spinor by

$$|\eta\rangle\langle\eta| = \frac{1}{2}(r_0 + \vec{r}\cdot\vec{\sigma}) = \frac{1}{2}R \tag{5.4.8}$$

The relativistic invariant, (2.2.3a) appears now as

$$\begin{aligned}\frac{1}{2}\operatorname{Tr}(R\bar{K}) &= \frac{1}{2}\operatorname{Tr}(|\eta\rangle\langle\eta \mid \bar{\xi}\rangle\langle\bar{\xi}|) \\ &= \langle\bar{\xi} \mid \eta\rangle\langle\eta \mid \bar{\xi}\rangle = |\langle\bar{\xi} \mid \eta\rangle|^2\end{aligned} \tag{5.4.9}$$

It follows from Equations 5.4.4 and 5.4.7 that even the amplitude is invariant

$$\langle\bar{\xi} \mid \eta\rangle = \text{ invariant} \tag{5.4.10}$$

Explicitly it is equal to

$$(-\xi_1, \xi_0)\begin{pmatrix}\eta_0 \\ \eta_1\end{pmatrix} = \xi_0\eta_1 - \xi_1\eta_0 = \begin{vmatrix}\xi_0 & \eta_0 \\ \xi_1 & \eta_1\end{vmatrix} \tag{5.4.11}$$

We turn now to the last two of the Equations 5.1.40 and write by analogy

$$|\xi\rangle\langle\bar{\xi}| \sim (\vec{E} + i\vec{B})\cdot\vec{\sigma} = F \tag{5.4.12}$$

$$|\bar{\xi}\rangle\langle\xi| \sim (\vec{E} - i\vec{B})\cdot\vec{\sigma} = -\bar{F} = F^\dagger \tag{5.4.13}$$

We see that the field quantities have, in view of Equations 5.4.4– 5.4.7 the correct transformation properties.

The occurrence of the same spinor in Equations ref{1}, 5.4.2, 5.4.12 and 5.4.13 ensures the expected orthogonality properties of the triad.

However, in Equations 5.4.12 and 5.4.13 we write proportionality instead of equality, because we have to admit a different normalization for the four-vector and the six-vector respectively. We are not ready to discuss the matter at this point.

If in Equation 5.4.10 we choose the two spinors to be identical, the invariant vanishes:

$$\langle \bar{\xi} \mid \xi \rangle = 0 \tag{5.4.14}$$

The same is true of the invariant of the electromagnetic field:

$$\begin{aligned} \frac{1}{2}\operatorname{Tr}(F\tilde{F}) &= -\frac{1}{2}\operatorname{Tr} F^2 \simeq -\frac{1}{2}\operatorname{Tr}(|\xi\rangle\langle\bar{\xi} \mid \xi\rangle\langle\bar{\xi}|) \\ &= -(\langle\bar{\xi} \mid \xi\rangle)^2 = 0 \end{aligned} \tag{5.4.15}$$

Thus from a single spinor we can build up only constructs corresponding toa plane wave. We do not enter here into the discussion of more complicated situations and note only that we cannot use the device of taking linear combination of conjugate spinors in the usual form $a_0|\xi\rangle + a_1|\bar{\xi}\rangle$, because the two terms have contragradient Lorentz transformation properties. We write them, displaying their Lorentz transformations as

$$\begin{pmatrix} |\xi'\rangle \\ \left|\bar{\xi}'\right\rangle \end{pmatrix} = \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} |\xi\rangle \\ |\bar{\xi}\rangle \end{pmatrix} \tag{5.4.16}$$

We have arrived at spinors of the Dirac type. We return to their discussion later.

Let us conclude this section by considering the relation of the formalism to the standard formalism of van der Waerden. (See e.g., [MTW73])

The point of departure is Equation 5.4.4 applied to two spinors yielding the determinantal invariant of 5.4.11. The first characteristic aspect of the theory is the rule for raising the indices:

$$\eta_1 = \eta^0, \quad \eta_0 = -\eta^1 \tag{5.4.17}$$

Hence the invariant appears as

$$\xi_0\eta^0 + \xi_1\eta^1 \tag{5.4.18}$$

The motivation for writing the invariant in this form is to harmonize the presentation with the standard tensor formalism. In contrast, our expression 5.4.10, is an extension of the bra-ket formalism of nonrelativistic quantum mechanics, that is also quite natural for the linear algebra of complex vector spaces.

A second distinctive feature is connected with the method of complexification. Van der Waerden takes the complex conjugate of the matrix V by taking the complex conjugate of its elements, whereas we deal with the Hermitian conjugate $V^\dagger$ and the complex reflection $\bar{V}$.

From the practical point of view we tend to develop unitary and relativistic spinors in as united a form as objectively possible.

5.5: Review of SU(2) and preview of quantization

Our introduction of the spinor concept at the beginning of Section 5.1 can be rationalized on the basis of the following guidelines. First, we require a certain economy and wish to avoid dealing with redundant parameters in specifying a rotating triad, just as we have previously solved the analogous problem for the rotation operator. Second, we wish to have an efficient formalism to represent the rotational problem.

We have seen that the matrices of $\mathcal{SU}(2)$ satisfy all these requirements, but we find ourselves saddled with a two-valuedness of the representation: $|\xi\rangle$ and $|\xi\rangle$ correspond to the same triad configuration. This is not a serious trouble, since our key relations $5.1.36$ and $5.1.37$ are quadratic in $|\xi\rangle$. Thus the two-valuedness appears here only as a computational aid that disappears in the final result.

