12

Elements of group theory

Group theory did not find a use in physics until the advent of modern quantum mechanics in 1925. In recent years group theory has been applied to many branches of physics and physical chemistry, notably to problems of molecules, atoms and atomic nuclei. Mostly recently, group theory has been being applied in the search for a pattern of 'family' relationships between elementary particles. Mathematicians are generally more interested in the abstract theory of groups, but the representation theory of groups of direct use in a large variety of physical problems is more useful to physicists. In this chapter, we shall give an elementary introduction to the theory of groups, which will be needed for understanding the representation theory.

Definition of a group (group axioms)

A group is a set of distinct elements for which a law of 'combination' is well defined. Hence, before we give 'group' a formal definition, we must first define what kind of 'elements' do we mean. Any collection of objects, quantities or operators form a set, and each individual object, quantity or operator is called an element of the set.

A group is a set of elements $A, B, C, \ldots$, finite or infinite in number, with a rule for combining any two of them to form a 'product', subject to the following four conditions:

(1) The product of any two group elements must be a group element; that is, if A and B are members of the group, then so is the product AB.
(2) The law of composition of the group elements is associative; that is, if A, B, and C are members of the group, then $(AB)C = A(BC)$.
(3) There exists a unit group element E, called the identity, such that $EA = AE = A$ for every member of the group.

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

(4) Every element has a unique inverse, A^{-1}, such that $AA^{-1} = A^{-1}A = E$.

The use of the word 'product' in the above definition requires comment. The law of combination is commonly referred as 'multiplication', and so the result of a combination of elements is referred to as a 'product'. However, the law of combination may be ordinary addition as in the group consisting of the set of all integers (positive, negative, and zero). Here $AB = A + B$, 'zero' is the identity, and $A^{-1} = (-A)$. The word 'product' is meant to symbolize a broad meaning of 'multiplication' in group theory, as will become clearer from the examples below.

A group with a finite number of elements is called a finite group; and the number of elements (in a finite group) is the order of the group.

A group containing an infinite number of elements is called an infinite group. An infinite group may be either discrete or continuous. If the number of the elements in an infinite group is denumerably infinite, the group is discrete; if the number of elements is non-denumerably infinite, the group is continuous.

A group is called Abelian (or commutative) if for every pair of elements A, B in the group, $AB = BA$. In general, groups are not Abelian and so it is necessary to preserve carefully the order of the factors in a group 'product'.

A subgroup is any subset of the elements of a group that by themselves satisfy the group axioms with the same law of combination.

Now let us consider some examples of groups.

Example 12.1

The real numbers 1 and -1 form a group of order two, under multiplication. The identity element is 1; and the inverse is $1/x$, where x stands for 1 or -1.

Example 12.2

The set of all integers (positive, negative, and zero) forms a discrete infinite group under addition. The identity element is zero; the inverse of each element is its negative. The group axioms are satisfied:

(1) is satisfied because the sum of any two integers (including any integer with itself) is always another integer.
(2) is satisfied because the associative law of addition $A + (B + C) = (A + B) + C$ is true for integers.
(3) is satisfied because the addition of 0 to any integer does not alter it.
(4) is satisfied because the addition of the inverse of an integer to the integer itself always gives 0, the identity element of our group: $A + (-A) = 0$.

Obviously, the group is Abelian since $A + B = B + A$. We denote this group by S_1.

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

The same set of all integers does not form a group under multiplication. Why? Because the inverses of integers are not integers and so they are not members of the set.

Example 12.3
The set of all rational numbers (p/q, with $q \neq 0$) forms a continuous infinite group under addition. It is an Abelian group, and we denote it by S_2. The identity element is 0; and the inverse of a given element is its negative.

Example 12.4
The set of all complex numbers ($z = x + iy$) forms an infinite group under addition. It is an Abelian group and we denote it by S_3. The identity element is 0; and the inverse of a given element is its negative (that is, $-z$ is the inverse of z).

The set of elements in S_1 is a subset of elements in S_2, and the set of elements in S_2 is a subset of elements in S_3. Furthermore, each of these sets forms a group under addition, thus S_1 is a subgroup of S_2, and S_2 a subgroup of S_3. Obviously S_1 is also a subgroup of S_3.

Example 12.5
The three matrices

$$\tilde{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$$

form an Abelian group of order three under matrix multiplication. The identity element is the unit matrix, $E = \tilde{A}$. The inverse of a given matrix is the inverse matrix of the given matrix:

$$\tilde{A}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \tilde{A}, \quad \tilde{B}^{-1} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} = \tilde{C}, \quad \tilde{C}^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} = \tilde{B}.$$

It is straightforward to check that all the four group axioms are satisfied. We leave this to the reader.

Example 12.6
The three permutation operations on three objects a, b, c

$$[1\ 2\ 3], [2\ 3\ 1], [3\ 1\ 2]$$

form an Abelian group of order three with sequential performance as the law of combination.

The operation [1 2 3] means we put the object a first, object b second, and object c third. And two elements are multiplied by performing first the operation on the right, then the operation on the left. For example

$$[2\ 3\ 1][3\ 1\ 2]abc = [2\ 3\ 1]cab = abc.$$

Thus two operations performed sequentially are equivalent to the operation [1 2 3]:

$$[2\ 3\ 1][3\ 1\ 2] = [1\ 2\ 3].$$

similarly

$$[3\ 1\ 2][2\ 3\ 1]abc = [3\ 1\ 2]bca = abc,$$

that is,

$$[3\ 1\ 2][2\ 3\ 1] = [1\ 2\ 3].$$

This law of combination is commutative. What is the identity element of this group? And the inverse of a given element? We leave the reader to answer these questions. The group illustrated by this example is known as a cyclic group of order 3, C_3.

It can be shown that the set of all permutations of three objects

$$[1\ 2\ 3],\ [2\ 3\ 1],\ [3\ 1\ 2],\ [1\ 3\ 2],\ [3\ 2\ 1],\ [2\ 1\ 3]$$

forms a non-Abelian group of order six denoted by S_3. It is called the symmetric group of three objects. Note that C_3 is a subgroup of S_3.

Cyclic groups

We now revisit the cyclic groups. The elements of a cyclic group can be expressed as power of a single element A, say, as $A, A^2, A^3, \ldots, A^{p-1}, A^p = E$; p is the smallest integer for which $A^p = E$ and is the order of the group. The inverse of A^k is A^{p-k}, that is, an element of the set. It is straightforward to check that all group axioms are satisfied. We leave this to the reader. It is obvious that cyclic groups are Abelian since $A^k A = AA^k$ $(k < p)$.

Example 12.7

The complex numbers 1, $i, -1, -i$ form a cyclic group of order 3. In this case, $A = i$ and $p = 3$: i^n, $n = 0, 1, 2, 3$. These group elements may be interpreted as successive 90° rotations in the complex plane $(0, \pi/2, \pi$, and $3\pi/2)$. Consequently, they can be represented by four 2×2 matrices. We shall come back to this later.

Example 12.8

We now consider a second example of cyclic groups: the group of rotations of an equilateral triangle in its plane about an axis passing through its center that brings

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

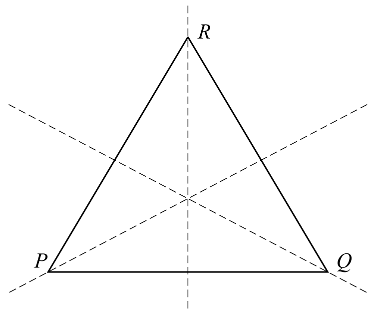

Figure 12.1.

it onto itself. This group contains three elements (see Fig. 12.1):

$E(=0^\circ)$: the identity; triangle is left alone;
$A(=120^\circ)$: the triangle is rotated through 120° counterclockwise, which sends P to Q, Q to R, and R to P;
$B(=240^\circ)$: the triangle is rotated through 240° counterclockwise, which sends P to R, R to Q, and Q to P;
$C(=360^\circ)$: the triangle is rotated through 360° counterclockwise, which sends P back to P, Q back to Q and R back to R.

