CHAPTER 10

INTRODUCTION TO GROUP THEORY

INTRODUCTION

Thus far in this book we have discussed objects (vectors) which combine in a way that is essentially *additive*, i.e., they combine commutatively. In investigating the properties of these objects we were led in a natural way to consider *transformations* which changed one vector into another. We observed that these transformations themselves could be considered as constituting a vector space. However, since the transformations took one vector into another, they could also be combined in a way which was essentially *multiplicative* in nature; i.e., the commutative law did not in general hold. In this concluding chapter we want to introduce some of the techniques by which one can study collections of objects which are characterized in terms of their multiplicative properties. Such a collection is called a *group*.

10.1 AN INDUCTIVE APPROACH

Let us first recall the definition of a group given previously in Chapter 3:

Definition 10.1. A *group*, G, is a collection of objects which can be combined via a closed operation, which we will denote by a dot. By a closed operation we mean one such that if $a, b \in G$, then $a \cdot b \in G$. The operation must furthermore satisfy the following three axioms:

1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (associative law)
2) there exists an identity element, e, such that $a \cdot e = a$ for all $a \in G$.
3) for every $a \in G$, there exists an inverse element, a^{-1}, such that $a \cdot a^{-1} = e$.

Note that in this definition we have required that the identity be only a *right* identity and that the inverse be only a *right* inverse. However, because of the other group properties, it follows that e is also a *left* identity $(e \cdot a = a)$ and a^{-1} is also a *left* inverse $(a^{-1} \cdot a = e)$.

To show this we note that since $a^{-1} \in G$, it must have an inverse, which we denote by $(a^{-1})^{-1}$. According to axiom (3),

$$(a^{-1}) \cdot (a^{-1})^{-1} = e\,. \tag{10.1}$$

But by Axioms (1), (2), and (3) we may write Eq. (10.1) as

$$\begin{aligned} e = a^{-1} \cdot (a^{-1})^{-1} &= (a^{-1} \cdot e) \cdot (a^{-1})^{-1} = [a^{-1} \cdot (a \cdot a^{-1})] \cdot (a^{-1})^{-1} \\ &= [(a^{-1} \cdot a) \cdot a^{-1}] \cdot (a^{-1})^{-1} = (a^{-1} \cdot a) \cdot [a^{-1} \cdot (a^{-1})^{-1}]\,. \end{aligned}$$

Using Eq. (10.1) again, we have

$$e = (a^{-1}\cdot a)\cdot e = a^{-1}\cdot(a\cdot e) \ .$$

By Axiom (2), $a\cdot e = a$, so finally we obtain

$$e = a^{-1}\cdot a \ . \tag{10.2}$$

Hence a^{-1} is a left inverse as well as a right inverse. This result contrasts with the situation for general linear operators, where the existence of a right inverse does not necessarily imply the existence of a left inverse.

Using this result, we can immediately see that e must be a left identity. By Axiom (2), e is a right identity, that is,

$$a\cdot e = a \ ,$$

but according to Eq. (10.2), this is equivalent to

$$a\cdot(a^{-1}\cdot a) = a \ .$$

Using Axioms (1) and (3), we find that

$$a = a\cdot(a^{-1}\cdot a) = (a\cdot a^{-1})\cdot a = e\cdot a \ ,$$

that is, we have

$$a = e\cdot a \ ,$$

so e is also a left identity.

Definition 10.2. A group, G is said to be *abelian* if $a\cdot b = b\cdot a$ for any $a, b\in G$.

Definition 10.3. A group with n elements is called a group of *order n*.

In the first three sections of this chapter, we shall consider only finite groups ($n < \infty$) in order to fix our thinking on the most basic aspects of group theory. In subsequent sections, we shall have some occasion to mention infinite groups; when we do, we shall always state explicitly that this is the case.

Let us now introduce a few groups, starting with the simplest ones.

Example 10.1. The most uncomplicated group which one can imagine is the one element group: $\{e\}$. It obviously satisfies all the requirements of Definition 10.1. It is not a particularly interesting group.

Example 10.2. Next we have the group with two distinct elements: $\{e, a\}$. Now $a\cdot a$ must belong to this group according to Definition 10.1. Thus we must have either $a\cdot a = a$ or $a\cdot a = e$. The former is immediately ruled out by the fact that since a has an inverse, $a\cdot a = a$ implies that $a = e$. This contradicts the assumption that the group has two distinct elements. Thus $a\cdot a = e$. The integers $\{1, -1\}$ form a group with this structure if they combine via ordinary arithmetical multiplication ($e = 1$, $a = -1$). Also, the integers $\{0, 1\}$ form a group with addition modulo two as the rule of combination: $0 + 0 = 0$, $0 + 1 = 1$, $1 + 0 = 1$, $1 + 1 = 0$. In this case, $e = 0$, $a = 1$. Another structure of this type is the group of permutations of two objects; e is the permutation which leaves the order of the objects unchanged, and a is the permutation which interchanges them.

This last example leads naturally to the following definition.

Definition 10.4. Two groups are said to be *isomorphic* if there exists a one-to-one operation-preserving correspondence between them. That is, suppose that G has elements $a, b, c, \cdots$, and G' has elements $a', b', c', \cdots$ If $(a \cdot b)' = (a' \times b')$ for all $a, b \in G$ and $a', b' \in G$, then the groups $\{G, \cdot\}$ and $\{G', \times\}$ are said to be isomorphic.

Evidently the three groups mentioned in Example 10.2 are isomorphic to each other: once we know everything about any *one* of them, we know everything about *all* of them.

Example 10.3. Let us now move on to groups with three distinct elements, which we denote by $\{e, a, b\}$. From the axioms we see immediately that $a \cdot b = e$, for if $a \cdot b = a$ or $a \cdot b = b$, then the existence of an inverse would imply $b = e$ in the first case and $a = e$ in the second case, which is contrary to our assumption that this group has three *distinct* elements. What about the product $a \cdot a$? Clearly, $a \cdot a = a$ implies that $a = e$, so either $a \cdot a = e$ or $a \cdot a = b$. But

$$a \cdot a = e \Longrightarrow a \cdot (a \cdot b) = b \Longrightarrow a = b\,,$$

since $a \cdot b = e$. But this, too, is impossible because we have assumed three distinct elements. Thus $a \cdot a = b$. Similarly $b \cdot b = a$. Adopting the obvious notation for quantities like $a \cdot a$, we may write our three element group as $\{e, a, a^2\}$, with $a^3 = e$. This is the *only* possibility for the three-element group. The postulates of Definition 10.1 restrict us to just one form. A simple example of this group is the set of three cube roots of unity, under ordinary multiplication of complex numbers.

Clearly, it is always possible to construct a group of the form $\{e, a, a^2, a^3, \cdots, a^{n-1}\}$ for any n; the basic model of such a group is the set of n nth roots of unity. In general, consider an arbitrary group, $\{G, \cdot\}$ and let $a \in G$. Form the sequence $a^0 = e, a, a^2, \cdots, a^i, \cdots$, and let n be the smallest *nonzero* integer such that $a^n = e$. Then the element a is said to be of *order n* (if n is infinite, we say that a is of infinite order). In this case, all the elements $a^0, a^1, a^2, \cdots, a^{n-1}$ are distinct, for if $a^i = a^j$ $(i < n, j < n)$, then $a^{|i-j|} = e$. But since $0 \leq i < n$ and $0 \leq j < n$, $|i - j| < n$, which contradicts our assumption that n was the *smallest* nonzero integer for which $a^n = e$.

Definition 10.5. A group whose elements can be written as $\{e, a, a^2, \cdots, a^{n-1}\}$ is called a *cyclic* group (of order n).

Thus far, all our groups have been of the cyclic type, so it is natural to ask: Are all groups cyclic? The answer is provided by the following example.

Example 10.4. We apply the methods already used to the possible groups of order four, which we denote by $\{e, a, b, c\}$. Let us begin by assuming $a^2 = b$;

then we can write the corresponding group as $\{e, a, a^2, c\}$. By an argument used before, $a \cdot c$ must equal either e or a^2. If $a \cdot c = a^2$, we conclude that $a = c$, which contradicts the assumption of four distinct elements. Thus $a \cdot c = e$, so it begins to look like we are obtaining a cyclic group. We therefore look at a^3. Suppose that $a^3 = e$; then $a^3 = a \cdot c$, so $a^2 = c$, which is impossible. If $a^3 = a$, then $a^2 = e$, which again is impossible. Similarly, if $a^3 = a^2$, $a = e$, which is impossible. Thus $a^3 = c$, and we indeed have found a cyclic group of order four: $\{e, a, a^2, a^3\}$, with $a^4 = e$.