The situation is different if we look at the formulas 5.1.42–5.1.66 of the same section (Section 5.1). These equations are linear, they have a quantum mechanical character and we know that they are indeed applicable in the proper context. It is a pragmatic fact the two-valuedness is not just a necessary nuisance, but has physical meaning. But to understand this meaning is a challenge which we can meet only in carefully chosen steps.

We wish to give a more physical interpretation to the triad, but avoid the impasse of the rigid body. First we associate the abstract Poincaré space with the physical system of the two-dimensional degenerate oscillator. The rotation in the Poincare´ space is associated with a phase shift between conjugate states, which is translated into a rotation of the Poincaré sphere, interpreted in turn as a change in orientation, or change of shape of the vibrational patterns in ordinary space.

It is only a mild exaggeration to say that our transition from the triad in Euclidean space to that in Poincare´ space is something like a “quantization,” in the sense as Schrödinger’ ¨ s wave equation associates a wave with a particle. (Planck’s h is to enter shortly!)

In this theory we have a good use for the conjugate spinors $|\xi\rangle$ and $|\bar{\xi}\rangle$ representing opposite polarizations, but we must identify $|\bar{\xi}\rangle = -|\xi\rangle$ with $|\xi\rangle$.

The foregoing is still nothing but a perfectly well-defined kinematic model. The next step is different. As a “second quantization” we introduce Planck’s h to define single photons. A beam splitting represented by a projection operator can be expressed in probabilistic terms.

Formally all this is easy and we would at once have a great deal of the quantum mechanical formalism involving the theory of measurement.

Next, we could take the two-valuedness of the spinor seriously and obtain the formalism of isospin and of the neutrino, say as in Section 17, Fermion States, in [Kae65].

Finally, instead of doubly degenerate vibrations, we could consider the triply degenerate vibrator and handle it by $SU(3)$ $[\text{Lip02}]$. We shall not consider these generalizations at this point, however. Before further expanding the formalism we should hope to understand better what we already have.

First, a formal remark. Our results thus far developed are uniquely determined by the spinor formalism of Section 5.1 and by the program of considering the Poincaré sphere as the basic configuration space to be described by conventional spherical coordinates α, β or ϕ, θ.

It is noteworthy that this modest conceptual equipment carries us so far. We have obtained spinors, density matrices and have discussed at least fleetingly coherence, incoherence, quantum theory of measurement and transformation theory.

What we do not get out of the theory is a specific interpretation of the underlying vibrational process since the formalism is thus far entirely independent of it. This fact gives us some understanding of the scope and limit of quantum mechanics. We can apply the formalism to phenomena we understand very little. However, since the same §$U(2)$ formalism applies to polarized light, spin, isospin, strangeness and other phenomena, we learn little about their distinctive aspects.

In order to overcome this limitation we need a deeper understanding of what a quantized angular momentum is in the framework of a dynamical problem.

The next chapter is devoted to a phenomenological discussion of the concepts of particle and wave. We shall attempt to obtain sufficient hints for developing a dynamic theory in the form of a phase space geometry in Chapter VI.