Notice that $C = E$. Thus there are only three elements represented by E, A, and B. This set forms a group of order three under addition. The reader can check that all four group axioms are satisfied. It is also obvious that operation B is equivalent to performing operation A twice ($240^\circ = 120^\circ + 120^\circ$), and the operation C corresponds to performing A three times. Thus the elements of the group may be expressed as the power of the single element A as E, A, A^2, A^3 $(=E)$: that is, it is a cyclic group of order three, and is generated by the element A.

The cyclic group considered in Example 12.8 is a special case of groups of transformations (rotations, reflection, translations, permutations, etc.), the groups of particular interest to physicists. A transformation that leaves a physical system invariant is called a symmetry transformation of the system. The set of all symmetry transformations of a system is a group, as illustrated by this example.

Group multiplication table

A group of order n has n^2 products. Once the products of all ordered pairs of elements are specified the structure of a group is uniquely determined. It is sometimes convenient to arrange these products in a square array called a group multiplication table. Such a table is indicated schematically in Table 12.1. The element that appears at the intersection of the row labeled A and the column labeled B is the product AB, (in the table A^2 means AA, etc). It should be noted that all the

Table 12.1. Group multiplication table

	E	A	B	C	$\dots$
E	E	A	B	C	$\dots$
A	A	A^2	AB	AC	$\dots$
B	B	BA	B^2	BC	$\dots$
C	C	CA	CB	C^2	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Table 12.2.

	E	X	Y
E	E	X	Y
X^{-1}	X^{-1}	E	$X^{-1}Y$
Y^{-1}	Y^{-1}	$Y^{-1}X$	E

elements in each row or column of the group multiplication must be distinct: that is, each element appears once and only once in each row or column. This can be proved easily: if the same element appeared twice in a given row, the row labeled A say, then there would be two distinct elements C and D such that $AC = AD$. If we multiply the equation by A^{-1} on the left, then we would have $A^{-1}AC = A^{-1}AD$, or $EC = ED$. This cannot be true unless $C = D$, in contradiction to our hypothesis that C and D are distinct. Similarly, we can prove that all the elements in any column must be distinct.

As a simple practice, consider the group C_3 of Example 12.6 and label the elements as follows

$$[1\ 2\ 3] \to E, \quad [2\ 3\ 1] \to X, \quad [3\ 1\ 2] \to Y.$$

If we label the columns of the table with the elements E, X, Y and the rows with their respective inverses, E, X^{-1}, Y^{-1}, the group multiplication table then takes the form shown in Table 12.2.

Isomorphic groups

Two groups are isomorphic to each other if the elements of one group can be put in one-to-one correspondence with the elements of the other so that the corresponding elements multiply in the same way. Thus if the elements $A, B, C, \dots$ of the group G

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

Table 12.3.

	1	i	-1	i
1	1	i	-1	i
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

or

	E	A	B	C
E	E	A	B	C
A	A	B	C	E
B	B	C	E	A
C	C	E	A	B

correspond respectively to the elements $A', B', C', \ldots$ of G', then the equation $AB = C$ implies that $A'B' = C'$, etc., and vice versa. Two isomorphic groups have the same multiplication tables except for the labels attached to the group elements. Obviously, two isomorphic groups must have the same order.

Groups that are isomorphic and so have the same multiplication table are the same or identical, from an abstract point of view. That is why the concept of isomorphism is a key concept to physicists. Diverse groups of operators that act on diverse sets of objects have the same multiplication table; there is only one abstract group. This is where the value and beauty of the group theoretical method lie; the same abstract algebraic results may be applied in making predictions about a wide variety physical objects.

The isomorphism of groups is a special instance of homomorphism, which allows many-to-one correspondence.

Example 12.9

Consider the groups of Problems 12.2 and 12.4. The group G of Problem 12.2 consists of the four elements $E = 1, A = i, B = -1, C = -i$ with *ordinary multiplication* as the rule of combination. The group multiplication table has the form shown in Table 12.3. The group G' of Problem 12.4 consists of the following four elements, with *matrix multiplication* as the rule of combination

$$E' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B' = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad C' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

It is straightforward to check that the group multiplication table of group G' has the form of Table 12.4. Comparing Tables 12.3 and 12.4 we can see that they have precisely the same structure. The two groups are therefore isomorphic.

Example 12.10

We stated earlier that diverse groups of operators that act on diverse sets of objects have the same multiplication table; there is only one abstract group. To illustrate this, we consider, for simplicity, an abstract group of order two, G_2: that

Table 12.4.

	E'	A'	B'	C'
E'	E'	A'	B'	C'
A'	A'	B'	C'	E'
B'	B'	C'	E'	A'
C'	C'	E'	E'	B'

is, we make no *a priori* assumption about the significance of the two elements of our group. One of them must be the identity E, and we call the other X. Thus we have

$$E^2 = E, \quad EX = XE = E.$$

Since each element appears once and only once in each row and column, the group multiplication table takes the form:

	E	X
E	E	X
X	X	E

We next consider some groups of operators that are isomorphic to G_2. First, consider the following two transformations of three-dimensional space into itself:

(1) the transformation E', which leaves each point in its place, and
(2) the transformation R, which maps the point (x, y, z) into the point $(-x, -y, -z)$. Evidently, $R^2 = RR$ (the transformation R followed by R) will bring each point back to its original position. Thus we have $(E')^2 = E'$, $RE' = E'R = RE' = R$, $R^2 = E'$; and the group multiplication table has the same form as G_2: that is, the group formed by the set of the two operations E' and R is isomorphic to G_2.

We now associate with the two operations E' and R two operators $\hat{O}_{E'}$ and $\hat{O}_R$, which act on real- or complex-valued functions of the spatial coordinates (x, y, z), $\psi(x, y, z)$, with the following effects:

$$\hat{O}_{E'}\psi(x, y, z) = \psi(x, y, z), \quad \hat{O}_R\psi(x, y, z) = \psi(-x, -y, -z).$$

From these we see that

$$(\hat{O}_{E'})^2 = \hat{O}_{E'}, \qquad \hat{O}_{E'}\hat{O}_R = \hat{O}_R\hat{O}_{E'} = \hat{O}_R, \qquad (\hat{O}_R)^2 = \hat{O}_R.$$

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

Obviously these two operators form a group that is isomorphic to G_2. These two groups (formed by the elements E', R, and the elements $\hat{O}_{E'}$ and $\hat{O}_R$, respectively) are the two representations of the abstract group G_2. These two simple examples cannot illustrate the value and beauty of the group theoretical method, but they do serve to illustrate the key concept of isomorphism.

Group of permutations and Cayley's theorem

In Example 12.6 we examined briefly the group of permutations of three objects. We now come back to the general case of n objects $(1, 2, \ldots, n)$ placed in n boxes (or places) labeled α_1, $\alpha_2, \ldots, \alpha_n$. This group, denoted by S_n, is called the symmetric group on n objects. It is of order n! How do we know? The first object may be put in any of n boxes, and the second object may then be put in any of $n-1$ boxes, and so forth: $n(n-1)(n-2) \times \cdots \times 3 \times 2 \times 1 = n!$.

