2 Algebraic Preliminaries

2.1 Groups [1]

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 $A \in \mathcal{G}$.

[1] This outline serves mainly to delimit the extent of abstract group theory to be used later. Supplementary reading is recommended. See for instance [Tin64, Wig59, BM41].

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 $\mathcal{S}(n)$, 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 (four-group), 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..., let the numbers of distinct elements in the sequence be p. It is easy to show that $A^p = E$. The sequence

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

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 $E, H_2, ...H_h$; the set of elements

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

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} 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:

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

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:

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

a problem to which we shall return later.

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 [2].

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.

[2] See [Whi64]

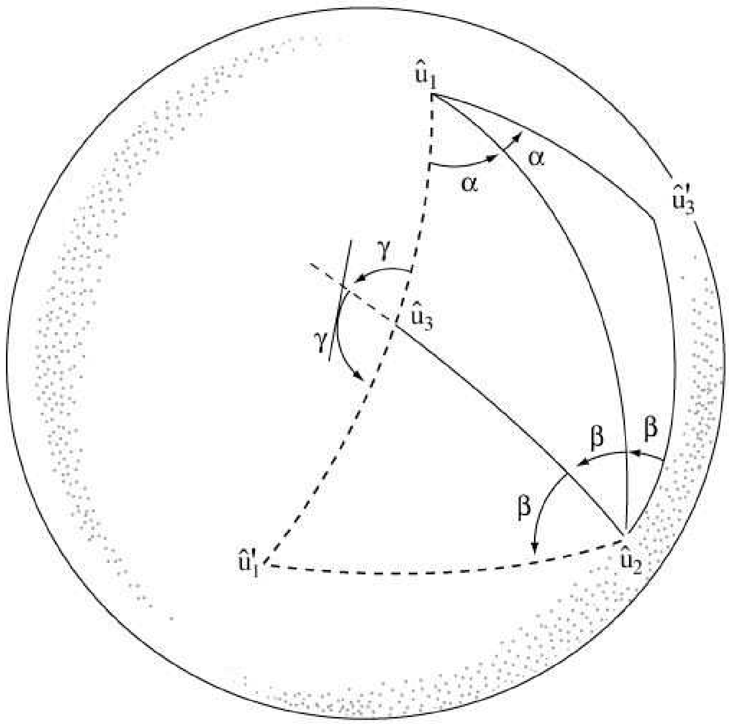

Figure 2.1: Composition of the Rotations of the Sphere. $\alpha = \phi_1/2$, $\beta = \phi_2/2$, $\gamma = \phi_3/2$.

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 [3].

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

$$\begin{aligned} \vec{x} &= \sum x^{i\prime} \left(\sum S_i^k \vec{e}_k \right) \\ &= \sum \left(\sum x^{i\prime} S_i^k \right) \vec{e}_k \end{aligned} \tag{2.3.7}$$

and hence in conjunction with Equation 2.3.5

$$x^k = \sum S_i^k x^{i\prime} \tag{2.3.8}$$

Note the characteristic "turning around" of the indices as we pass from Equation 2.3.4 to Equation 2.3.8 with a simultaneous interchange of the roles of the old and the new frame [4]. 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

$$X = \begin{pmatrix} x^1 \\ \vdots \\ x^k \end{pmatrix} \quad E = \begin{pmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_k \end{pmatrix} \tag{2.3.9}$$

Equation 2.3.6 appears then as a matrix product

$$\vec{x} = X^T E = X^T S^{-1} S E = X'^T E' \tag{2.3.10}$$

where the superscript stands for "transpose."

We ensure consistency by setting

$$E' = SE \tag{2.3.11}$$

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

$$X' = S^{-1T} X \tag{2.3.13}$$

Thus we arrive in a lucid fashion at the results contained in Equations 2.3.4 and 2.3.8. 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.

[3] These transformations form the general linear group $\mathcal{GL}(n, R)$, or $\mathcal{GL}(n, C)$

[4] See [Hal58], p. 66

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 algebras 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 [5].

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 four-dimensional 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:

$$x_1y_1 + x_2y_2 + \ldots + x_ny_n = X^TY \tag{2.5.1}$$

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

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

[5]The new vector calculus was not received with undivided enthusiasm. Professor Tait referred to it as "…a sort of hermaphrodite monster compounded of the notations of Hamilton and Grassmann." Quoted by Gibbs in Collected Works, Volume II, Part II, p 155.

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 $detO = |O| = \pm 1$. 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 $\mathcal{SO}(n)$ 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[6]. 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 $\mathcal{O}(k, l)$.

In this category the Lorentz group $\mathcal{SO}(3, 1)$ 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 $\mathcal{SO}(3, 1)$ 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 $\mathcal{SO}(3, 1)$ as for $\mathcal{SO}(4)$. 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 $\mathcal{SO}(3)$ 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.

[6] This chapter was not included in the Spring 1976 notes - Editor.