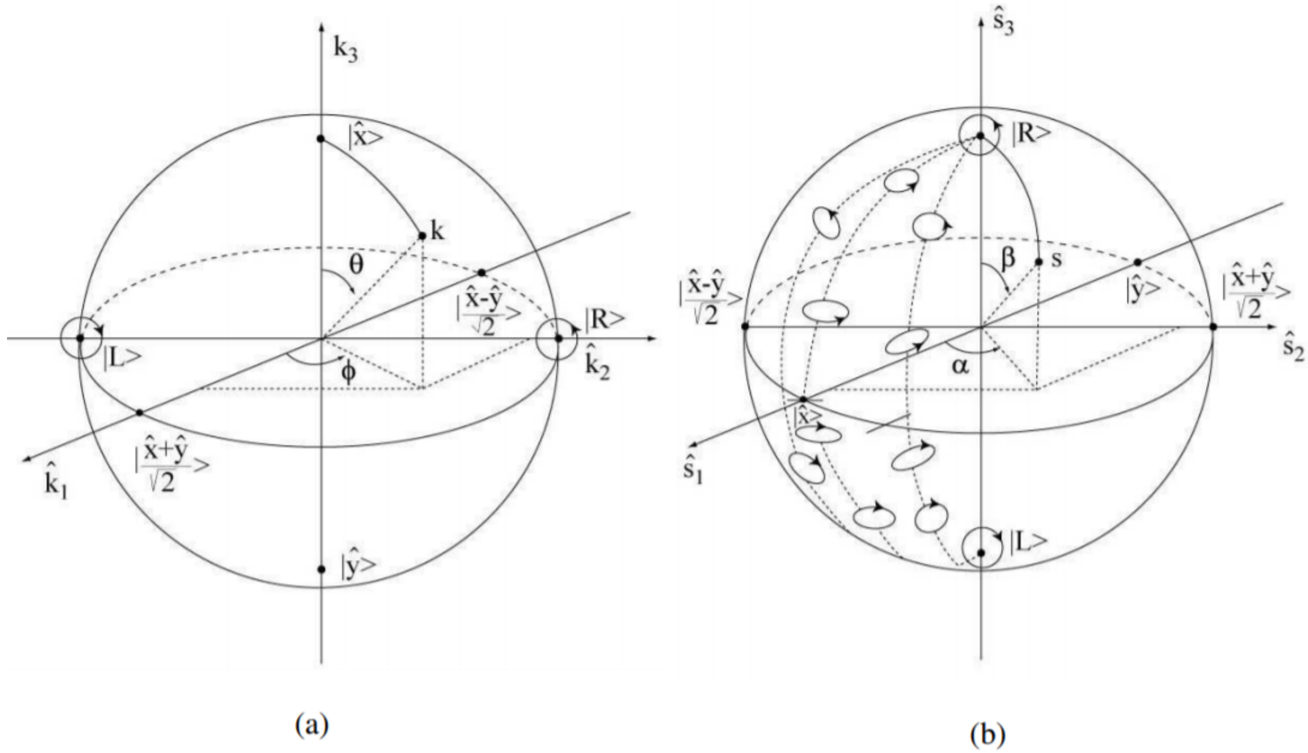

Figure 5.5.3: Representation of Polarization in the Poincaré Sphere. Connection between the schemes: (a) $\hat{k}(\phi, \theta)$ (b) $\hat{s}(\alpha, \beta)$ scheme

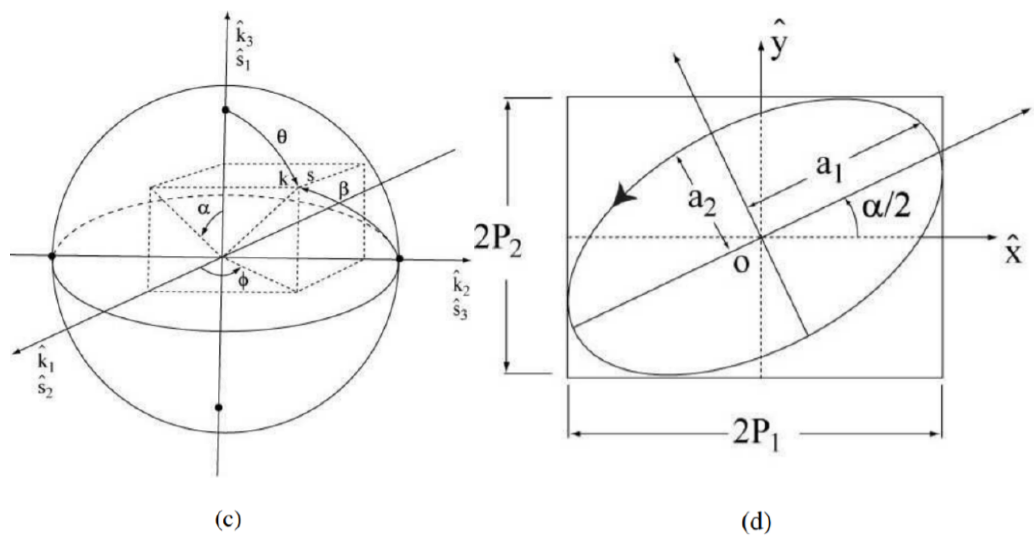

Figure 5.5.4: Representation of Polarization in the Poincare´ Sphere. Connection between the schemes (c) and (d).

CHAPTER OVERVIEW

6: Supplementary Material on the Pauli Algebra

6.1: Useful formulas

$A=a_0 1+\vec{a}\cdot\vec{\sigma}\tilde{A}=a_0 1-\vec{a}\cdot\vec{\sigma}A^{\dagger}=a_0^* 1+\vec{a}^*\cdot\vec{\sigma}\bar{A}=\tilde{A}^{\dagger}=a_0^* 1-\vec{a}^*\cdot\vec{\sigma}$

$$\frac{1}{2}\mathrm{Tr}(A)=a_0,\quad |A|=a_0^2-\vec{a}^2 1\frac{1}{2}\mathrm{Tr}(A\tilde{A}) \tag{6.1.1}$$

$$\frac{1}{2}\mathrm{Tr}(A\tilde{B})=a_0b_0-\vec{a}\cdot\vec{b} \tag{6.1.2}$$

$$A^{-1}=\frac{\tilde{A}}{|A|}\quad \text{for}\quad |A|=1: A^{-1}=\tilde{A} \tag{6.1.3}$$

$$(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma})=\vec{a}\cdot\vec{b}1+i(\vec{a}\times\vec{b})\cdot\vec{\sigma} \tag{6.1.4}$$

$$\text{For } \vec{a}\|\vec{b}\quad \frac{a_1}{b_1}=\frac{a_2}{b_2}=\frac{a_3}{b_3}\quad \vec{a}\times\vec{b}=0 \tag{6.1.5}$$

$$(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma})-(\vec{b}\cdot\vec{\sigma})(\vec{a}\cdot\vec{\sigma})=[(\vec{a}\cdot\vec{\sigma}),(\vec{b}\cdot\vec{\sigma})]=0 \tag{6.1.6}$$

$$\text{For}\quad A=a_0 1+\vec{a}\cdot\vec{\sigma},\quad B=b_0 1+\vec{b}\cdot\vec{\sigma} \tag{6.1.7}$$

$$[A,B]=0\quad \text{iff}\quad \vec{a}\|\vec{b} \tag{6.1.8}$$

$$\begin{aligned}\text{For}\qquad &\vec{a}\perp\vec{b},\quad \vec{a}\cdot\vec{b}\\ &\{\vec{a}\cdot\vec{\sigma},\vec{b}\cdot\vec{\sigma}\}\quad\equiv\quad(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma})+(\vec{b}\cdot\vec{\sigma})(\vec{a}\cdot\vec{\sigma})=0\end{aligned} \tag{6.1.9}$$

$$A(\vec{b}\cdot\vec{\sigma})=(\vec{b}\cdot\vec{\sigma})\tilde{A} \tag{6.1.10}$$

$$U=U\left(\hat{u},\frac{\phi}{2}\right)=\cos\frac{\phi}{2}1-i\sin\frac{\phi}{2}\hat{n}\cdot\vec{\sigma}=\exp\left(-i\frac{\phi}{2}\hat{n}\cdot\vec{\sigma}\right) \tag{6.1.11}$$

$$H=H\left(\hat{h},\frac{\mu}{2}\right)=\cosh\frac{\mu}{2}1+\sinh\frac{\mu}{2}\hat{h}\cdot\vec{\sigma}=\exp\left(\frac{\mu}{2}\hat{h}\cdot\vec{\sigma}\right) \tag{6.1.12}$$