We now define, following common practice, a permutation symbol P

$$P = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \end{pmatrix}, \tag{12.1}$$

which shifts the object in box 1 to box α_1, the object in box 2 to box α_2, and so forth, where $\alpha_1 \alpha_2 \cdots \alpha_n$ is some arrangement of the numbers $1, 2, 3, \ldots, n$. The old notation in Example 12.6 can now be written as

$$[2\ 3\ 1] = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

For n objects there are $n!$ permutations or arrangements, each of which may be written in the form (12.1). Taking a specific example of three objects, we have

$$P_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$

$$P_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad P_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad P_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

For the product of two permutations $P_i P_j$ $(i, j = 1, 2, \ldots, 6)$, we first perform the one on the right, P_j, and then the one on the left, P_i. Thus

$$P_3 P_6 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = P_4.$$

To the reader who has difficulty seeing this result, let us explain. Consider the first column. We first perform P_6, so that 1 is replaced by 3, we then perform P_3 and 3 is replaced by 2. So by the combined action 1 is replaced by 2 and we have the first column

$$\begin{pmatrix} 1 & \cdots & \cdots \\ 2 & \cdots & \cdots \end{pmatrix}.$$

We leave the other two columns to be completed by the reader.

Each element of a group has an inverse. Thus, for each permutation P_i there is P_i^{-1}, the inverse of P_i. We can use the property $P_i P_i^{-1} = P_1$ to find P_i^{-1}. Let us find P_6^{-1}:

$$P_6^{-1} = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = P_2.$$

It is straightforward to check that

$$P_6 P_6^{-1} = P_6 P_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = P_1.$$

The reader can verify that our group S_3 is generated by the elements P_2 and P_3, while P_1 serves as the identity. This means that the other three distinct elements can be expressed as distinct multiplicative combinations of P_2 and P_3:

$$P_4 = P_2^2 P_3, \quad P_5 = P_2 P_3, \quad P_6 = P_2^2.$$

The symmetric group S_n plays an important role in the study of finite groups. Every finite group of order n is isomorphic to a subgroup of the permutation group S_n. This is known as Cayley's theorem. For a proof of this theorem the interested reader is referred to an advanced text on group theory.

In physics, these permutation groups are of considerable importance in the quantum mechanics of identical particles, where, if we interchange any two or more these particles, the resulting configuration is indistinguishable from the original one. Various quantities must be invariant under interchange or permutation of the particles. Details of the consequences of this invariant property may be found in most first-year graduate textbooks on quantum mechanics that cover the application of group theory to quantum mechanics.

Subgroups and cosets

A subset of a group G, which is itself a group, is called a subgroup of G. This idea was introduced earlier. And we also saw that C_3, a cyclic group of order 3, is a subgroup of S_3, a symmetric group of order 6. We note that the order of C_3 is a factor of the order of S_3. In fact, we will show that, in general,

> the order of a subgroup is a factor of the order of the full group (that is, the group from which the subgroup is derived).

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

This can be proved as follows. Let G be a group of order n with elements $g_1(=E)$, $g_2, \ldots, g_n$. Let H, of order m, be a subgroup of G with elements $h_1(=E)$, $h_2, \ldots, h_m$. Now form the set gh_k $(0 \leq k \leq m)$, where g is any element of G not in H. This collection of elements is called the left-coset of H with respect to g (the *left*-coset, because g is at the left of h_k).

If such an element g does not exist, then $H = G$, and the theorem holds trivially. If g does exist, than the elements gh_k are all different. Otherwise, we would have $gh_k = gh_\ell$, or $h_k = h_\ell$, which contradicts our assumption that H is a group. Moreover, the elements gh_k are not elements of H. Otherwise, $gh_k = h_j$, and we have

$$g = h_j/h_k.$$

This implies that g is an element of H, which contradicts our assumption that g does not belong to H.

This left-coset of H does not form a group because it does not contain the identity element ($g_1 = h_1 = E$). If it did form a group, it would require for some h_j such that $gh_j = E$ or, equivalently, $g = h_j^{-1}$. This requires g to be an element of H. Again this is contrary to assumption that g does not belong to H.

Now every element g in G but not in H belongs to some coset gH. Thus G is a union of H and a number of non-overlapping cosets, each having m different elements. The order of G is therefore divisible by m. This proves that the order of a subgroup is a factor of the order of the full group. The ratio n/m is the index of H in G.

It is straightforward to prove that a group of order p, where p is a prime number, has no subgroup. It could be a cyclic group generated by an element a of period p.

Conjugate classes and invariant subgroups

Another way of dividing a group into subsets is to use the concept of classes. Let a, b, and u be any three elements of a group, and if

$$b = u^{-1}au,$$

b is said to be the transform of a by the element u; a and b are conjugate (or equivalent) to each other. It is straightforward to prove that conjugate has the following three properties:

(1) Every element is conjugate with itself (reflexivity). Allowing u to be the identity element E, then we have $a = E^{-1}aE$.
(2) If a is conjugate to b, then b is conjugate to a (symmetry). If $a = u^{-1}bu$, then $b = uau^{-1} = (u^{-1})^{-1}a(u^{-1})$, where u^{-1} is an element of G if u is.

(3) If a is conjugate with both b and c, then b and c are conjugate with each other (transitivity). If $a = u^{-1}bu$ and $b = v^{-1}cv$, then $a = u^{-1}v^{-1}cvu = (vu)^{-1}c(vu)$, where u and v belong to G so that vu is also an element of G.

We now divide our group up into subsets, such that all elements in any subset are conjugate to each other. These subsets are called classes of our group.

Example 12.11
The symmetric group S_3 has the following six distinct elements:

$$P_1 = E,\ P_2,\ P_3,\ P_4 = P_2^2P_3,\ P_5 = P_2P_3,\ P_6 = P_2^2,$$

which can be separated into three conjugate classes:

$$\{P_1\}; \quad \{P_2, P_6\}; \quad \{P_3, P_4, P_5\}.$$

We now state some simple facts about classes without proofs:

(a) The identity element always forms a class by itself.
(b) Each element of an Abelian group forms a class by itself.
(c) All elements of a class have the same period.

Starting from a subgroup H of a group G, we can form a set of elements $uh^{-1}u$ for each u belong to G. This set of elements can be seen to be itself a group. It is a subgroup of G and is isomorphic to H. It is said to be a conjugate subgroup to H in G. It may happen, for some subgroup H, that for all u belonging to G, the sets H and uhu^{-1} are identical. H is then an invariant or self-conjugate subgroup of G.

Example 12.12
Let us revisit S_3 of Example 12.11, taking it as our group $G = S_3$. Consider the subgroup $H = C_3 = \{P_1, P_2, P_6\}$. The following relation holds

$$P_2\begin{pmatrix} P_1 \\ P_2 \\ P_6 \end{pmatrix}P_2^{-1} = P_1P_2\begin{pmatrix} P_1 \\ P_2 \\ P_6 \end{pmatrix}P_2^{-1}P_1^{-1} = (P_1P_2)\begin{pmatrix} P_1 \\ P_2 \\ P_6 \end{pmatrix}(P_1P_2)^{-1}$$
$$= P_1^2P_2\begin{pmatrix} P_1 \\ P_2 \\ P_6 \end{pmatrix}(P_1^2P_2)^{-1} = \begin{pmatrix} P_1 \\ P_6 \\ P_2 \end{pmatrix}.$$

Hence $H = C_3 = \{P_1, P_2, P_6\}$ is an invariant subgroup of S_3.

Group representations

In previous sections we have seen some examples of groups which are isomorphic with matrix groups. Physicists have found that the representation of group elements by matrices is a very powerful technique. It is beyond the scope of this text to make a full study of the representation of groups; in this section we shall make a brief study of this important subject of the matrix representations of groups.