The same result is obtained if we assume that $a^2 = c$, except that it then follows by the same reasoning as above that $b = a^3$. As mentioned above, $a^2 = a$ is not permissible. However, there remains the possibility that $a^2 = e$. Clearly, either $a \cdot b = e$ or $a \cdot b = c$. If $a \cdot b = e$, then $a^2 \cdot b = a$; since $a^2 = e$, we have $a = b$, which is impossible. Thus $a \cdot b = c$. By the same process of elimination, we find that $b \cdot a = c$, $a \cdot c = b$ and $c \cdot a = b$. Since $a \cdot b = c$ and $b \cdot a = c$, we have

$$(b \cdot a) \cdot (a \cdot b) = c \cdot c \Longrightarrow b \cdot a^2 \cdot b = c^2 .$$

But $a^2 = e$, so we have immediately $b^2 = c^2$, which means that c^2 cannot equal b or c. Thus either $c^2 = e$ or $c^2 = a$. We first examine the case $c^2 = a$. Then $a \cdot c^2 = a^2 = e$, but $a \cdot c = b$ so we have $b \cdot c = e$, which gives us explicit expressions for the products of all elements of the group. We summarize by writing

$$\begin{gathered} a^2 = e , \qquad b^2 = c^2 = a , \\ b \cdot c = c \cdot b = e , \\ a \cdot b = b \cdot a = c , \qquad a \cdot c = c \cdot a = b . \end{gathered} \tag{10.3}$$

We now want to know if this group of order four is different from the cyclic group of order four which we have already found several times. A very convenient way of analyzing simple finite groups is by means of a *multiplication table*. We illustrate it by using the *cyclic group* of order four:

	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

(10.4)

The meaning of this table is simple: in the ij-box one writes the product of the ith element of the group times the jth element. The reader can check for himself that this table represents accurately the fourth-order cyclic group first discussed in Example 10.4. Now what about the group whose multiplicative properties are given by Eqs. (10.3)? We have for the multiplication table

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	a	e
c	c	b	e	a

(10.5)

This looks completely different from the table for the fourth-order cyclic group. However, it may nevertheless be *isomorphic* to the fourth-order cyclic group. This will be the case if we can relabel the elements in such a way that the multiplication table is the same as that of the fourth-order cyclic group. If in Table (10.5) we redefine the elements according to

$$e \to e', \qquad a \to b',$$
$$b \to a', \qquad c \to c',$$

then Table (10.5) becomes

	e'	a'	b'	c'
e'	e'	a'	b'	c'
a'	a'	b'	c'	e'
b'	b'	c'	e'	a'
c'	c'	e'	a'	b'

(10.6)

Since Table (10.6) is the same as Table (10.4), we still have not found a group with a multiplicative structure different from that of the cyclic group of order four.

However, in the class of four-element groups, one possibility remains. We have been forced to conclude that $a^2 = e$, $a \cdot b = c$, $b \cdot a = c$, $a \cdot c = b$, $c \cdot a = b$, and $b^2 = c^2$. When we let $b^2 = c^2 = a$, we were led back to the cyclic group. There remains, however, the choice $b^2 = c^2 = e$. Then

$$a \cdot b = c \Longrightarrow a \cdot b^2 = c \cdot b \Longrightarrow c \cdot b = a,$$

and similarly

$$b \cdot a = c \Longrightarrow b^2 \cdot a = b \cdot c \Longrightarrow b \cdot c = a.$$

This completes the multiplication table:

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

(10.7)

The reader can check that there is no rearrangement of these elements which will give the Table (10.4). Thus we have finally found a noncyclic group. Note, however, that it is an abelian group; this is reflected in the fact that the array in Table (10.7) is symmetric with respect to the diagonal. This group is called the four-group. A simple realization of the four-group is the set of all transformations of a rectangle which leave the rectangle's orientation in space unchanged (remember that in a rectangle the right and left sides are indistinguishable, as are the top and bottom sides). If the rectangle sits on an xy-plane with its center at the origin, there are four such transformations: (i) the identity transformation, (ii) rotation through 180° about an axis perpendicular to the rectangle and passing through its center, (iii) reflection through the x-axis, and (iv) reflection through the y-axis. Note that we omit rotations through 360°, 540°, etc., since they just duplicate the transformations (i) and (ii) above. Figure 10.1 illustrates the configuration. We leave it to the reader to show that this is indeed a group and that its multiplication table is the same as that given in Table (10.7).

From Table (10.7) it also appears that the elements $\{e, a\}$ of the four-group form a group by themselves ($e \cdot e = e$, $e \cdot a = a$, $a \cdot e = a$, $a \cdot a = e$), as do $\{e, b\}$ and $\{e, c\}$. This leads to the following definition.

Definition 10.6. A set of elements H, which is contained in G, is said to be a *subgroup* of G if

i) the product of any pair of elements in H is in H,

ii) if $a \in H$, then $a^{-1} \in H$.

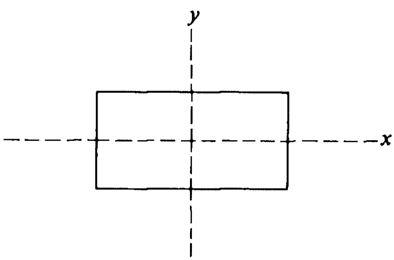

Fig. 10.1

Any subgroup is also a group because its elements are already in G, and hence obey the associative law; moreover, $a, a^{-1} \in H$ implies that $a \cdot a^{-1} = e$ also belongs to H. Clearly any group is a subgroup of itself, and the identity element is a subgroup of any group. As mentioned above, the four-group has $\{e, a\}$ as a subgroup, as well as isomorphic subgroups $\{e, b\}$ and $\{e, c\}$. Clearly, in any group, G, the collection $\{e, a, a^2, \cdots, a^{n-1}\}$, where n is the order of a and $a \in G$, is an abelian subgroup of G.

Another interesting feature of Table (10.7), which it shares with all the other multiplication tables written above, is the fact that each row (or column) contains each element of the group once and only once. This is not accidental; we have in fact the following result.

Theorem 10.1. If G is a group of order n, with elements $e, a_2, a_3, \cdots, a_n$, then every element of G occurs once and only once in the sequence

$$ea_i,\ a_2a_i,\ \cdots,\ a_na_i\ ,$$

for any i, and similarly for

$$a_ie,\ a_ia_2,\ \cdots,\ a_ia_n\ .$$

Note that in the statement of this theorem we have omitted the dot in denoting the product of group elements. We shall do this throughout the chapter, except when this omission might cause confusion.

Proof. If some element occurred twice, we would have $a_\mu a_i = a_\nu a_i$; this would imply $a_\mu = a_\nu$, contrary to our assumption of distinct group elements. Since there are n elements in the sequence and no element occurs twice, each element occurs once and only once. This explains the above mentioned structure of the group multiplication table. This very simple looking theorem is actually of central importance in the proof of most of the results of this chapter.

One could continue inductively, in the manner of this section, developing the groups of order five, order six, etc., but this would not be very instructive. We may remark, however, that not all groups are abelian like the ones which we have discussed so far; at order six one finds the first nonabelian group, the group of permutations of three objects. It should also be emphasized that not all groups have a finite number of elements. For example, the integers under addition ($e = 0$, $n^{-1} = -n$) form an (abelian) group of infinite order; the set of all unitary $n \times n$ matrices is an infinite nonabelian group.

10.2 THE SYMMETRIC GROUPS

By way of further illustrating group structure, we now discuss one of the most important groups of mathematics and physics, the group of permutations on n objects, called the symmetric group, S_n. The structure of the groups S_n is extraordinarily rich, and in this section we will consider only the simplest aspects of these groups.

There are $n!$ permutations of n objects, so S_n is of order $n!$. A typical one of these $n!$ elements will be denoted by

$$p = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ m_1 & m_2 & m_3 & \cdots & m_n \end{pmatrix},$$

where the set $\{m_1, m_2, m_3, \cdots, m_n\}$ is some arrangement of the first n integers. This symbol means that 1 is replaced by m_1, 2 by m_2, etc. For example, the permutation

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix},$$

acting on the arrangement $\{1\ 2\ 3\ 4\ 5\ 6\}$ produces the arrangement $\{6\ 5\ 4\ 3\ 2\ 1\}$; p acting on $\{2\ 5\ 3\ 4\ 6\ 1\}$ produces $\{5\ 2\ 4\ 3\ 1\ 6\}$. In this example, 1 and 6 go into each other in a closed manner: 1 is replaced by 6, and 6 is replaced by 1. Such a structure is called a *cycle* and will be written simply as (16). Similarly, the above example contains the cycles (25) and (34). We can thus write

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix} = (16)(25)(34).$$

When we perform such a factorization, we say that we factor the given permutation into *disjoint* cycles, or cycles having no elements in common.