U unitary unimodular, H Hermitian and positive.

6.2: Lorentz Invariance and Bilateral Multiplication

For Hermitian matrices: $K^\dagger = K, \bar{K} = \tilde{K}$ and the same for R. Why bilateral multiplication? To eliminate nonphysical factors indicated as $\underbrace{}$

$$\begin{pmatrix} e^{(\mu-i\phi)/2} & 0 \\ 0 & e^{-(\mu-i\phi)/2} \end{pmatrix} \begin{pmatrix} k_0+k_3 & k_1-ik_2 \\ k_1+ik_2 & k_2-k_3 \end{pmatrix} =$$

$$\begin{pmatrix} e^{\mu/2}(k_0+u_3)\underbrace{e^{-i\phi/2}} & \underbrace{e^{\mu/2}}(k_1-ik_2)e^{-i\phi/2} \\ \underbrace{e^{-\mu/2}}(k_1+ik_2)e^{-i\phi/2} & e^{\mu/2}(k_0-k_3)\underbrace{e^{i\phi/2}} \end{pmatrix} \times$$

$$\begin{pmatrix} k_0+k_3 & k_1-ik_2 \\ k_1+ik_2 & k_2-k_3 \end{pmatrix} \begin{pmatrix} e^{(\mu+i\phi)/2} & 0 \\ 0 & e^{-(\mu+i\phi)/2} \end{pmatrix} =$$

$$\begin{pmatrix} e^{\mu/2}(k_0+k_3)\underbrace{e^{i\phi/2}} & \underbrace{e^{-\mu/2}}(k_1-ik_2)e^{-i\phi/2} \\ \underbrace{e^{-\mu/2}}(k_1+ik_2)e^{i\phi/2} & e^{\mu/2}(k_0-k_3)\underbrace{e^{-i\phi/2}} \end{pmatrix} \times$$

$$\begin{pmatrix} e^{(\mu-i\phi)/2} & 0 \\ 0 & e^{-(\mu-i\phi)/2} \end{pmatrix} \begin{pmatrix} k_0+k_3 & k_1-ik_2 \\ k_1+ik_2 & k_2-k_3 \end{pmatrix} \begin{pmatrix} e^{(\mu+i\phi)/2} & 0 \\ 0 & e^{-(\mu+i\phi)/2} \end{pmatrix} =$$

$$\begin{pmatrix} e^{\mu/2}(k_0+k_3) & e^{-i\phi/2}(k_1-ik_2) \\ e^{i\phi/2}(k_1+ik_2) & e^{\mu/2}(k_0-k_3) \end{pmatrix}$$

Or, in 4×4 matrix form:

$$\begin{pmatrix} k_1' \\ k_2' \\ k_3' \\ k_0' \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 & 0 \\ \sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & \cosh\mu & \sinh\mu \\ 0 & 0 & \sinh\mu & \cosh\mu \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_0 \end{pmatrix} \tag{6.2.1}$$

Circular rotation around the z-axis by ϕ and hyperbolic rotation along the same asix by th ehyperbolic angle μ: Lorentz four-screw: $\mathcal{L}(\phi, \hat{z}, \mu)$. These transformations form an Abelian group.

In the Pauli algebra the formal simplicity of these relations is maintained even for arbitrary axial directions. To be sure, obtaining explicit results from the bilateral products may become cumbersome. However, the standard vectorial results can be easily extracted.

6.3: Typical Examples

Example 1

$$\begin{gathered} K' = HKH, \quad H = \exp\left(\frac{\mu}{2}\hat{h}\cdot\vec{\sigma}\right) \\ \vec{k} = \vec{k}_{\|} + \vec{k}_{\perp} \quad \vec{k}_{\|} = (\vec{k}\cdot\hat{h})\hat{h} \end{gathered} \tag{6.3.1}$$

By using (6a) and (7b):

$$\begin{gathered} \vec{k}_{\|}\cdot\vec{\sigma}H = H\vec{k}_{\|}\cdot\vec{\sigma}, \quad \vec{k}_{\perp}\cdot\vec{\sigma}H = H^{-1}\vec{k}_{\perp}\cdot\vec{\sigma} \\ \vec{k}'_{\|} = \vec{k}_{\|} = k\hat{h} \end{gathered} \tag{6.3.2}$$

$$\begin{aligned} \left(k'_0 + \vec{k}'_{\|}\cdot\vec{\sigma}\right) &= H^2\left(k_0 + \vec{k}_{\|}\cdot\vec{\sigma}\right) \\ &= (\cosh\mu + \sinh\mu\hat{h}\cdot\vec{\sigma})\left(k_0 + \vec{k}_{\|}\cdot\vec{\sigma}\right) \end{aligned} \tag{6.3.3}$$

$$\begin{aligned} k'_0 &= k_0\cosh\mu + k\sinh\mu \\ k' &= k_0\sinh\mu + k\cosh\mu \end{aligned} \tag{6.3.4}$$