If to every element of a group G, $g_1, g_2, g_3, \ldots$, we can associate a non-singular square matrix $D(g_1), D(g_2), D(g_3), \ldots$, in such a way that

$$g_i g_j = g_k \qquad \text{implies} \qquad D(g_i)D(g_j) = D(g_k), \tag{12.2}$$

then these matrices themselves form a group G', which is either isomorphic or homomorphic to G. The set of such non-singular square matrices is called a representation of group G. If the matrices are $n \times n$, we have an n-dimensional representation; that is, the order of the matrix is the dimension (or order) of the representation D_n. One trivial example of such a representation is the unit matrix associated with every element of the group. As shown in Example 12.9, the four matrices of Problem 12.4 form a two-dimensional representation of the group G of Problem 12.2.

If there is one-to-one correspondence between each element of G and the matrix representation group G', the two groups are isomorphic, and the representation is said to be faithful (or true). If one matrix D represents more than one group element of G, the group G is homomorphic to the matrix representation group G' and the representation is said to be unfaithful.

Now suppose a representation of a group G has been found which consists of matrices $D = D(g_1), D(g_2), D(g_3), \ldots, D(g_p)$, each matrix being of dimension n. We can form another representation D' by a similarity transformation

$$D'(g) = S^{-1}D(g)S, \tag{12.3}$$

S being a non-singular matrix, then

$$\begin{aligned} D'(g_i)D'(g_j) &= S^{-1}D(g_i)SS^{-1}D(g_j)S \\ &= S^{-1}D(g_i)D(g_j)S \\ &= S^{-1}D(g_i g_j)S \\ &= D'(g_i g_j). \end{aligned}$$

In general, representations related in this way by a similarity transformation are regarded as being equivalent. However, the forms of the individual matrices in the two equivalent representations will be quite different. With this freedom in the choice of the forms of the matrices it is important to look for some quantity that is an invariant for a given transformation. This is found in considering the traces of the matrices of the representation group because the trace of a matrix is invariant under a similarity transformation. It is often possible to bring, by a similarity transformation, each matrix in the representation group into a diagonal form

$$S^{-1}DS = \begin{pmatrix} D^{(1)} & 0 \\ 0 & D^{(2)} \end{pmatrix}, \tag{12.4}$$

where $D^{(1)}$ is of order $m, m < n$ and $D^{(2)}$ is of order $n - m$. Under these conditions, the original representation is said to be reducible to $D^{(1)}$ and $D^{(2)}$. We may write this result as

$$D = D^{(1)} \oplus D^{(2)} \tag{12.5}$$

and say that D has been decomposed into the two smaller representation $D^{(1)}$ and $D^{(2)}$; D is often called the direct sum of $D^{(1)}$ and $D^{(2)}$.

A representation $D(g)$ is called irreducible if it is not of the form (12.4) and cannot be put into this form by a similarity transformation. Irreducible representations are the simplest representations, all others may be built up from them, that is, they play the role of 'building blocks' for the study of group representation.

In general, a given group has many representations, and it is always possible to find a unitary representation – one whose matrices are unitary. Unitary matrices can be diagonalized, and the eigenvalues can serve for the description or classification of quantum states. Hence unitary representations play an especially important role in quantum mechanics.

The task of finding all the irreducible representations of a group is usually very laborious. Fortunately, for most physical applications, it is sufficient to know only the traces of the matrices forming the representation, for the trace of a matrix is invariant under a similarity transformation. Thus, the trace can be used to identify or characterize our representation, and so it is called the character in group theory. A further simplification is provided by the fact that the character of every element in a class is identical, since elements in the same class are related to each other by a similarity transformation. If we know all the characters of one element from every class of the group, we have all of the information concerning the group that is usually needed. Hence characters play an important part in the theory of group representations. However, this topic and others related to whether a given representation of a group can be reduced to one of smaller dimensions are beyond the scope of this book. There are several important theorems of representation theory, which we now state without proof.

(1) A matrix that commutes with all matrices of an irreducible representation of a group is a multiple of the unit matrix (perhaps null). That is, if matrix A commutes with $D(g)$ which is irreducible,

$$D(g)A = AD(g)$$

for all g in our group, then A is a multiple of the unit matrix.

(2) A representation of a group is irreducible if and only if the only matrices to commute with all matrices are multiple of the unit matrix.

Both theorems (1) and (2) are corollaries of Schur's lemma.

(3) Schur's lemma: Let $D^{(1)}$ and $D^{(2)}$ be two irreducible representations of (a group G) dimensionality n and n', if there exists a matrix A such that

$$AD^{(1)}(g) = D^{(2)}(g)A \qquad \text{for all } g \text{ in the group } G$$

then for $n \neq n'$, $A = 0$; for $n = n'$, either $A = 0$ or A is a non-singular matrix and $D^{(1)}$ and $D^{(2)}$ are equivalent representations under the similarity transformation generated by A.

(4) Orthogonality theorem: If G is a group of order h and $D^{(1)}$ and $D^{(2)}$ are any two inequivalent irreducible (unitary) representations, of dimensions d_1 and d_2, respectively, then

$$\sum_g [D^{(i)}_{\alpha\beta}(g)]^* D^{(j)}_{\gamma\delta}(g) = \frac{h}{d_1}\delta_{ij}\delta_{\alpha\gamma}\delta_{\beta\delta},$$

where $D^{(i)}(g)$ is a matrix, and $D^{(i)}_{\alpha\beta}(g)$ is a typical matrix element. The sum runs over all g in G.

Some special groups

Many physical systems possess symmetry properties that always lead to certain quantity being invariant. For example, translational symmetry (or spatial homogeneity) leads to the conservation of linear momentum for a closed system, and rotational symmetry (or isotropy of space) leads to the conservation of angular momentum. Group theory is most appropriate for the study of symmetry. In this section we consider the geometrical symmetries. This provides more illustrations of the group concepts and leads to some special groups.

Let us first review some symmetry operations. A plane of symmetry is a plane in the system such that each point on one side of the plane is the mirror image of a corresponding point on the other side. If the system takes up an identical position on rotation through a certain angle about an axis, that axis is called an axis of symmetry. A center of inversion is a point such that the system is invariant under the operation $\mathbf{r} \to -\mathbf{r}$, where $\mathbf{r}$ is the position vector of any point in the system referred to the inversion center. If the system takes up an identical position after a rotation followed by an inversion, the system possesses a rotation–inversion center.

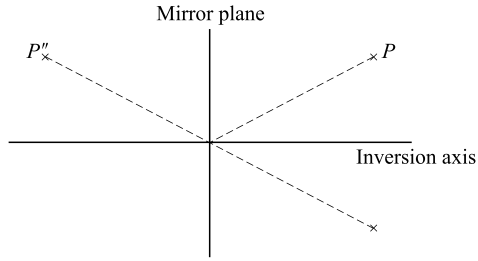

Figure 12.2.

Some symmetry operations are equivalent. As shown in Fig. 12.2, a two-fold inversion axis is equivalent to a mirror plane perpendicular to the axis.

There are two different ways of looking at a rotation, as shown in Fig. 12.3. According to the so-called active view, the system (the body) undergoes a rotation through an angle θ, say, in the clockwise direction about the x_3-axis. In the passive view, this is equivalent to a rotation of the coordinate system through the same angle but in the counterclockwise sense. The relation between the new and old coordinates of any point in the body is the same in both cases:

$$\left.\begin{aligned} x_1' &= x_1 \cos\theta + x_2 \sin\theta, \\ x_2' &= -x_1 \sin\theta + x_2 \cos\theta, \\ x_3' &= x_3, \end{aligned}\right\} \tag{12.6}$$

where the prime quantities represent the new coordinates.