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

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

that can be brought to the canonical form

$$x_1 x_1^* + x_2 x_2^* + \ldots + x_n x_n^* = X^\dagger X \tag{2.5.4}$$

where $X^\dagger = X^{*T}$ is the Hermitian adjoint of X and the star stands for the conjugate complex. Expression 2.5.4 is invariant under transformations by matrices that satisfy the condition

$$U^\dagger U = \mathcal{I} \tag{2.5.5}$$

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 $\mathcal{SL}(n, R)$ 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[7].

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

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.

[7]See [Cox69], Section 13.4 Equiaffinites; [VY18], Vol. II, p. 105, 291; [Cur68], Chapter 5.

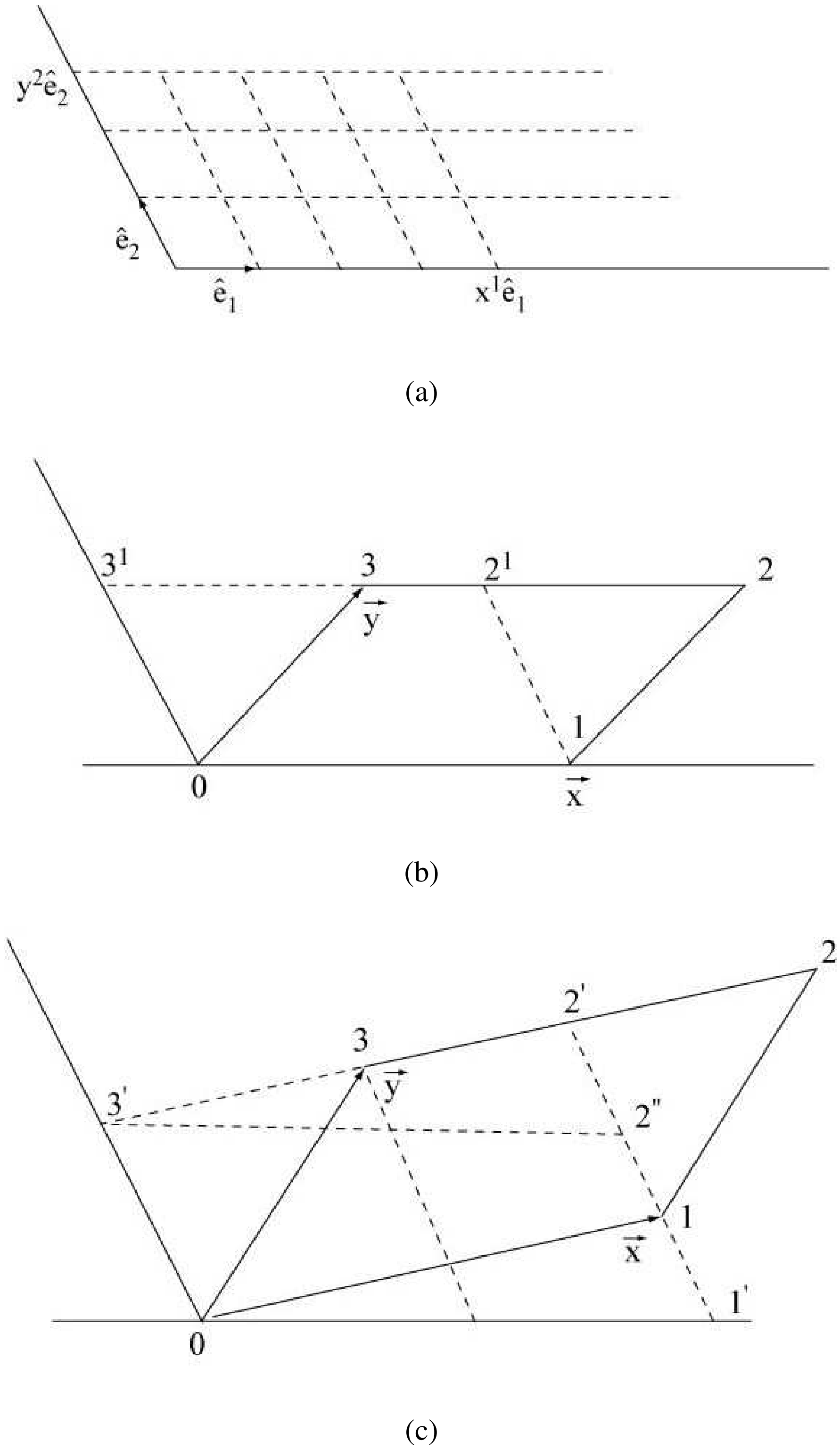

Figure 2.2: Translational congruence and equal area.

Consider at first the vectors

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

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:

$$x^1y^2 = \begin{vmatrix} x^1 & 0 \\ 0 & y^2 \end{vmatrix} \tag{2.6.1}$$

If the vectors

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

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

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

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[8].

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,

[8] This chapter was not included in the Spring 1976 notes - Editor.

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.