Example 2

$$\begin{gathered} K' = UKU^{-1}, U = \exp\left(-i\frac{\phi}{2}\hat{u}\cdot\vec{\sigma}\right) \\ \vec{k} = \vec{k}_{\|} + \vec{k}_{\perp} \quad \vec{k}_{\|} = (\vec{k}\cdot\hat{u})\hat{u} \end{gathered} \tag{6.3.5}$$

$$\vec{k}_{\|}\cdot\vec{\sigma}U^{-1} = U^{-1}\vec{k}_{\|}\cdot\vec{\sigma}, \quad \vec{k}_{\perp}\cdot\vec{\sigma}U^{-1} = U\vec{k}_{\perp}\cdot\vec{\sigma} \tag{6.3.6}$$

$$\vec{k}'_{\|} = \vec{k}_{\|} \tag{6.3.7}$$

$$\begin{aligned} \vec{k}'_{\perp}\cdot\vec{\sigma} &= \left(\cos\frac{\phi}{2}1 - i\sin\frac{\phi}{2}\hat{u}\cdot\vec{\sigma}\right)^2\vec{k}_{\perp}\cdot\vec{\sigma} \\ &= (\cos\phi 1 - i\sin\phi\hat{u}\cdot\vec{\sigma})\vec{k}_{\perp}\cdot\vec{\sigma} \\ \vec{k}'_{\perp} &= \cos\phi\vec{k}_{\perp} + \sin\phi\hat{u}\times\vec{k}_{\perp} \end{aligned} \tag{6.3.8}$$

6.4: On the us of Involutions

The existence of the three involutions (see Equations A.1.1 above), provides a great deal of flexilbity. However, the most efficient use of these concepts calls for some care.

For any matrix of $\mathcal{A}_2$,

$$A^{-1}=\frac{\tilde{A}}{|A|} \quad |A|=\frac{1}{2}\operatorname{Tr}(A\tilde{A}) \tag{6.4.1}$$

In the case of Hermitian matrices we have two alternatives:

$$k_0r_0-\vec{k}\cdot\vec{r}=\frac{1}{2}\operatorname{Tr}(K\tilde{R}) \tag{6.4.2}$$

or

$$k_0r_0-\vec{k}\cdot\vec{r}=\frac{1}{2}\operatorname{Tr}(K\bar{R}) \tag{6.4.3}$$

It will appear, however from later discussions, that the complex reflection of Equation 6.4.3 is more appropriate to describe the transition from contravariant to covariant entities.

A case in point is the formal representation of the mirroring of a four-vector in a plane with the normal along $\hat{x}_1$. We have

$$\begin{aligned} K' &= \sigma_1\bar{K}\sigma_1=\sigma_1\left(k_0 1-k_1\sigma_1-k_2\sigma_2-k_3\sigma_3\right)\sigma_1 \\ &=\sigma_1^2\left(k_0 1-k_1\sigma_1+k_2\sigma_2+k_3\sigma_3\right) \\ &=k_0 1-k_1\sigma_1+k_2\sigma_2+k_3\sigma_3 \end{aligned} \tag{6.4.4}$$

More generally the mirroring in a plane with normal x is achieve by means of the operation

$$K'=\hat{a}\cdot\vec{\sigma}\bar{K}\hat{a}\cdot\vec{\sigma} \tag{6.4.5}$$

Again, we could have chosen $\tilde{K}$ instead of $\bar{K}$.

However, Eq (22) generalizes to the inversion of the electromagnetic six-vector $\vec{f}=\vec{E}+i\vec{B}$:

$$\left(\vec{E}'+i\vec{B}'\right)\cdot\vec{\sigma}=\overline{(\vec{E}+i\vec{B})\cdot\vec{\sigma}}=(-\vec{E}+i\vec{B})\cdot\vec{\sigma} \tag{6.4.6}$$

This relation takes into account the fact that $\vec{E}$ is a polar and $\vec{B}$ an axial vector.

6.5: On Parameterization and Integration

The explicit performance of the bilateral multiplication provides the connection between the parameters of rotation and the elements of the 4×4 matrices. We consider here only the pure rotation generated by

$$U = \exp\left(-i\frac{\phi}{2}\hat{u}\cdot\vec{\sigma}\right) \tag{6.5.1}$$

Let

$$l_0 = \cos\phi/2, \quad l_1 = \sin\phi/2\hat{u}_1 \tag{6.5.2}$$

$$l_2 = \sin\phi/2\hat{u}_3, \quad l_3 = \sin\phi/2\hat{u}_3 \tag{6.5.3}$$

$$u_1 = \cos(\hat{u}\cdot\hat{x}_1),\ldots, \text{ etc.} \tag{6.5.4}$$

$$u_1^2 + u_2^2 + u_3^2 = 1 \tag{6.5.5}$$

$$\begin{pmatrix} k_1' \\ k_2' \\ k_3' \end{pmatrix} = \begin{pmatrix} l_0^2+l_1^2-l_2^2-l_3^2 & 2\left(l_1l_2-l_0l_3\right) & 2\left(l_1l_3+l_0l_2\right) \\ 2\left(l_1l_2+l_0l_3\right) & l_0^2-l_1^2+l_2^2-l_3^2 & 2\left(l_2l_3-l_0l_1\right) \\ 2\left(l_1l_3-l_0l_3\right) & 2\left(l_2l_3+l_0l_1\right) & l_0^2-l_1^2-l_2^2+l_3^2 \end{pmatrix} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} \tag{6.5.6}$$

Such expression are, of course not very practical. One usually considers infinitesimal relations with the parameters $d\phi\mu_k$. Integration of the infinitesimal operations into those of the finite group can be achieved within the general theory of Lie groups and Lie algebras.