A general rotation, reflection, or inversion can be represented by a linear transformation of the form

$$\left.\begin{aligned} x_1' &= \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3, \\ x_2' &= \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3, \\ x_3' &= \alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3. \end{aligned}\right\} \tag{12.7}$$

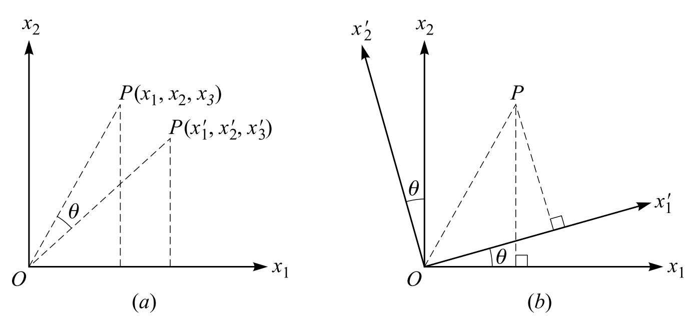

Figure 12.3. (a) Active view of rotation; (b) passive view of rotation.

Equation (12.7) can be written in matrix form

$$\tilde{x}' = \tilde{\lambda}\tilde{x} \tag{12.8}$$

with

$$\tilde{\lambda} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix}, \quad \tilde{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \tilde{x}' = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}.$$

The matrix $\tilde{\lambda}$ is an orthogonal matrix and the value of its determinant is ± 1. The '-1' value corresponds to an operation involving an odd number of reflections. For Eq. (12.6) the matrix $\tilde{\lambda}$ has the form

$$\tilde{\lambda} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{12.6a}$$

For a rotation, an inversion about an axis, or a reflection in a plane through the origin, the distance of a point from the origin remains unchanged:

$$r^2 = x_1^2 + x_2^2 + x_3^2 = x_1'^2 + x_2'^2 + x_3'^2. \tag{12.9}$$

The symmetry group D_2, D_3

Let us now examine two simple examples of symmetry and groups. The first one is on twofold symmetry axes. Our system consists of six particles: two identical particles A located at $\pm a$ on the x-axis, two particles, B at $\pm b$ on the y-axis, and two particles C at $\pm c$ on the z-axis. These particles could be the atoms of a molecule or part of a crystal. Each axis is a twofold symmetry axis. Clearly, the identity or unit operator (no rotation) will leave the system unchanged. What rotations can be carried out that will leave our system invariant? A certain combination of rotations of π radians about the three coordinate axes will do it. The orthogonal matrices that represent rotations about the three coordinate axes can be set up in a similar manner as was done for Eq. (12.6a), and they are

$$\tilde{\alpha}(\pi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \tilde{\beta}(\pi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \tilde{\gamma}(\pi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where $\tilde{\alpha}$ is the rotational matrix about the x-axis, and $\tilde{\beta}$ and $\tilde{\gamma}$ are the rotational matrices about y- and z-axes, respectively. Of course, the identity operator is a unit matrix

$$\tilde{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

These four elements form an Abelian group with the group multiplication table shown in Table 12.5. It is easy to check this group table by matrix multiplication. Or you can check it by analyzing the operations themselves, a tedious task. This demonstrates the power of mathematics: when the system becomes too complex for a direct physical interpretation, the usefulness of mathematics shows.

Table 12.5.

	$\tilde{E}$	$\tilde{\alpha}$	$\tilde{\beta}$	$\tilde{\gamma}$
$\tilde{E}$	$\tilde{E}$	$\tilde{\alpha}$	$\tilde{\beta}$	$\tilde{\gamma}$
$\tilde{\alpha}$	$\tilde{\alpha}$	$\tilde{E}$	$\tilde{\gamma}$	$\tilde{\beta}$
$\tilde{\beta}$	$\tilde{\beta}$	$\tilde{\gamma}$	$\tilde{E}$	$\tilde{\alpha}$
$\tilde{\gamma}$	$\tilde{\gamma}$	$\tilde{\beta}$	$\tilde{\alpha}$	$\tilde{E}$

This symmetry group is usually labeled D_2, a dihedral group with a twofold symmetry axis. A dihedral group D_n with an n-fold symmetry axis has n axes with an angular separation of $2\pi/n$ radians and is very useful in crystallographic study.

We next consider an example of threefold symmetry axes. To this end, let us revisit Example 12.8. Rotations of the triangle of 0°, 120°, 240°, and 360° leave the triangle invariant. Rotation of the triangle of 0° means no rotation, the triangle is left unchanged; this is represented by a unit matrix (the identity element). The other two orthogonal rotational matrices can be set up easily:

$$\tilde{A} = R_z(120^\circ) = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix},$$

$$\tilde{B} = R_z(240^\circ) = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{pmatrix},$$

and

$$\tilde{E} = R_z(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We notice that $\tilde{C} = R_z(360^\circ) = \tilde{E}$. The set of the three elements $(\tilde{E}, \tilde{A}, \tilde{B})$ forms a cyclic group C_3 with the group multiplication table shown in Table 12.6. The z-

Table 12.6.

	$\tilde{E}$	$\tilde{A}$	$\tilde{B}$
$\tilde{E}$	$\tilde{E}$	$\tilde{A}$	$\tilde{B}$
$\tilde{A}$	$\tilde{A}$	$\tilde{B}$	$\tilde{E}$
B	$\tilde{B}$	$\tilde{E}$	$\tilde{A}$

axis is a threefold symmetry axis. There are three additional axes of symmetry in the xy plane: each corner and the geometric center O defining an axis; each of these is a twofold symmetry axis (Fig. 12.4). Now let us consider reflection operations. The following successive operations will bring the equilateral angle onto itself (that is, be invariant):

- $\tilde{E}$ the identity; triangle is left unchanged;
- $\tilde{A}$ triangle is rotated through 120° clockwise;
- $\tilde{B}$ triangle is rotated through 240° clockwise;
- $\tilde{C}$ triangle is reflected about axis OR (or the y-axis);
- $\tilde{D}$ triangle is reflected about axis OQ;
- $\tilde{F}$ triangle is reflected about axis OP.

Now the reflection about axis OR is just a rotation of 180° about axis OR, thus

$$\tilde{C} = R_{OR}(180^\circ) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Next, we notice that reflection about axis OQ is equivalent to a rotation of 240° about the z-axis followed by a reflection of the x-axis (Fig. 12.5):

$$\tilde{D} = R_{OQ}(180^\circ) = \tilde{C}\tilde{B} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{pmatrix} = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{pmatrix}.$$

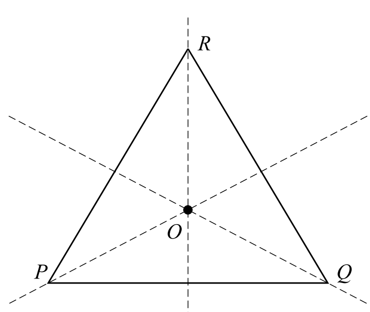

Figure 12.4.

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

Table 12.7.