Of course, not all cycles contain only two elements. A cycle containing l elements is called an l-cycle, or a cycle of length l. For example, the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 1 & 2 & 5 & 3 \end{pmatrix}$$

can be written as a disjoint product of a three-cycle, a two-cycle, and a one-cycle:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 1 & 2 & 5 & 3 \end{pmatrix} = (163)(24)(5).$$

The three-cycle, (163), is to be read as "1 is replaced by 6, 6 is replaced by 3 and 3 is replaced by 1." This structure closes on itself just like a two-cycle. Note that

$$(163) = (316) = (631),$$

but

$$(163) \neq (136).$$

Thus as the name "cycle" suggests, if one writes the elements of a cycle clockwise around a circle, one can start performing the sequence of replacements at *any* element and proceed clockwise back to the starting point. However, if one proceeds in a counterclockwise direction, one gets a different permutation.

The two-cycles, usually referred to as transpositions, are especially important because any cycle (and hence any permutation) can be written as a product of two-cycles (which are not, in general, disjoint). For example,

$$(163) = (13)(16) \, ,$$

since when (163) acts on say $\{136\}$, it produces $\{613\}$, and (16) acting on $\{136\}$ produces $\{631\}$, which in turn becomes $\{613\}$ when acted on by (13). Here we adopt the convention that when a string of cycles is written, the one on the right acts first, and the one on the left acts last. If the cycles are disjoint, it clearly makes no difference in which order they act, but if they are not disjoint, the order is crucial. For example,

$$(163)\{136\} = (13)(16)\{136\} = \{613\} \, ,$$

whereas

$$(16)(13)\{136\} = \{361\} = (136)\{136\} \, .$$

Also, the decomposition into two-cycles is *not* unique. In the above case,

$$(163) = (13)(16)$$

and

$$(163) = (316) = (36)(31) \, .$$

In general, a decomposition of an n-cycle into transpositions can be written as

$$(1\ 2\ 3\ \cdots\ n) = (1n)(1\ n-1)\cdots(13)(12) \, ,$$

as may be readily verified. The parity of a permutation is defined to be $(-1)^N$, where N is the number of transpositions in a given permutation. If N is odd, we speak of an odd permutation; if N is even, we speak of an even permutation.

If now by a *product* of permutations, we mean simply the two permutations carried out successively, then obviously the product of two permutations is another permutation. All the other group axioms are trivially satisfied so the collection of all permutations on n objects is a group, and this group will contain $n!$ elements. By our above discussion, it is easily seen that only S_1 and S_2 are abelian groups. S_n $(n \geq 3)$ is nonabelian. For example, (12) and (13) belong to S_3, and $(12)(13) \neq (13)(12)$.

Note that in our convention for writing a permutation, we have quite a bit of freedom in ordering; in the permutation symbol, only the vertical relationship matters, not the horizontal. Thus

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 5 & 4 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 6 & 3 & 1 & 5 \\ 3 & 5 & 1 & 4 & 6 & 2 \end{pmatrix} .$$

Also, we have used numbers in our symbols merely for convenience (there is an arbitrarily large number of them). One could equally well use apple, pear, orange, $\cdots$, or Greek, Russian, French, $\cdots$, instead of one, two, three, $\cdots$

Example 10.5. The simplest symmetric groups are S_1, which contains the single element e, and S_2, which contains the elements e and (12). More interesting is S_3. Its six elements are

$$e,\ (12),\ (13),\ (23),\ (123),\ (321)\ .$$

As the reader can easily show, the multiplication table for S_3 is

	e	(12)	(13)	(23)	(123)	(321)
e	e	(12)	(13)	(23)	(123)	(321)
(12)	(12)	e	(321)	(123)	(23)	(13)
(13)	(13)	(123)	e	(321)	(12)	(23)
(23)	(23)	(321)	(123)	e	(13)	(12)
(123)	(123)	(13)	(23)	(12)	(321)	e
(321)	(321)	(23)	(12)	(13)	e	(123)

If we make the identifications $e = I$, $(12) = A$, $(13) = B$, $(23) = C$, $(123) = D$, $(321) = F$, then this table is identical to the one given in Table (3.4). Note that this group has quite a few subgroups. Clearly $\{e, (12)\}$ is a subgroup of order two, and $\{e, (13)\}$ and $\{e, (23)\}$ are isomorphic to it. $\{e, (123), (321)\}$ is a subgroup of order three, which is isomorphic to the cyclic group of order three discussed in Example 10.3. Inspection of the multiplication table shows that S_3 is a nonabelian group.

One of the most remarkable facts about S_n is embodied in the following theorem of Cayley.

Theorem 10.2. Every group G of order n is isomorphic to a subgroup of S_n.

Proof. The proof of this result is based on Theorem 10.1. Call the elements of G $a_1, a_2, a_3, \cdots, a_n$. Let a_i be any element of G. Then according to Theorem 10.1, the collection $\{a_ia_1, a_ia_2, \cdots, a_ia_n\}$ is a rearrangement of $\{a_1, a_2, \cdots, a_n\}$ in which every element occurs once and only once. Therefore, let us make the correspondence, which is clearly one-to-one,

$$a_i \to P_{a_i} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_ia_1 & a_ia_2 & \cdots & a_ia_n \end{pmatrix}.$$

Similarly, for $a_j \in G$,

$$a_j \to P_{a_j} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_ja_1 & a_ja_2 & \cdots & a_ja_n \end{pmatrix}.$$

For the element of G which is given by the product $a_i a_j$, we will have

$$a_i a_j \to P_{a_i a_j} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_i a_j a_1 & a_i a_j a_2 & \cdots & a_i a_j a_n \end{pmatrix}.$$

The crucial point in demonstrating that we have an isomorphism is to show that $P_{a_i} P_{a_j} = P_{a_i a_j}$. This result is easily obtained—it is merely an exercise in notation. The reader should keep this in mind as he plows through the manipulations which follow.

We have already remarked that the horizontal ordering of permutation symbols is unimportant. For example,

$$\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_i a_1 & a_i a_2 & a_i a_3 & \cdots & a_i a_n \end{pmatrix} = \begin{pmatrix} a_3 & a_n & a_1 & \cdots & a_2 \\ a_i a_3 & a_i a_n & a_i a_1 & \cdots & a_i a_2 \end{pmatrix}.$$

In other words, we can rearrange the top row in any way we please provided we make a similar rearrangement of the bottom row. Therefore, since $\{a_j a_1, a_j a_2, \cdots, a_j a_n\}$ is a rearrangement of $\{a_1, a_2, \cdots, a_n\}$, we may write

$$\begin{aligned} P_{a_i} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_i a_1 & a_i a_2 & \cdots & a_i a_n \end{pmatrix} &= \begin{pmatrix} a_j a_1 & a_j a_2 & \cdots & a_j a_n \\ a_i(a_j a_1) & a_i(a_j a_2) & \cdots & a_i(a_j a_n) \end{pmatrix} \\ &= \begin{pmatrix} a_j a_1 & a_j a_2 & \cdots & a_j a_n \\ a_i a_j a_1 & a_i a_j a_2 & \cdots & a_i a_j a_n \end{pmatrix}. \end{aligned}$$

Thus

$$\begin{aligned} P_{a_i} P_{a_j} &= \begin{pmatrix} a_j a_1 & a_j a_2 & \cdots & a_j a_n \\ a_i a_j a_1 & a_i a_j a_2 & \cdots & a_i a_j a_n \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_j a_1 & a_j a_2 & \cdots & a_j a_n \end{pmatrix} \\ &= \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_i a_j a_1 & a_i a_j a_2 & \cdots & a_i a_j a_n \end{pmatrix}, \end{aligned}$$

since the multiplication can now be determined simply by inspection. Thus

$$P_{a_i} P_{a_j} = P_{a_i a_j},$$

and the proof of isomorphism is complete.