In our approach the integration is achieved by explicit construction for the special case of the restricted Lorentz group. This is the first step in our program of using group theory to supplement or replace method of differential equations.

CHAPTER OVERVIEW

7: Homework Assignments

7.1: Assignments 1–7

These are the homework assignents from the Spring 1977 version of the Physics 8.352 course. In the original notes the problems were numbered consecutively but given in separate assignments.

1. (a) Establish all the abstract groups having an order $2 \leq N \leq 6$. Compute typical products. Which groups are Abelian? Indicate at least two isomorphic realizations for each group.

(b) Identify the subgroups. Which are invariant?

2. Write down the permutations of n = 3 and n = 4 objects. Arrange the result in a compact fashion. Consider at first the subgroup of even permutations (the alternating group). Make use of cycles.

3. Find the joint effect of two mirror planes (see Figure B.1). Consider also parallel mirrors.

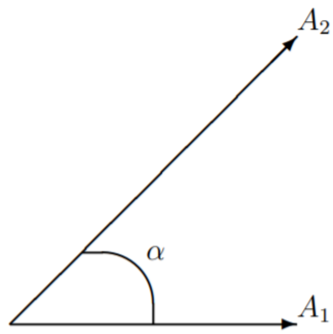

Figure B.1: Mirror Planes.

4. A spherical wave pulse diverges from the space-time point $(0,0,0,0)$ in the inertial frame Σ. Consider a frame Σ' moving along the z direction with the velocity $\beta = \tanh\mu$. The observer in Σ' sees also spherical wave fronts. However, the space-time points making up a surface $r' = ct' = \text{const}$ do not look synchronous, hence spherical in Σ. Show that the surfaces are ellipsoids of revolution with one common focus. Find the major and minor axes a, b, and the eccentricity in terms of r' and β. Find also the lengths of the perihelion and the aphelion. Use polar coordinates.

5. Consider the composition of rotations in the $\mathcal{SU}(2)$ formalism:

$$U'' = U'U$$

where $U = l_0 = -i\vec{l}\cdot\vec{\sigma}$, with

$$l_0 = \cos\frac{\phi}{2}, \quad \vec{l} = \sin\frac{\phi}{2}\hat{u}$$

(a) Express $\left\{l_0'', \vec{l}''\right\}$ in terms of $\left\{l_0', \vec{l}'\right\}$ and $\left\{l_0, \vec{l}\right\}$.

(b) Refer to the Rodriues-Hamilton theorem (Figure 2.1) and obtain the cosine law of spherical trigonometry.

(c) Obtain the sine law.

6. Check your general expressions by applying the special cases:

(a) $U'' = UU = U^2$

(b) $\begin{array}{l}\hat{u} = \frac{1}{\sqrt{3}}(1,1,1), \phi = \frac{2\pi}{3} \\ \hat{u}' = \frac{1}{\sqrt{3}}(1,0,0), \phi = \frac{\pi}{2}\end{array}$

Note that U and U' generate symmetry operations on the cube.

7. Consider the one-dimensional motion of a particle of rest mass m, under the influence of a force eE_z. At $t=0$ the particle is at rest. Show that the trajectory is represented in the z, ct plane as a hyperbola and find the semi-diameter. Develop the analogy with the cyclotron problem as far as you can. Discuss the significance of the approximation

$$\gamma^{-1} = \sqrt{1-\beta^2} \simeq 1$$

8. Consider an electromagnetic field

$\vec{f} = \vec{E} + i\vec{B}$

in a small space-time region. The Lorentz invariant of the field is:

$f^2 = E^2 - B^2 + 2iE \cdot B = I_1 + iI_2 = g^2 \exp(2i\psi)$

(a) Consider the case $f^2 \neq 0$. In this case, a canonical frame exists in which $E_{can} \| B_{can}$ and $\zeta = B_{\text{can}} / E_{\text{can}}$, the pitch, is a real number (which could be 0 or ∞) . Discuss the possible values of ζ according to the signs of I_1 and I_2. Summarize your conclusions in a table such as that shown in Table B.1.

$I_2 \mid I_1$	$+$	0	$-$
$+$			
0		$-$	
$-$			

Table B.1: Table for Problem 8

(b) Express E_{can}, B_{can}, ζ in terms of I_1, I_2 and g, ψ.

(c) Assume $\zeta \neq 0, \infty, \infty$. Take $\hat{x}$ along E_{can}. Consider a passive Lorentz transformation in the $\hat{z}$ direction, to a frame of velocity $v(\beta = v/c = \tanh \mu)$ with respect to the canonical frame. Find $\tan\theta_E, \tan\theta_B, \tan(\theta_E - \theta_B)$ in terms of β, ζ and also μ, ψ where θ_E and θ_B are the angles by which the electric and magnetic fields rotate under the Lorentz transformation, as shown in Figure B.2.