	$\tilde{E}$	$\tilde{A}$	$\tilde{B}$	$\tilde{C}$	$\tilde{D}$	$\tilde{F}$
$\tilde{E}$	$\tilde{E}$	$\tilde{A}$	$\tilde{B}$	$\tilde{C}$	$\tilde{D}$	$\tilde{F}$
$\tilde{A}$	$\tilde{A}$	$\tilde{B}$	$\tilde{E}$	$\tilde{D}$	$\tilde{F}$	$\tilde{C}$
$\tilde{B}$	$\tilde{B}$	$\tilde{E}$	$\tilde{A}$	$\tilde{F}$	$\tilde{C}$	$\tilde{D}$
$\tilde{C}$	$\tilde{C}$	$\tilde{F}$	$\tilde{D}$	$\tilde{E}$	$\tilde{B}$	$\tilde{A}$
$\tilde{D}$	$\tilde{D}$	$\tilde{C}$	$\tilde{F}$	$\tilde{A}$	$\tilde{E}$	$\tilde{B}$
$\tilde{F}$	$\tilde{F}$	$\tilde{D}$	$\tilde{C}$	$\tilde{B}$	$\tilde{A}$	$\tilde{E}$

Similarly, reflection about axis OP is equivalent to a rotation of 180° followed by a reflection of the x-axis:

$$\tilde{F} = R_{OP}(180^\circ) = \tilde{C}\tilde{A} = \begin{pmatrix} 1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}.$$

The group multiplication table is shown in Table 12.7. We have constructed a six-element non-Abelian group and a 2×2 irreducible matrix representation of it. Our group is known as D_3 in crystallography, the dihedral group with a threefold axis of symmetry.

One-dimensional unitrary group $U(1)$

We now consider groups with an infinite number of elements. The group element will contain one or more parameters that vary continuously over some range so they are also known as continuous groups. In Example 12.7, we saw that the complex numbers $(1, i, -1, -i)$ form a cyclic group of order 3. These group elements may be interpreted as successive 90° rotations in the complex plane $(0, \pi/2, \pi, 3\pi/2)$, and so they may be written as $e^{i\varphi}$ with $\varphi = 0, \pi/2, \pi, 3\pi/2$. If φ is allowed to vary continuously over the range $[0, 2\pi]$, then we will have, instead of a four-member cyclic group, a continuous group with multiplication for the composition rule. It is straightforward to check that the four group axioms are all

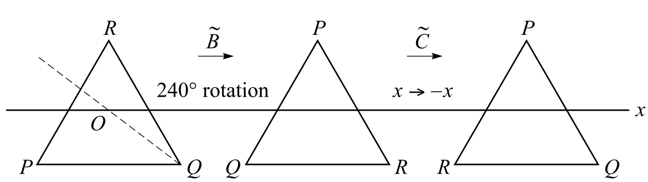

Figure 12.5.

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

met. In quantum mechanics, $e^{i\varphi}$ is a complex phase factor of a wave function, which we denote by $U(\varphi)$. Obviously, $U(0)$ is an identity element. Next,

$$U(\varphi)U(\varphi') = e^{i(\varphi+\varphi')} = U(\varphi+\varphi'),$$

and $U(\varphi+\varphi')$ is an element of the group. There is an inverse: $U^{-1}(\varphi) = U(-\varphi)$, since

$$U(\varphi)U(-\varphi) = U(-\varphi)U(\varphi) = U(0) = E$$

for any φ. The associative law is satisfied:

$$[U(\varphi_1)U(\varphi_2)]U(\varphi_3) = e^{i(\varphi_1+\varphi_2)}e^{i\varphi_3} = e^{i(\varphi_1+\varphi_2+\varphi_3)} = e^{i\varphi_1}e^{i(\varphi_2+\varphi_3)}$$
$$= U(\varphi_1)[U(\varphi_2)U(\varphi_3)].$$

This group is a one-dimensional unitary group; it is called $U(1)$. Each element is characterized by a continuous parameter φ, $0 \le \varphi \le 2\pi$; φ can take on an infinite number of values. Moreover, the elements are differentiable:

$$dU = U(\varphi + d\varphi) - U(\varphi) = e^{i(\varphi+d\varphi)} - e^{i\varphi}$$
$$= e^{i\varphi}(1 + id\varphi) - e^{i\varphi} = ie^{i\varphi}d\varphi = iUd\varphi$$

or

$$dU/d\varphi = iU.$$

Infinite groups whose elements are differentiable functions of their parameters are called Lie groups. The differentiability of group elements allows us to develop the concept of the generator. Furthermore, instead of studying the whole group, we can study the group elements in the neighborhood of the identity element. Thus Lie groups are of particular interest. Let us take a brief look at a few more Lie groups.

Orthogonal groups $SO(2)$ *and* $SO(3)$

The rotations in an n-dimensional Euclidean space form a group, called $O(n)$. The group elements can be represented by $n \times n$ orthogonal matrices, each with $n(n-1)/2$ independent elements (Problem 12.12). If the determinant of O is set to be $+1$ (rotation only, no reflection), then the group is often labeled $SO(n)$. The label O_n^+ is also often used.

The elements of $SO(2)$ are familiar; they are the rotations in a plane, say the xy plane:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \tilde{R}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

This group has one parameter: the angle θ. As we stated earlier, groups enter physics because we can carry out transformations on physical systems and the physical systems often are invariant under the transformations. Here $x^2 + y^2$ is left invariant.

We now introduce the concept of a generator and show that rotations of $SO(2)$ are generated by a special 2×2 matrix $\tilde{\sigma}_2$, where

$$\tilde{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Using the Euler identity, $e^{i\theta} = \cos\theta + i \sin\theta$, we can express the 2×2 rotation matrices $R(\theta)$ in exponential form:

$$\tilde{R}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \tilde{I}_2 \cos\theta + i\tilde{\sigma}_2 \sin\theta = e^{i\theta\tilde{\sigma}_2},$$

where $\tilde{I}_2$ is a 2×2 unit matrix. From the exponential form we see that multiplication is equivalent to addition of the arguments. The rotations close to the identity element have small angles $\theta \cong 0$. We call $\tilde{\sigma}_2$ the generator of rotations for $SO(2)$.

It has been shown that any element g of a Lie group can be written in the form

$$g(\theta_1, \theta_2, \ldots, \theta_n) = \exp\left(\sum_{i=1} i\theta_i F_i\right).$$

For n parameters there are n of the quantities F_i, and they are called the generators of the Lie group.

Note that we can get $\tilde{\sigma}_2$ from the rotation matrix $\tilde{R}(\theta)$ by differentiation at the identity of $SO(2)$, that is, $\theta \cong 0$. This suggests that we may find the generators of other groups in a similar manner.

For $n = 3$ there are three independent parameters, and the set of 3×3 orthogonal matrices with determinant $+1$ also forms a group, the $SO(3)$, its general member may be expressed in terms of the Euler angle rotation

$$R(\alpha, \beta, \gamma) = R_{z'}(0, 0, \alpha) R_y(0, \beta, 0) R_z(0, 0, \gamma),$$

where R_z is a rotation about the z-axis by an angle γ, R_y a rotation about the y-axis by an angle β, and $R_{z'}$ a rotation about the z'-axis (the new z-axis) by an angle α. This sequence can perform a general rotation. The separate rotations can be written as

$$\tilde{R}_y(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}, \quad \tilde{R}_z(\gamma) = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & o & 1 \end{pmatrix},$$

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

$$\tilde{R}_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}.$$

The $SO(3)$ rotations leave $x^2 + y^2 + z^2$ invariant.