This result encompasses the remark made in Example 10.5, where it was noted that S_3 has a subgroup $\{e, (123), (321)\}$ which is isomorphic to the group of order three. Regarding the groups of order four, we take first the four-group. Using the multiplication table of Table (10.7), we find immediately the correspondence

$$P_e = \begin{pmatrix} e & a & b & c \\ e & a & b & c \end{pmatrix}, \qquad P_a = \begin{pmatrix} e & a & b & c \\ a & e & c & b \end{pmatrix},$$

$$P_b = \begin{pmatrix} e & a & b & c \\ b & c & e & a \end{pmatrix}, \qquad P_c = \begin{pmatrix} e & a & b & c \\ c & b & a & e \end{pmatrix},$$

by using the rule of correspondence of Cayley's theorem. In the compact cycle notation, these can be written as

$$P_e = e, \qquad P_a = (ea)(bc), \qquad P_b = (eb)(ac), \qquad P_c = (ec)(ab).$$

It is a simple matter to check that the multiplication table of Table (10.7) is obeyed.

In the case of the cyclic group of order four, we find, using Table (10.4),

$$P_e = \begin{pmatrix} e & a & b & c \\ e & a & b & c \end{pmatrix}, \qquad P_a = \begin{pmatrix} e & a & b & c \\ a & b & c & e \end{pmatrix},$$

$$P_b = \begin{pmatrix} e & a & b & c \\ b & c & e & a \end{pmatrix}, \qquad P_c = \begin{pmatrix} e & a & b & c \\ c & e & a & b \end{pmatrix}.$$

As above, this can be written as

$$P_e = e\,, \qquad P_a = (eabc)\,, \qquad P_b = (eb)(ac)\,, \qquad P_c = (ecba)\,.$$

This case gives us our first example of a four-cycle.

In looking at these two illustrations, one notes that except for the identity, all the permutations leave no symbol unchanged. A subgroup of S_n with this property is called a *regular subgroup* or a subgroup of regular permutations. It is clear from the construction used in proving Theorem 10.2 that every group of order n is isomorphic to a regular subgroup of S_n.

We close this section with a few results about the regular subgroups of S_n.

Lemma. In a regular subgroup, no two elements take a given symbol into the same symbol.

Proof. Suppose that p_1 and p_2 $(p_1 \neq p_2)$ belong to a regular subgroup and that both p_1 and p_2 take a into b. But then $p_1 p_2^{-1}$ $(\neq e)$ leaves b unchanged. However, $p_1 p_2^{-1}$ also belongs to the regular subgroup (since the subgroup is closed), which contradicts the assumption that the subgroup is regular and hence that every element (except the identity) changes *all* the symbols.

Lemma. In a regular subgroup, if we decompose a given permutation into disjoint cycles, then each cycle must have the same length. For example, in S_5, (12) (345) could not belong to a regular subgroup.

Proof. Suppose that p is an element of a regular subgroup which can be decomposed into two cycles of length l_1 and l_2 with $l_1 < l_2$. Now since a cycle of length l must satisfy $(a_1 a_2 \cdots a_l)^l = e$, p^{l_1} leaves all the symbols contained in the first cycle unchanged. However, p^{l_1} must change some of the symbols contained in the second cycle (all of them, in fact), since if we have a cycle of length $\lambda > l'$, then $(a_1 a_2 \cdots a_\lambda)^{l'} \neq e$. Hence p^{l_1}, which is a permutation belonging to the subgroup containing p, changes some symbols, but not others. This contradicts our assumption that p belongs to a regular subgroup.

This lemma is illustrated by the results just obtained for the regular subgroups of S_4 corresponding to the four-group and to the cyclic group of order four. Using this lemma, we can also prove the following very powerful theorem.

Theorem 10.3. Every group $\{G, \cdot\}$ of order n is cyclic if n is a prime number.

Proof. By the previous lemma, the subgroup of S_n to which G is isomorphic must contain elements which are either a product of n one-cycles or just one n-cycle, since n is prime. The identity is a product of n one-cycles, so G consists of an identity, plus $n - 1$ elements which correspond isomorphically to n-cycles. Call these elements $a_2, a_3, \cdots, a_n$. Now since a_i corresponds to an n-cycle, each element of the sequence $\{a_i, a_i^2, a_i^3, \cdots, a_i^{n-1}\}$ is distinct (note that $a_i^n = e$), since if the elements were not all distinct, say $a_i^\mu = a_i^\nu$, then we would have $a_i^m = e$ for $m < n$, which is impossible because a_i corresponds to an n-cycle. Thus the distinct elements $\{e, a_i, a_i^2, \cdots, a_i^{n-1}\}$ exhaust all elements of the group. That is, we have a cyclic group of order n. Clearly this argument works for any element a_i.

For example, the group of order three is isomorphic to the subgroup $\{e, (123), (321)\}$ of S_3. It is easily checked that

$$(123)^1 = (123)\ , \qquad (123)^2 = (321)\ , \qquad (123)^3 = e\ ,$$

so we could write our group as

$$\{e, (123), (123)^2\}\ .$$

But also,

$$(321)^1 = (321)\ , \qquad (321)^2 = (123)\ , \qquad (321)^3 = e\ ,$$

so we could equally well reorder the group and write it as

$$\{e, (321), (321)^2\}\ .$$

Theorem 10.3 tells us, among other things, that although we might be tempted to imagine that for n large the number of groups of order n is also large, this is in fact not the case. There is only one group of order 97, namely, the cyclic group.

10.3 COSETS, CLASSES, AND INVARIANT SUBGROUPS

Having discussed in some detail a special category of finite groups, let us now look at some of the most important general properties which are common to all groups.

Definition 10.7. Let A be a subgroup of G, with elements $e, a_2, a_3, \cdots, a_m$ ($m \leq n$, where n is the order of G) and let $b \in G$, but $b \notin A$. Then the m distinct elements $be = b, ba_2, ba_3, \cdots, ba_m$ form a *left coset* of A, which we denote symbolically by bA. Similarly, $b, a_2b, a_3b, \cdots, a_mb$ form a right coset, Ab.

The notation bA is just a shorthand for the collection of objects formed by multiplying every element of A on the left by b. Note that a coset is *not* a subgroup, since it clearly cannot contain the identity element. If $ba_i = e$, then $b = a_i^{-1}$, which implies $b \in A$, contrary to the assumption in Definition 10.7. In fact, bA contains no elements of A at all, for if $ba_i = a_j$, then $b = a_ja_i^{-1}$, and hence $b \in A$, again contrary to assumption.

Lemma. Two left cosets of a subgroup A either contain all the same elements or else have no elements in common.

Proof. Let xA be one coset of A and yA be another. Suppose that $xa_i = ya_j$; then $y^{-1}x = a_j a_i^{-1}$, so $y^{-1}x \in A$. Thus $y^{-1}x$ applied to the group A must just be a rearrangement of A, according to Theorem 10.1. Hence A can be written as $\{y^{-1}xa_1, y^{-1}xa_2, \cdots, y^{-1}xa_m\}$, and the collection $\{y(y^{-1}xa_1), y(y^{-1}xa_2), \cdots, y(y^{-1}xa_m)\}$ must be identical with yA (the ordering of the elements of yA is immaterial). But this collection is just $\{xa_1, xa_2, \cdots, xa_m\}$, so it is equal to xA. Hence xA and yA are the same, and we conclude that if two cosets have one element in common, they have all elements in common. This is what we set out to prove.

Theorem 10.4 (Lagrange's theorem). The order, g, of the group G is an integral multiple of the order of any subgroup, A.

Proof. Consider all the distinct cosets of A, which we denote by $b_1A, b_2A, \cdots, b_{\mu-1}A$. (It is purely for convenience that we take the number of cosets to be $\mu - 1$.) Let h be the order of A; thus there are h elements in each coset. But every element of G must occur either in A or in one of its $\mu - 1$ distinct cosets, and no element can appear more than once. Hence $h + (\mu - 1)h = \mu h = g$, which is what was to be shown.

We call μ the *index* of the subgroup A. From this result it is apparent that the order of any element of a group must be an integral divisor of the order of the group, since each element, a, of order ν generates a subgroup $\{e, a, a^2, \cdots, a^{\nu-1}\}$ of order ν. This leads us to conclude again, as in Theorem 10.3, that the groups of prime order must be cyclic.

Example 10.6. Consider S_3, whose elements are e, (12), (13), (23), (123), and (321). The left cosets of the subgroup $\{e, (12)\} = A$ are:

$$\begin{aligned}
(13)A &= \{(13), (13)(12)\} &&= \{(13), (123)\}\,, \\
(23)A &= \{(23), (23)(12)\} &&= \{(23), (321)\}\,, \\
(123)A &= \{(123), (123)(12)\} &&= \{(123), (13)\}\,, \\
(321)A &= \{(321), (321)(12)\} &&= \{(321), (23)\}\,.
\end{aligned}$$

Thus there are two distinct left cosets of A:

$$(13)A = \{(13), (123)\}\,, \qquad (23)A = \{(23), (321)\}\,.$$

We may write symbolically

$$S_3 = A + (13)A + (23)A\,,$$

or equivalently,

$$S_3 = A + (123)A + (321)A\,.$$

Note that the right cosets of A are *not* the same as the left cosets:

$$\begin{aligned}
A(13) &= \{(13), (12)(13)\} = \{(13), (321)\}\,, \\
A(23) &= \{(23), (12)(23)\} = \{(23), (123)\}\,,
\end{aligned}$$

and similarly for $A(123)$ and $A(321)$.