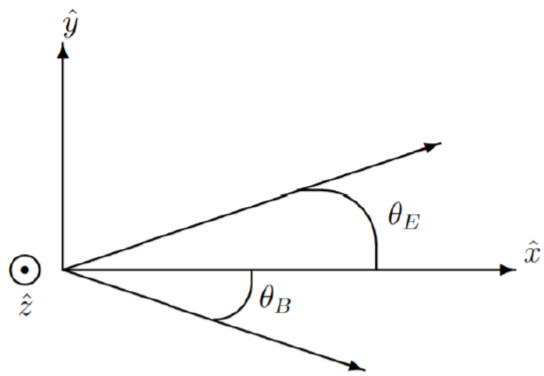

Figure B.2: Problem 8 coordinate frame and angles.

(d) Consider now the cases $\zeta = 0; \zeta = \infty$. Take $\hat{x}$ in the direction of the non-vanishing canonical field. Discuss the effect of a Lorentz transformation similar to that considered in (c). Give the ratio of the magnitudes of the electric and magnetic fields after the Lorentz transformation.

9. (a) Find the polar decomposition of the matrix

$$\begin{pmatrix} 1 & \delta \\ 0 & 1 \end{pmatrix}$$

Verify the relation (11b) on p. II-53. Consider the cases $\delta = 1$ and $\delta << 1$.

(b) Find

$\mathcal{P}_{\hat{a}}(\vec{p} \cdot \vec{\sigma})\mathcal{P}_{\hat{a}}$

where

$\mathcal{P}_{\hat{a}} = \frac{1}{2}(1 + \hat{a} \cdot \vec{\sigma})$

10. Verify Eq's (23) - (26) on II-42, 43.

11. Show that the field matrix $F = (\vec{E} + i\vec{B}) \cdot \vec{\sigma}$ can be derived from the matrix equivalent of the four-potential. What, if any, conditions are to be imposed on the latter?

12. (a) Express the reflection of a four-vector $K = k_0 1 + \vec{k} \cdot \vec{\sigma}$ in a moving plane. The normal of the plane is $\hat{a}$. Its velocity is $v = v\hat{a}$ with $v/c = \tanh \mu$. (Hint:transform to the rest frame of the mirror.)

(b) Show that the combination of two mirrors $\vec{v}_1 = v_1\hat{a}_1$, and $\vec{v}_2 = v_2\hat{a}_2$ yields a Lorentz transformation.

13. Verify the equivalence of Equations (4) and (5) in Section 4.2 by transforming each factor from space- to the body-frame.

14. Show that the relation

$$|\xi\rangle\langle\xi| = \frac{1}{2}(1+\hat{k}\cdot\vec{\sigma}) \tag{7.1.1}$$

can be obtained through stereographic projection. Hint: Project the sphere $k_1^2+k_2^2+k_3^2=1$ from the south pole to the equatorial plane interpreted as the complex z-plane. Express k_1, k_2, k_2 in terms of $z, z*$ and set $z=\xi_1/\xi_0$ with $|\xi_0|^2+|\xi_1|^2=1$.

15. Find the unitary matrix U that connects tvo given set of spinors with each other:

$$(|\eta\rangle, |\bar{\eta}\rangle) = (|\xi\rangle, |\bar{\xi}\rangle)U \tag{7.1.2}$$

Express first its elements, then its components in terms of $\xi_0, \xi_1, \eta_0, \eta_1$.

16. The Pauli algebra can be considered as a generalization of elementary vector algebra and the knowledge of the latter is helpful in matrix manipulation.

However, one can approach the problem also from the converse point of view and derive the vector relations through matrix operations. Define

$A=\vec{a}\cdot\vec{\sigma}, \quad B=\vec{b}\cdot\vec{\sigma}, \quad C=\vec{c}\cdot\vec{\sigma}$

and associate

$$\vec{a}\cdot\vec{b} \quad \text{with} \quad \frac{1}{2}\{A,B\} = \frac{1}{2}(AB+BA) \tag{7.1.3}$$

$$\vec{a}\times\vec{b} \text{ with } \frac{1}{2i}\{A,B\} = \frac{1}{2i}(AB-BA) \tag{7.1.4}$$

Consider the Jacobi identity

$$[[A,B],C]+[[B,C],A]+[[C,A],B]=0 \tag{7.1.5}$$

and the condition for associativity:

$$A(BC)-(AB)C=0 \tag{7.1.6}$$

(Equation 7.1.5 is easily verified for commutators. For its significance see [Hal74] .) Translate Equation 7.1.5 and 7.1.6 by means of Equations 7.1.3 and 7.1.4, and obtain the familiar relations for triple vector products.

17. Give explicit spinorial expressions for the following polarization forms: $|x\rangle$ (linear polarization along the x-axis) ; $|\theta/2\rangle$ (polarized at the angle θ/2 with the x-axis); $|R\rangle$ (right circularly polarized).

(a) Use the $\hat{\kappa}(\phi,\theta,\psi)$ scheme and assign $\phi=\psi=\theta=0$ to $|x\rangle=(1,0)$. Express $|\theta/2\rangle, |\theta/2\rangle, |R\rangle, |\bar{R}\rangle$ in terms of $|x\rangle$ and $|\bar{x}\rangle$.

(b) Use the $\hat{s}(\alpha,\beta,\gamma)$ scheme. Assign $\beta=0, \alpha=\gamma=\pi/2$ to $|R\rangle$. Express the above mentioned spinors in terms of $|R\rangle$ and $|\bar{R}\rangle$. Note that the results of (a) and (b) are consistent with each other.

18. Give the matrix representations of a quarter wave, plate, a half wave plate, a rotator and a plane polarizer in both the $\hat{k}$ and the $\hat{s}$ schemes.