The rotations $R_z(\gamma)$ form a group, called the group R_z, which is an Abelian subgroup of $SO(3)$. To find the generator of this group, let us take the following differentiation

$$-i\, d\tilde{R}_z(\gamma)/d\gamma\big|_{\gamma=0} = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \equiv \tilde{S}_z,$$

where the insertion of i is to make $\tilde{S}_z$ Hermitian. The rotation $R_z(\delta\gamma)$ through an infinitesimal angle $\delta\gamma$ can be written in terms of $\tilde{S}_z$:

$$R_z(\delta\gamma) = \tilde{I}_3 + \frac{dR_z(\gamma)}{d\gamma}\bigg|_{\gamma=0} \delta\gamma + O((\delta\gamma)^2) = \tilde{I}_3 + i\delta\gamma\tilde{S}_z.$$

A finite rotation $R(\gamma)$ may be constructed from successive infinitesimal rotations

$$R_z(\delta\gamma_1 + \delta\gamma_2) = (\tilde{I}_3 + i\delta\gamma_1\tilde{S}_z)(\tilde{I}_3 + i\delta\gamma_2\tilde{S}_z).$$

Now let $(\delta\gamma = \gamma/N$ for N rotations, with $N \to \infty$, then

$$R_z(\gamma) = \lim_{N\to\infty}\left[\tilde{I}_3 + (i\gamma/N)\tilde{S}_z\right]^N = \exp(i\tilde{S}_z),$$

which identifies $\tilde{S}_z$ as the generator of the rotation group R_z. Similarly, we can find the generators of the subgroups of rotations about the x-axis and the y-axis.

The $SU(n)$ groups

The $n \times n$ unitary matrices $\tilde{U}$ also form a group, the $U(n)$ group. If there is the additional restriction that the determinant of the matrices be $+1$, we have the special unitary or unitary unimodular group, $SU(n)$. Each $n \times n$ unitary matrix has $n^2 - 1$ independent parameters (Problem 12.14).

For $n = 2$ we have $SU(2)$ and possible ways to parameterize the matrix U are

$$\tilde{U} = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix},$$

where a, b are arbitrary complex numbers and $|a|^2 + |b|^2 = 1$. These parameters are often called the Cayley–Klein parameters, and were first introduced by Cayley and Klein in connection with problems of rotation in classical mechanics.

Now let us write our unitary matrix in exponential form:

$$\tilde{U} = e^{i\tilde{H}},$$

where $\tilde{H}$ is a Hermitian matrix. It is easy to show that $e^{i\tilde{H}}$ is unitary:

$$(e^{i\tilde{H}})^{+}(e^{i\tilde{H}}) = e^{-i\tilde{H}^{+}}e^{i\tilde{H}} = e^{i(\tilde{H}-\tilde{H}^{+})} = 1.$$

This implies that any $n \times n$ unitary matrix can be written in exponential form with a particularly selected set of n^2 Hermitian $n \times n$ matrices, $\tilde{H}_j$

$$\tilde{U} = \exp\left(i \sum_{j=1}^{n^2} \theta_j \tilde{H}_j\right),$$

where the θ_j are real parameters. The $n^2 \tilde{H}_j$ are the generators of the group $U(n)$. To specialize to $SU(n)$ we need to meet the restriction det $\tilde{U} = 1$. To impose this restriction we need to use the identity

$$\det e^{\tilde{A}} = e^{\mathrm{Tr}\tilde{A}}$$

for any square matrix $\tilde{A}$. The proof is left as homework (Problem 12.15). Thus the condition det $\tilde{U} = 1$ requires $\mathrm{Tr}\,\tilde{H} = 0$ for every $\tilde{H}$. Accordingly, the generators of $SU(n)$ are any set of $n \times n$ traceless Hermitian matrices.

For $n = 2$, $SU(n)$ reduces to $SU(2)$, which describes rotations in two-dimensional complex space. The determinant is $+1$. There are three continuous parameters ($2^2 - 1 = 3$). We have expressed these as Cayley–Klein parameters. The orthogonal group $SO(3)$, determinant $+1$, describes rotations in ordinary three-dimensional space and leaves $x^2 + y^2 + z^2$ invariant. There are also three independent parameters. The rotation interpretations and the equality of numbers of independent parameters suggest these two groups may be isomorphic or homomorphic. The correspondence between these groups has been proved to be two-to-one. Thus $SU(2)$ and $SO(3)$ are isomorphic. It is beyond the scope of this book to reproduce the proof here.

The $SU(2)$ group has found various applications in particle physics. For example, we can think of the proton (p) and neutron (n) as two states of the same particle, a nucleon N, and use the electric charge as a label. It is also useful to imagine a particle space, called the strong isospin space, where the nucleon state points in some direction, as shown in Fig. 12.6. If (or assuming that) the theory that describes nucleon interactions is invariant under rotations in strong isospin space, then we may try to put the proton and the neutron as states of a spin-like doublet, or $SU(2)$ doublet. Other hadrons (strong-interacting particles) can also be classified as states in $SU(2)$ multiplets. Physicists do not have a deep understanding of why the Standard Model (of Elementary Particles) has an $SU(2)$ internal symmetry.

For $n = 3$ there are eight independent parameters ($3^2 - 1 = 8$), and we have $SU(3)$, which is very useful in describing the color symmetry.

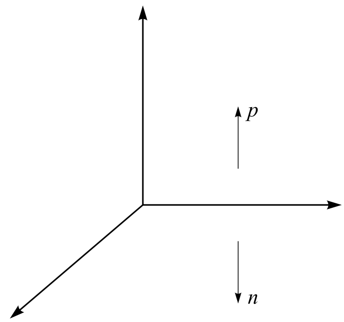

Figure 12.6. The strong isospin space.

Homogeneous Lorentz group

Before we describe the homogeneous Lorentz group, we need to know the Lorentz transformation. This will bring us back to the origin of special theory of relativity. In classical mechanics, time is absolute and the Galilean transformation (the principle of Newtonian relativity) asserts that all inertial frames are equivalent for describing the laws of classical mechanics. But physicists in the nineteenth century found that electromagnetic theory did not seem to obey the principle of Newtonian relativity. Classical electromagnetic theory is summarized in Maxwell's equations, and one of the consequences of Maxwell's equations is that the speed of light (electromagnetic waves) is independent of the motion of the source. However, under the Galilean transformation, in a frame of reference moving uniformly with respect to the light source the light wave is no longer spherical and the speed of light is also different. Hence, for electromagnetic phenomena, inertial frames are not equivalent and Maxwell's equations are not invariant under Galilean transformation. A number of experiments were proposed to resolve this conflict. After the Michelson–Morley experiment failed to detect ether, physicists finally accepted that Maxwell's equations are correct and have the same form in all inertial frames. There had to be some transformation other than the Galilean transformation that would make both electromagnetic theory and classical mechanical invariant.

This desired new transformation is the Lorentz transformation, worked out by H. Lorentz. But it was not until 1905 that Einstein realized its full implications and took the epoch-making step involved. In his paper, 'On the Electrodynamics of Moving Bodies' (*The Principle of Relativity*, Dover, New York, 1952), he developed the Special Theory of Relativity from two fundamental postulates, which are rephrased as follows:

(1) The laws of physics are the same in all inertial frame. No preferred inertial frame exists.
(2) The speed of light in free space is the same in all inertial frames and is independent of the motion of the source (the emitting body).

These postulates are often called Einstein's principle of relativity, and they radically revised our concepts of space and time. Newton's laws of motion abolish the concept of absolute space, because according to the laws of motion there is no absolute standard of rest. The non-existence of absolute rest means that we cannot give an event an absolute position in space. This in turn means that space is not absolute. This disturbed Newton, who insisted that there must be some absolute standard of rest for motion, remote stars or the ether system. Absolute space was finally abolished in its Maxwellian role as the ether. Then absolute time was abolished by Einstein's special relativity. We can see this by sending a pulse of light from one place to another. Since the speed of light is just the distance it has traveled divided by the time it has taken, in Newtonian theory, different observers would measure different speeds for the light because time is absolute. Now in relativity, all observers agree on the speed of light, but they do not agree on the distance the light has traveled. So they cannot agree on the time it has taken. That is, time is no longer absolute.