Another subgroup of S_3 is $B = \{e, (123), (321)\}$. Its left cosets are:

$$(12)B = \{(12), (12)(123), (12)(321)\}$$
$$= \{(12), (23), (13)\} ;$$

$(13)B$ and $(23)B$ must equal $(12)B$ according to the previous lemma. It also follows from that lemma that the right coset $B(12)$ is equal to $(12)B$. Thus there is only one distinct coset of B. In this case we may write symbolically

$$S_3 = B + (12)B .$$

An important relationship which can exist between group elements is that of conjugacy.

Definition 10.8. An element $b \in G$ is said to be *conjugate* to $a \in G$ if there exists some $x \in G$ such that

$$b = xax^{-1} .$$

This relationship of two group elements is analogous to that of similarity between matrices, which was discussed in Chapter 3. According to this definition, a is clearly conjugate to itself and if a is conjugate to b, then b is conjugate to a. We also have the following simple result.

Lemma. If a is conjugate to b and b is conjugate to c, then a is conjugate to c.

Proof. Since a is conjugate to b, there exists $x \in G$ such that $a = xbx^{-1}$; similarly, there exists $y \in G$ such that $b = ycy^{-1}$. Combining these two results, we find that

$$a = x(ycy^{-1})x^{-1} = xycy^{-1}x^{-1} = (xy)c(xy)^{-1} ,$$

since $(xy)^{-1} = y^{-1}x^{-1}$ as may be checked by direct multiplication. But $xy \in G$, so a is conjugate to c, as required.

Thus the relation of conjugacy is reflexive, symmetric, and transitive. This suggests the following definition.

Definition 10.9. All the elements of a group which are conjugate to each other form an *equivalence class*, referred to hereafter simply as a *class*.

According to this definition any two elements of a class must be of the same order. For suppose that a is of order n ($a^n = e$) and that b is conjugate to a ($b = xax^{-1}$ for some $x \in G$). Then $b^n = (xax^{-1})^n = xax^{-1}xax^{-1} \cdots xax^{-1} = xa^nx^{-1} = xex^{-1} = e$. We also have $b^m = xa^mx^{-1}$ for any $m < n$. Now if $b^m = e$, it follows that $a^m = e$, which is impossible since a is of order n and $m < n$. Thus n is the smallest integer for which $b^n = e$; that is, b is also of order n. Obviously, the identity forms a class by itself. For any other group element, a, we form the sequence

$$eae^{-1} = a,\ a_2aa_2^{-1},\ a_3aa_3^{-1},\ \cdots,\ a_naa_n^{-1} .$$

The elements of this sequence are all conjugate to each other (of course some elements may occur more than once), and hence form a class. In this manner, the elements of any group can be divided into classes. For abelian groups, this procedure is very simple: every element constitutes a class by itself, since all the elements commute.

Example 10.7. Consider once again the group S_3: $\{e, (12), (13), (23), (123), (321)\}$. The respective inverses are e, (12), (13), (23), (321) and (123). Of course, e constitutes a class by itself. Now let us conjugate (12) by all the elements of S_3:

i) $e(12)e^{-1} = (12)$,

ii) $(12)(12)(12)^{-1} = (12)$,

iii) $(13)(12)(13)^{-1} = (13)(12)(13) = (13)(132) = (13)(321)$
$= (13)(31)(32) = (23)$,

iv) $(23)(12)(23)^{-1} = (23)(12)(23) = (23)(21)(23) = (23)(231)$
$= (23)(312) = (23)(32)(31) = (13)$,

v) $(123)(12)(123)^{-1} = (123)(12)(321) = (13)(12)(12)(321)$
$= (13)(321) = (13)(31)(32) = (23)$,

vi) $(321)(12)(321)^{-1} = (321)(12)(123) = (321)(12)(231)$
$= (321)(12)(21)(23) = (321)(23) = (31)(32)(23)$
$= (13)$.

The above calculations illustrate some of the techniques that are useful in the manipulation of cycles. In particular, we have repeatedly made use of two important properties of cycles: (1) that the square of any two-cycle gives the identity element; and (2) that one can order a three-cycle (or generally, an n-cycle) in more ways than one to suit a given situation. Thus we conclude that (12), (13), and (23) form a class. This leaves only (123) and (321) to place in classes. We have

$$e(123)e^{-1} = (123)$$
$$(12)(123)(12)^{-1} = (12)(13)(12)(12) = (12)(13) = (132) = (321) .$$

We need go no further, since we see that (123) and (321) are conjugate to each other. They are therefore in the same class. Since this exhausts the elements of the group, we have determined all the classes of S_3. We summarize as follows:

$$\mathscr{C}_1 = e ,$$
$$\mathscr{C}_2 = \{(12), (13), (23)\} ,$$
$$\mathscr{C}_3 = \{(123), (321)\} .$$

We remark here without proof that in the case of the permutation groups, the division into classes corresponds always to the division according to cycle structure, just as in the above case. That means, for example, that in the case of

S_4 the $4! = 24$ elements fall into the classes:

$$\begin{aligned}
\mathscr{C}_1 &= e\,, \\
\mathscr{C}_2 &= \{(12), (13), (14), (23), (24), (34)\}\,, \\
\mathscr{C}_3 &= \{(123), (321), (124), (421), (134), (431), (234), (432)\}\,, \\
\mathscr{C}_4 &= \{(12)(34), (13)(24), (14)(23)\}\,, \\
\mathscr{C}_5 &= \{(1234), (1243), (1324), (1342), (1423), (1432)\}\,.
\end{aligned}$$

Now let H be a subgroup of G. The set $H' = aHa^{-1}$, where $a \in G$, is also a subgroup of G. For if $x \in H$ and $y \in H$, axa^{-1} and aya^{-1} are two elements of H'. But

$$(axa^{-1})(aya^{-1}) = a(xy)a^{-1}\,,$$

and since $xy \in H$, $a(xy)a^{-1} \in H'$. This proves that H' is closed; the other group properties of H' may readily be verified. Evidently, if $a \in H$, then H' is just a one-to-one mapping of H onto itself. $H' = aHa^{-1}$ is said to be a *conjugate subgroup* of H in G.

Definition 10.10. If for all $a \in G$, $aHa^{-1} = H$, H is said to be an *invariant* subgroup.

Under these circumstances, $aH = Ha$, so we arrive at an alternative formulation: H is an invariant subgroup if the left and right cosets formed with any $a \in G$ are the same. Thus the subgroup B of S_3 discussed in Example 10.6 is an invariant subgroup. The subgroup A of S_3 is not invariant. Clearly, in any group, the identity is an invariant subgroup, as is the whole group itself.

Definition 10.11. A group which has no invariant subgroups save the whole group and the identity is a *simple* group.

Lemma. If H is a subgroup of G, H is invariant if and only if it contains the elements of G in complete classes; that is, if H contains one element of a class, then it must contain them all.

Proof. Assume first that H is invariant; then for any $x \in H$ and $a \in G$, $axa^{-1} \in H$. Hence if $\mathscr{C}$ is a class which contains x, all members of $\mathscr{C}$ are also in H, by Definition 10.9. By the same argument, if $y \in \mathscr{C}'$ ($\mathscr{C}' \neq \mathscr{C}$), all members of $\mathscr{C}'$ must also be in H. Thus H contains only complete classes and may contain more than one class. In fact, since e always constitutes a class by itself, only the most trivial invariant subgroup (the group consisting of the identity alone) will contain just one class. Conversely, suppose that H contains only complete classes. Then every $x \in H$ belongs to some class $\mathscr{C}$; furthermore, every element of $\mathscr{C}$ is contained in H. This means that for *any* $a \in G$, $axa^{-1} \in H$. Since this argument holds for any x, H is invariant.

One of the most interesting results of group theory is that the *cosets* of an invariant subgroup themselves constitute a group!

Theorem 10.5. The collection consisting of an invariant subgroup H and all its distinct cosets is itself a group, called the factor group of G, usually denoted by G/H. (Remember that the left and right cosets of an *invariant* subgroup are identical.) Multiplication of two cosets aH and bH is defined as the set of all distinct products $z = xy$, with $x \in aH$ and $y \in bH$; the identity element of the factor group is the subgroup H itself.