19. (a) We know of an optical instrument only that it transforms $|R\rangle$ into $|\bar{R}\rangle$ and vice versa. Find the most general matrix operator consistent with this fact

(b) Sharpen this answer by using the additional information that the instrument passes a beam $|x\rangle$ unchanged. What is the name of this device?

20. Consider an arbitrary Hermitian 2×2 matrix: $S=s_0+\vec{s}\cdot\vec{\sigma}$ with $s_0^2-\vec{s}^2\neq 0$ in general.

(a) Show that it is possible to decompose S into a sum of two matrices with determinant zero. That is:

$S=K'+K''$

where

$$K' = k'_0 + \vec{k}' \cdot \vec{\sigma} \quad k'^2_0 - \vec{k}'^2 = 0$$

$$K'' = k''_0 + \vec{k}'' \cdot \vec{\sigma} \quad k''^2_0 - \vec{k}''_0 2 = 0$$

(b) Show that if one imposes:

$$\vec{k}' = k'\hat{k}$$

$$\vec{k}'' = k''\hat{k}$$

$$\vec{k}' \text{ and } \vec{k}'' \quad \text{parallel}$$

the decomposition becomes unique. Find $k'_0, k''_0, k', k'', \hat{k}$.

21. Consider an approximately monochromatic beam of unpolarized light, it has been suggested that such a beam be considered as a random sequence of elliptically polarized light, whereby the parameters of ellipticity α, β vary slowly compared to $1/\omega$ but fast compared to the time of observation (see [Hur45]). This author shows that the average ellipticity is given by the median value

$$\left(\frac{a_2}{a_1}\right)_m = \tan(15^\circ)$$

This result can be obtained very simply. Assume that all representative points of the Poincare´ sphere are equally probable. Consider the quantity:

$$S = \frac{2a_1a_2}{a_1^2 + a_2^2}$$

for an arbitrary point on the sphere. Take the average of $|S|$ over the Poincaré sphere, using the statistical assumption above. Deduce the value

$$\left(\frac{a_2}{a_1}\right)_0$$

corresponding to $\langle |S| \rangle$.

Index

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Book: Applied Geometric Algebra (Tisza)

Webpages: 42

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC-SA 4.0: 73.8% (31 pages)
- Undeclared: 26.2% (11 pages)

By Page

- Book: Applied Geometric Algebra (Tisza) - *CC BY-NC-SA 4.0*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Introduction - *CC BY-NC-SA 4.0*
    - 1.1: Introduction - *Undeclared*
  - 2: Algebraic Preliminaries - *CC BY-NC-SA 4.0*
    - 2.1: Groups - *CC BY-NC-SA 4.0*
    - 2.2: The geometry of the three-dimensional rotation group. The Rodrigues-Hamilton theorem - *CC BY-NC-SA 4.0*
    - 2.3: The n-dimensional vector space V(n) - *CC BY-NC-SA 4.0*
    - 2.4: How to multiply vectors? Heuristic considerations - *CC BY-NC-SA 4.0*
    - 2.5: A Short Survey of Linear Groups - *CC BY-NC-SA 4.0*
    - 2.6: The unimodular group SL(n, R) and the invariance of volume - *CC BY-NC-SA 4.0*
    - 2.7: On “alias” and “alibi”. The Object Group - *CC BY-NC-SA 4.0*
  - 3: The Lorentz Group and the Pauli Algebra - *CC BY-NC-SA 4.0*
    - 3.1: Introduction - *CC BY-NC-SA 4.0*
    - 3.2: The Corpuscular Aspects of Light - *CC BY-NC-SA 4.0*
    - 3.3: On Circular and Hyperbolic Rotations - *CC BY-NC-SA 4.0*
    - 3.4: The Pauli Algebra - *CC BY-NC-SA 4.0*
  - 4: Pauli Algebra and Electrodynamics - *CC BY-NC-SA 4.0*
    - 4.1: Lorentz transformation and Lorentz force - *CC BY-NC-SA 4.0*
    - 4.2: The Free Maxwell Field - *CC BY-NC-SA 4.0*
  - 5: Spinor Calculus - *CC BY-NC-SA 4.0*
    - 5.1: From Triads and Euler Angles to Spinors - A Heuristic Introduction - *CC BY-NC-SA 4.0*
    - 5.2: Rigid Body Rotation - *CC BY-NC-SA 4.0*
    - 5.3: Polarized Light - *CC BY-NC-SA 4.0*
    - 5.4: Relativistic triads and spinors. A preliminary discussion - *CC BY-NC-SA 4.0*
    - 5.5: Review of SU(2) and preview of quantization - *CC BY-NC-SA 4.0*
  - 6: Supplementary Material on the Pauli Algebra - *CC BY-NC-SA 4.0*
    - 6.1: Useful formulas - *CC BY-NC-SA 4.0*
    - 6.2: Lorentz Invariance and Bilateral Multiplication - *CC BY-NC-SA 4.0*
    - 6.3: Typical Examples - *CC BY-NC-SA 4.0*
    - 6.4: On the us of Involutions - *CC BY-NC-SA 4.0*
    - 6.5: On Parameterization and Integration - *CC BY-NC-SA 4.0*
  - 7: Homework Assignments - *CC BY-NC-SA 4.0*
    - 7.1: Assignments 1–7 - *Undeclared*
  - Back Matter - *Undeclared*
    - Index - *Undeclared*
    - Glossary - *Undeclared*
    - Detailed Licensing - *Undeclared*