We now come to the Lorentz transformation, and suggest that the reader to consult books on special relativity for its derivation. For two inertial frames with their corresponding axes parallel and the relative velocity v along the $x_1(=x)$ axis, the Lorentz transformation has the form:

$$\begin{aligned} x_1' &= \gamma(x_1 + i\beta x_4), \\ x_2' &= x_2, \\ x_3' &= x_3, \\ x_4' &= \gamma(x_4 - i\beta x_1), \end{aligned}$$

where $x_4 = ict$, $\beta = v/c$, and $\gamma = 1/\sqrt{1-\beta^2}$. We will drop the two directions perpendicular to the motion in the following discussion.

For an infinitesimal relative velocity δv, the Lorentz transformation reduces to

$$\begin{aligned} x_1' &= x_1 + i\delta\beta x_4, \\ x_4' &= x_4 - i\delta\beta x_1, \end{aligned}$$

where $\delta\beta = \delta v/c$, $\gamma = 1/\sqrt{1-(\delta\beta)^2} \approx 1$. In matrix form we have

$$\begin{pmatrix} x_1' \\ x_4' \end{pmatrix} = \begin{pmatrix} 1 & i\delta\beta \\ -i\delta\beta & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_4 \end{pmatrix}.$$

We can express the transformation matrix in exponential form:

$$\begin{pmatrix} 1 & i\delta\beta \\ -i\delta\beta & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \delta\beta\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = \tilde{I} + \delta\beta\tilde{\sigma},$$

where

$$\tilde{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tilde{\sigma} = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Note that $\tilde{\sigma}$ is the negative of the Pauli spin matrix $\tilde{\sigma}_2$. Now we have

$$\begin{pmatrix} x_1' \\ x_4' \end{pmatrix} = (\tilde{I} + \delta\beta\tilde{\sigma})\begin{pmatrix} x_1 \\ x_4 \end{pmatrix}.$$

We can generate a finite transformation by repeating the infinitesimal transformation N times with $N\delta\beta = \theta$:

$$\begin{pmatrix} x_1' \\ x_4' \end{pmatrix} = \left(\tilde{I} + \frac{\theta\tilde{\sigma}}{N}\right)^N \begin{pmatrix} x_1 \\ x_4 \end{pmatrix}.$$

In the limit as $N \to \infty$,

$$\lim_{N\to\infty}\left(\tilde{I} + \frac{\theta\tilde{\sigma}}{N}\right)^N = e^{\theta\tilde{\sigma}}.$$

Now we can expand the exponential in a Maclaurin series:

$$e^{\theta\tilde{\sigma}} = \tilde{I} + \theta\tilde{\sigma} + (\theta\tilde{\sigma})^2/2! + (\theta\tilde{\sigma})^3/3! + \cdots$$

and, noting that $\tilde{\sigma}^2 = 1$ and

$$\sinh\theta = \theta + \theta^3/3! + \theta^5/5! + \theta^7/7! + \cdots,$$

$$\cosh\theta = 1 + \theta^2/2! + \theta^4/4! + \theta^6/6! + \cdots,$$

we finally obtain

$$e^{\theta\tilde{\sigma}} = \tilde{I}\cosh\theta + \tilde{\sigma}\sinh\theta.$$

Our finite Lorentz transformation then takes the form:

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} \cosh\theta & i\sinh\theta \\ -i\sinh\theta & \cosh\theta \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

and $\tilde{\sigma}$ is the generator of the representations of our Lorentz transformation. The transformation

$$\begin{pmatrix} \cosh\theta & i\sinh\theta \\ -i\sinh\theta & \cosh\theta \end{pmatrix}$$

can be interpreted as the rotation matrix in the complex x_4x_1 plane (Problem 12.16).

It is straightforward to generalize the above discussion to the general case where the relative velocity is in an arbitrary direction. The transformation matrix will be a 4×4 matrix, instead of a 2×2 matrix one. For this general case, we have to take x_2- and x_3-axes into consideration.

Problems

12.1. Show that
(*a*) the unit element (the identity) in a group is unique, and
(*b*) the inverse of each group element is unique.

12.2. Show that the set of complex numbers $1, i, -1$, and $-i$ form a group of order four under multiplication.

12.3. Show that the set of all rational numbers, the set of all real numbers, and the set of all complex numbers form infinite Abelian groups under addition.

12.4. Show that the four matrices

$$\tilde{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \tilde{D} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

form an Abelian group of order four under multiplication.

12.5. Show that the set of all permutations of three objects

$$[1\ 2\ 3],\ [2\ 3\ 1],\ [3\ 1\ 2],\ [1\ 3\ 2],\ [3\ 2\ 1],\ [2\ 1\ 3]$$

forms a non-Abelian group of order six, with sequential performance as the law of combination.

12.6. Given two elements A and B subject to the relations $A^2 = B^2 = E$ (the identity), show that:
(*a*) $AB \neq BA$, and
(*b*) the set of six elements E, A, B, A^2, AB, BA form a group.

12.7. Show that the set of elements $1, A, A^2, \ldots, A^{n-1}$, $A^n = 1$, where $A = e^{2\pi i/n}$ forms a cyclic group of order n under multiplication.

12.8. Consider the rotations of a line about the z-axis through the angles $\pi/2, \pi, 3\pi/2$, and 2π in the xy plane. This is a finite set of four elements, the four operations of rotating through $\pi/2, \pi, 3\pi/2$, and 2π. Show that this set of elements forms a group of order four under addition.

12.9. Construct the group multiplication table for the group of Problem 12.2.

12.10. Consider the possible rearrangement of two objects. The operation E_p leaves each object in its place, and the operation I_p interchanges the two objects. Show that the two operations form a group that is isomorphic to G_2.

https://doi.org/10.1017/CBO9780511755781.013 Published online by Cambridge University Press

Next, we associate with the two operations two operators $\hat{O}_{E_p}$ and $\hat{O}_{I_p}$, which act on the real or complex function $f(x_1, y_1, z_1; x_2, y_2, z_2)$ with the following effects:

$$\hat{O}_{E_p} f = f, \quad \hat{O}_{I_p} f(x_1, y_1, z_1; x_2, y_2, z_2) = f(x_2, y_2, z_2; x_1, y_1, z_1).$$

Show that the two operators form a group that is isomorphic to G_2.

12.11. Verify that the multiplication table of S_3 has the form:

	P_1	P_2	P_3	P_4	P_5	P_6
P_1	P_1	P_2	P_3	P_4	P_5	P_6
P_2	P_2	P_1	P_6	P_5	P_6	P_4
P_3	P_3	P_4	P_5	P_6	P_2	P_1
P_4	P_4	P_5	P_3	P_1	P_6	P_2
P_5	P_5	P_3	P_4	P_2	P_1	P_6
P_6	P_6	P_2	P_1	P_3	P_4	P_5

12.12. Show that an $n \times n$ orthogonal matrix has $n(n-1)/2$ independent elements.

12.13. Show that the 2×2 matrix σ_2 can be obtained from the rotation matrix $R(\sigma)$ by differentiation at the identity of $SO(2)$, that is, $\theta = 0$.

12.14. Show that an $n \times n$ unitary matrix has $n^2 - 1$ independent parameters.

12.15. Show that $\det e^{\tilde{A}} = e^{\mathrm{Tr}\,\tilde{A}}$ where $\tilde{A}$ is any square matrix.

12.16. Show that the Lorentz transformation

$$x_1' = \gamma(x_1 + i\beta x_4),$$
$$x_2' = x_2,$$
$$x_3' = x_3,$$
$$x_4' = \gamma(x_4 - i\beta x_1)$$

corresponds to an imaginary rotation in the $x_4 x_1$ plane. (A detailed discussion of this can be found in the book *Classical Mechanics*, by Tai L. Chow, John Wiley, 1995.)