Proof. Consider any coset aH. Then since $HH = H$,

$$(aH)H = aHH = aH\,; \qquad H(aH) = HaH = aHH = aH\,,$$

so H is indeed the identity. Here we have used $aH = Ha$, which follows from the fact that H is invariant. Now take any two cosets aH and bH. We have

$$(aH)(bH) = aHbH = abHH = abH\,,$$

so the product of two cosets yields another coset. Finally, the inverse of aH is clearly $a^{-1}H$, since H is the identity and

$$(aH)(a^{-1}H) = aa^{-1}HH = HH = H\,.$$

This completes the proof that G/H is indeed a group.

A simple illustration of this theorem is provided by S_3. We have already noted that $B = \{e, (123), (321)\}$ is an invariant subgroup. Its coset can be written as $(12)B$. Thus the factor group S_3/B consists of two elements:

$$E = B = \{e, (123), (321)\}\,, \qquad A = (12)B = \{(12), (13), (23)\}\,,$$

with $A^2 = E$. In discussing the factor group we will use capital letters to denote factor group elements. When multiplying two such elements we will denote group multiplication by a dot to avoid possible confusion with matrix multiplication.

From another point of view, the factor group is a mapping of one group G onto another group G'. This mapping preserves group products, but is not necessarily one to one. Such a map is called a homomorphism; it is to be contrasted with the isomorphism, which is a product-preserving, one-to-one map. We say that G is homomorphic to G'. In the example under discussion we can consider e, (123) and (321) as being mapped onto E, the identity; and (12), (13) and (23) as being mapped onto A ($A^2 = E$). According to the proof of Theorem 10.5, this mapping must preserve products. For example, (13)(12) is an element of the product of A by A. Since $(13)(12) = (123)$, this product belongs to E, as it should since $A \cdot A = E$. Similarly, $(123)(12) \in E \cdot A$, which is again as it should be since $(123)(12) = (13)$, which belongs to A. The reader can check that all the multiplications work out as they are supposed to according to Theorem 10.5. We also leave it to the reader to show that in *any* homomorphic mapping of G onto G', the elements of G which are mapped onto the identity of G' must form an invariant subgroup of G. The other elements of G' then must be the images of the cosets of the invariant subgroup.

Example 10.8. To illustrate further some of the ideas of this section and also to look at a slightly more complicated group than we have examined so far, let us analyze the group S_4. As mentioned earlier, S_4 can be divided into five classes according to cycle structure:

$$\begin{aligned}
\mathscr{C}_1 &= e\,,\\
\mathscr{C}_2 &= (12),\ (13),\ (14),\ (23),\ (24),\ (34)\,,\\
\mathscr{C}_3 &= (123),\ (124),\ (134),\ (234),\ (321),\ (421),\ (431),\ (432)\,,\\
\mathscr{C}_4 &= (12)(34),\ (13)(24),\ (14)(23)\,,\\
\mathscr{C}_5 &= (1234),\ (1243),\ (1324),\ (1342),\ (1423),\ (1432)\,.
\end{aligned}$$

We would like to determine whether there are any nontrivial invariant subgroups of S_4. Since we know that such an invariant subgroup can contain only complete classes and since furthermore the order of the subgroup must be an integral divisor of the order ($4! = 24$) of S_4, there are, according to the previous lemma only two possibilities:

$$H = \mathscr{C}_1 + \mathscr{C}_4\,, \qquad H' = \mathscr{C}_1 + \mathscr{C}_3 + \mathscr{C}_4\,.$$

H is of order four and H' is of order twelve. Both these orders are integral divisors of 24; $H'' = \mathscr{C}_1 + \mathscr{C}_2$, for example, is excluded since $24/7 \neq$ integer. In addition, H'' is not even a subgroup. Now H' is just the subgroup of even permutations. Since any coset must have as many elements as the related subgroup, H' can have only one coset, the set $\mathscr{C}$ of all odd permutations (note that this is *not* a subgroup). In particular, the left and right cosets must be identical, so H' is indeed an invariant subgroup. The factor group S_4/H' (consisting of the two sets H' and $\mathscr{C}$) is isomorphic to the two-element cyclic group. It is obvious that the two-element cyclic group is a factor group of any S_n, since half the elements of S_n are even permutations and half are odd permutations.

That H is also a group is seen by noting that according to the discussion following the proof of Theorem 10.2, this collection of four elements is isomorphic to the four-group. Now what are the cosets of H? The computation is tedious but straightforward. We have

$$\begin{aligned}
(12)H &= \{(12),\ (34),\ (1324),\ (1423)\}\,,\\
(13)H &= \{(13),\ (24),\ (1234),\ (1432)\}\,,\\
(23)H &= \{(23),\ (14),\ (1243),\ (1342)\}\,.
\end{aligned}$$

According to the first lemma of this section, we must also have

$$\begin{aligned}
(12)H &= (34)H = (1324)H = (1423)H\,,\\
(13)H &= (24)H = (1234)H = (1432)H\,,\\
(23)H &= (14)H = (1243)H = (1342)H\,.
\end{aligned}$$

Now consider $(123)H$. We find that

$$(123)H = \{(123),\ (134),\ (432),\ (421)\}\,.$$

This contains half of $\mathscr{C}_3$, so we know (why?) that $(321)H$ must be the only remaining coset of H. In fact,

$$(321)H = \{(321), (234), (124), (431)\} .$$

This collection of six objects, $\{H, (12)H, (13)H, (23)H, (123)H, (321)H\}$, is isomorphic to S_3, and by inspection we may make the identifications:

$$\begin{aligned} &H \to e , \qquad (12)H \to (12) , \qquad (13)H \to (13) , \\ &(23)H \to (23) , \qquad (123)H \to (123) , \qquad (321)H \to (321) . \end{aligned}$$

For example, $(1324) \in (12)H$ and $(1234) \in (13)H$. A simple calculation gives

$$\begin{aligned} (1324)(1234) &= (4132)(1234) = (42)(43)(41)(14)(13)(12) \\ &= (42)(43)(13)(12) = (432)(123) = (324)(312) \\ &= (34)(32)(32)(31) = (34)(31) = (314) = (431) , \end{aligned}$$

which belongs to $(321)H$. Since $(12)(13) = (132) = (321)$, we see that the isomorphism holds:

$$[(12)H][(13)H] = (321)H \qquad \text{and} \qquad (12)(13) = (321) .$$

10.4 SYMMETRY AND GROUP REPRESENTATIONS

The role of group theory in physics is intimately related to the symmetries of the world around us. The importance of such symmetries as translational and rotational invariance in giving rise to conservation of linear and angular momentum has long been known and is familiar to the student from classical mechanics. With the development of quantum mechanics, in which the physical world is separated from us by the intermediary of the "wave function," group theory became particularly significant. To see why this is the case, we must first say precisely what is meant by such phrases as "translational invariance" and "rotational invariance."

Suppose that we have a group, G, of operators, $U_1, U_2, U_3, \cdots$, which can act on elements, x, of a vector space, V. By a group of operators we simply mean a collection of operators on V which obeys the group axioms of Definition 10.1. The results of $U_1, U_2, U_3, \cdots$ acting on x will, as usual, be denoted by $U_1x, U_2x, U_3x, \cdots$ In three-dimensional space an example of such a group is the collection $\{T_{\mathbf{a}}\}$, where $\mathbf{a}$ is any three-dimensional vector, and $T_{\mathbf{a}}$ acts on vectors according to the rule

$$T_{\mathbf{a}}\mathbf{r} = \mathbf{r} - \mathbf{a} . \tag{10.8}$$

Thus $T_{\mathbf{a}}$ is a translation operator. Note that there is an infinitude of such operators, since $\mathbf{a}$ can be *any* vector. The entire collection is clearly a group; closure follows because the product of any two translations is again a translation. In fact, from the defining equation, we see that this group is abelian since

$$T_{\mathbf{a}}T_{\mathbf{b}} = T_{\mathbf{a}+\mathbf{b}} = T_{\mathbf{b}}T_{\mathbf{a}} .$$

Obviously, such operators can be generalized to any number of dimensions. Now we suppose further that on this vector space we have defined functions, $f, g, \cdots$, which assign to each vector some complex number. In three-dimensional Euclidean space, $f(\mathbf{r}) = x^4 + y^4 + z^4$ is an example of such a function.

In general, we denote the action of a function f on a vector x in V by $f(x)$. Let U_i be any element of the group of transformations written above. What can we say about the action of f on the transformed vector $U_i^{-1}x$? Let us define an operator $\mathscr{U}_i$ which acts on functions of x in such a manner that for all f

$$\mathscr{U}_i f(x) = f(U_i^{-1}x) .$$

Now consider the quantity $\mathscr{U}_i \mathscr{U}_j f(x)$. We have from the above equation

$$\mathscr{U}_i \mathscr{U}_j f(x) = \mathscr{U}_i f(U_j^{-1}x) = f(U_j^{-1}U_i^{-1}x) = f([U_iU_j]^{-1}x) .$$

Thus, if $U_iU_j = U_k$, then $\mathscr{U}_i \mathscr{U}_j = \mathscr{U}_k$, so that the elements $\mathscr{U}_1, \mathscr{U}_2, \cdots$ form a group, $\mathscr{G}$, which is isomorphic to G.

To illustrate what we have in mind, let us extend the example of the $\{T_\mathbf{a}\}$, defined in Eq. (10.8), a bit further. We ask: Is it possible to find an operator $\mathscr{T}_\mathbf{a}$ such that

$$\mathscr{T}_\mathbf{a} f(\mathbf{r}) = f(T_\mathbf{a}^{-1}\mathbf{r}) = f(\mathbf{r} + \mathbf{a}) ?$$

In general, it will be necessary to say something about the functions $f(x)$, $g(x)$, $\cdots$ before we can obtain an explicit expression for $\mathscr{T}_\mathbf{a}$. For the sake of simplicity, suppose that we take our functions to be analytic. Then we can write a Taylor series for $f(\mathbf{r} + \mathbf{a})$:

$$f(\mathbf{r} + \mathbf{a}) = \sum_{n=0}^{\infty} \frac{1}{n!} [\mathbf{a} \cdot \nabla]^n f(\mathbf{r}) .$$

Symbolically, this can be written as

$$f(\mathbf{r} + \mathbf{a}) = e^{\mathbf{a} \cdot \nabla} f(\mathbf{r}) ,$$

so we see that

$$\mathscr{T}_\mathbf{a} = e^{\mathbf{a} \cdot \nabla} . \tag{10.9}$$

If one were quantum-mechanically minded, one would write this as

$$\mathscr{T}_\mathbf{a} = e^{i\mathbf{a} \cdot \mathbf{p}/\hbar} , \tag{10.10}$$

where $\mathbf{p} = -i\hbar\nabla$ is the quantum-mechanical momentum operator.

Similarly, if one consider rotations, R_α^z, of a coordinate system about the z-axis, so that $(r, \theta, \phi) \to (r, \theta, \phi - \alpha)$, then the corresponding operator would be

$$\mathscr{R}_\alpha^z = e^{\alpha(\partial/\partial\phi)} = e^{i\alpha L_z/\hbar} , \tag{10.11}$$

where $L_z = -i\hbar\partial/\partial\phi$ is the quantum-mechanical operator for the z-component of angular momentum. For rotation through angles β and γ about the y-axis and x-axis respectively, one has

$$\mathscr{R}_\beta^y = e^{i\beta L_y/\hbar}, \qquad \mathscr{R}_\gamma^x = e^{i\gamma L_x/\hbar}. \tag{10.12}$$

Groups of the type characterized by the collection $\{\mathscr{T}_{\mathbf{a}}\}$ or $\{\mathscr{R}_\alpha^z\}$ are called transformation groups. Both of these examples involve infinite groups. We have been able to parametrize the translation group by $\mathbf{a}$ $(= a_x, a_y, a_z)$, and the group of rotations about the z-axis by α. The parameters in question vary continuously, and the groups are therefore called *continuous* groups.

Now we make contact with our work in earlier chapters. Suppose that the functions which we have been discussing above belong to a vector space, H, and that acting on H we have some linear operator, A_x. Acting on any $f(x) \in H$, A_x gives another vector $g(x) \in H$:

$$A_x f(x) = g(x). \tag{10.13}$$

The subscript x on A serves to remind us that A_x is an operator on a space of functions whose action depends on the point at which the function is evaluated. For example, we might have

$$A_{\mathbf{r}} = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Let $\mathscr{U} \in \mathscr{G}$ act on Eq. (10.13). Then we find that

$$\mathscr{U} A_x f(x) = \mathscr{U} g(x) = g(U^{-1}x).$$

Since $\mathscr{U}^{-1}\mathscr{U} = I$, we have equivalently,

$$\mathscr{U} A_x \mathscr{U}^{-1} \mathscr{U} f(x) = g(U^{-1}x),$$

or

$$\mathscr{U} A_x \mathscr{U}^{-1} f(U^{-1}x) = g(U^{-1}x).$$

But since $A_x f(x) = g(x)$, we have also

$$A_{U^{-1}x} f(U^{-1}x) = g(U^{-1}x),$$

so we conclude that

$$\mathscr{U} A_x \mathscr{U}^{-1} f(U^{-1}x) = A_{U^{-1}x} f(U^{-1}x),$$

for any $f \in H$, or

$$\mathscr{U} A_x \mathscr{U}^{-1} = A_{U^{-1}x}.$$

$\mathscr{U} A_x \mathscr{U}^{-1}$ is called the *transformed* operator; the above equation tells us that the transformed operator at the point x is equal to the untransformed operator at the point $U^{-1}x$. It is often simpler to find the transformed operator by evaluating $A_{x'}$, where $x' = U^{-1}x$, rather than $\mathscr{U} A_x \mathscr{U}^{-1}$. Now if it happens that the transformed operator at the point x equals the untransformed operator at the point x, that is, if

$$\mathscr{U} A_x \mathscr{U}^{-1} = A_x$$

for all $\mathscr{U} \in \mathscr{G}$, then we say that A_x is invariant under the action of the group $\mathscr{G}$. Similarly, if for all $\mathscr{U} \in \mathscr{G}$, $\mathscr{U} f(x) = f(x)$, then we say that f is invariant under $\mathscr{G}$. Note that the criterion for the invariance of A_x can also be written as

$$A_x \mathscr{U} = \mathscr{U} A_x .$$

In other words, if all $\mathscr{U} \in \mathscr{G}$ commute with A_x, then A_x is invariant under $\mathscr{G}$.

It is easily seen that the operator

$$A_{\mathbf{r}} = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

is invariant under $\mathscr{T}_{\mathbf{a}}$ for all $\mathbf{a}$, where $\mathscr{T}_{\mathbf{a}}$ is defined by Eq. (10.9). This follows simply from the fact that the partial derivatives with respect to x, y and z commute with each other, and hence $\exp(\mathbf{a}\cdot\nabla)$ commutes with ∇^2. Thus we may say that the Laplacian is translation invariant. A similar situation prevails for the rotation operators $\mathscr{R}^z_\alpha$, $\mathscr{R}^y_\beta$ and $\mathscr{R}^x_\gamma$ as defined by Eqs. (10.11) and (10.12). According to Eq. (5.89) ∇^2 is simply related to $L^2 = L_x^2 + L_y^2 + L_z^2$, and since L_x, L_y and L_z all commute with L^2 (why?), we see that the rotation operators will also commute with L^2 and hence with ∇^2. Thus, ∇^2 is rotationally invariant.

As an example of a function which is rotationally invariant we mention $f(\mathbf{r}) = g(r)$, where g is any function and r is the magnitude of $\mathbf{r}$. This follows immediately from the fact that L_x, L_y and L_z depend only on the angles θ and ϕ which specify the orientation of $\mathbf{r}$ and not on the magnitude of $\mathbf{r}$ [see Eqs. (5.87)].

To get a feeling for how the above formalism works, it is useful to examine an operator which is *not* invariant under some group. For example, let us consider the operator X_x defined by $X_x f(x) \equiv x f(x)$. We shall investigate how this operator transforms under the x-translation operators $T_a (T_a x = x - a$, for all $x)$. We can compute the transform of X_x in two ways. The simplest method is to make use of the relation

$$\mathscr{T}_a X_x \mathscr{T}_a^{-1} = X_{T_a^{-1}x} .$$

Since $T_a^{-1}x = x + a$, we see immediately that

$$\mathscr{T}_a X_x \mathscr{T}_a^{-1} = X_x + a .$$

We can also proceed by letting $\mathscr{T}_a X_x \mathscr{T}_a^{-1}$ act on an arbitrary function:

$$\begin{aligned} \mathscr{T}_a X_x \mathscr{T}_a^{-1} f(x) &= \mathscr{T}_a X_x f(T_a x) = \mathscr{T}_a X_x f(x-a) \\ &= \mathscr{T}_a x f(x-a) = (x+a) f(x) \\ &= (X_x + a)\, f(x) . \end{aligned}$$

This is exactly the result obtained above; clearly the operator X_x is not invariant under the group of translations. For a slightly more complicated example of the transformation of operators, the reader should look at Problem 10.10.

Now let us look at the eigenvalue problem

$$A_x f(x) = \lambda f(x) . \tag{10.14}$$

We assume that we have a physical problem in mind, so that A_x is self-adjoint, and we also assume that A_x is invariant under some group of operators $\mathscr{G}$, a typical element of which we denote by $\mathscr{U}$. The question now arises: If the operator A_x is invariant under $\mathscr{G}$, are its eigenvectors invariant under $\mathscr{G}$? To answer this question we let $\mathscr{U}$ operate on both sides of the above equation to obtain

$$\mathscr{U} A_x f(x) = \lambda \mathscr{U} f(x) .$$

Since $\mathscr{U} A_x = A_x \mathscr{U}$,

$$A_x \mathscr{U} f(x) = \lambda \mathscr{U} f(x) .$$

This equation tells us that if $f(x)$ is an eigenfunction of A_x belonging to λ, then so is $\mathscr{U} f(x)$. This does *not* mean that $f(x) = \mathscr{U} f(x)$. However, if λ is a non-degenerate eigenvalue, then we conclude that $\mathscr{U} f(x)$ is just a simple multiple of $f(x)$:

$$\mathscr{U} f(x) = D(U)\, f(x) ,$$

where we write $D(U)$ to emphasize that the value of D may depend on which element of the group G is being used.

If, however, λ has multiplicity μ, then we must write

$$A_x f_i(x) = \lambda f_i(x) , \qquad i = 1, 2, \cdots, \mu ,$$
$$A_x \mathscr{U} f_i(x) = \lambda \mathscr{U} f_i(x) , \qquad i = 1, 2, \cdots, \mu .$$

In this case, we must conclude that $\mathscr{U} f_i(x)$ is a linear combination of the elements of the set $\{f_j(x)\}$:

$$\mathscr{U} f_i(x) = \sum_{j=1}^{\mu} D_{ji}(U) f_j(x) . \tag{10.15}$$

Now let us consider $\mathscr{U}_1 \mathscr{U}_2 f_i(x)$, where $\mathscr{U}_1, \mathscr{U}_2 \in \mathscr{G}$. By the previous equation we have

$$\mathscr{U}_1 \mathscr{U}_2 f_i(x) = \mathscr{U}_1 \sum_{j=1}^{\mu} D_{ji}(U_2)\, f_j(x) = \sum_{j=1}^{\mu} D_{ji}(U_2)\, \mathscr{U}_1 f_j(x) .$$

Since

$$\mathscr{U}_1 f_j(x) = \sum_{k=1}^{\mu} D_{kj}(U_1)\, f_k(x) ,$$

we have

$$\mathscr{U}_1 \mathscr{U}_2 f_i(x) = \sum_{j=1}^{\mu} \sum_{k=1}^{\mu} D_{kj}(U_1) D_{ji}(U_2) f_k(x) .$$

But also,

$$\mathscr{U}_1 \mathscr{U}_2 f_i(x) = \sum_{k=1}^{\mu} D_{ki}(U_1 U_2)\, f_k(x) .$$

Combining these last two equations gives the result

$$\sum_{k=1}^{\mu} \left[D_{ki}(U_1 U_2) - \sum_{j=1}^{\mu} D_{kj}(U_1)\, D_{ji}(U_2) \right] f_k(x) = 0 .$$

But the set $\{f_k(x):\ k = 1, 2, \cdots, \mu\}$ is linearly independent, so

$$D_{ki}(U_1U_2) = \sum_{j=1}^{\mu} D_{kj}(U_1)D_{ji}(U_2)$$

for all $i, k = 1, 2, \cdots, \mu$. Thus the matrices $D(U_1), D(U_2), \cdots$ form a group which is homomorphic to the group G and the group $\mathscr{G}$. Such a collection of matrices is called a *representation* of the group G. A representation of a group G is just a homomorphic mapping of G onto a collection of finite-dimensional matrices.

Thus we see that there is a very close relationship between symmetries and degeneracy in physical eigenvalue problems. For example, in quantum mechanics it is often the case that the Hamiltonian which describes some physical system is rotationally invariant. We would then expect the eigenfunctions corresponding to some degenerate energy level to transform among themselves under rotations and thus give rise to a representation of the rotation group. Since degeneracy tends to be the rule rather than the exception in the eigenvalue problems of physics, it is not surprising that the study of group representations has come to play a very important role in physics. It is to the study and classification of such representations that we now turn our attention.

10.5 IRREDUCIBLE REPRESENTATIONS

Let us begin with a few general inferences about the possibility of finding representations of groups. Clearly, for any group there is always one available homomorphism, namely, the mapping which assigns the one-dimensional matrix, 1, to every element of the group. All group properties are trivially satisfied, but this representation is not very useful. However, it is *always* a possible representation and cannot be neglected when we classify representations. Also, if we can find some n-dimensional representation, then a related one-dimensional representation may always be obtained by mapping each element of G into the determinant of its representative matrix. This follows immediately from the relation

$$\det(AB) = (\det A)(\det B) .$$

Another possibility arises when the group, G, has an invariant subgroup, H, whose associated factor group, G/H, has known representations. Since the factor group is homomorphic to G, we can find representations of G by assigning to each element $U \in G$ the matrix representative of that element of G/H onto which U is mapped.

A third interesting representation can be obtained from the group table itself. This is called the *regular* representation and we will make use of it later in this chapter. The basic idea is contained in Theorem 10.1: If we take some element of G, say U_ν, and multiply every group element by U_ν, we rearrange the group elements. Symbolically,

$$U_\nu U_j = U_i , \tag{10.16}$$

where i and j run through all the group elements. We translate this ordering into matrix form by saying that U_ν will be represented by a $g \times g$ matrix, $D_{ij}(U_\nu)$, where g is the order of the group. Each column will contain all zeroes except for one "1," the position of the "1" being determined from Eq. (10.16) as follows: If $U_\nu U_j = U_i$, then the ij-element of $D(U_\nu)$ is 1 and all other elements of the ith row and jth column are zero. This rule is consistent since if $U_\nu U_j = U_i$ we cannot also have $U_\nu U_k = U_i$ unless $j = k$ and, similarly, if $U_\nu U_j = U_i$ we cannot also have $U_\nu U_j = U_k$ unless $i = k$. In this representation the identity is given by the $g \times g$ unit matrix, and all the other elements of G are represented by matrices which have only zeroes on the diagonal.

We now show that this is indeed a product-preserving mapping. Consider the matrix equation

$$M = D(U_\nu)D(U_\mu) .$$

In components this is

$$M_{ij} = \sum_{k=1}^{g} D_{ik}(U_\nu)D_{kj}(U_\mu) , \qquad i, j = 1, 2, \cdots, g .$$

It follows directly from the nature of matrix multiplication that M must be the same type of matrix as $D(U_\nu)$ and $D(U_\mu)$; that is, each row and column of M can have only one nonzero element, which must equal 1. Therefore, we need only check that this one nonzero element is in the proper place. Hence suppose that $M_{ij} = 1$. Therefore, for some unique k,

$$D_{ik}(U_\nu) = 1, \qquad D_{kj}(U_\mu) = 1 .$$

These equations, in turn, mean that

$$U_\nu U_k = U_i , \qquad U_\mu U_j = U_k .$$

Hence

$$U_\nu U_\mu U_j = U_i ,$$

which says that the ij-element of $D(U_\nu U_\mu)$ is unity. Thus we can identify M_{ij} with $D_{ij}(U_\nu U_\mu)$, for *all* i and j. Hence

$$D(U_\nu U_\mu) = D(U_\nu)D(U_\mu) ,$$

so we have a representation, as claimed above.

As an example, we consider the cyclic group of order four, whose multiplication table is given in Table (10.4). We find that if we call $a_1 = e$, $a_2 = a$, $a_3 = b = a^2$, $a_4 = c = a^3$, then, for example,

$$\begin{aligned} aa_1 &= ae = a = a_2 , \\ aa_2 &= aa = b = a_3 , \\ aa_3 &= aa^2 = c = a_4 , \\ aa_4 &= aa^3 = e = a_1 , \end{aligned}$$

so that the 21-, 32-, 43- and 14-elements of $D(a)$ are equal to 1, and all the other elements vanish. A similar calculation gives $D(b)$ and $D(c)$. $D(e)$ is, of course, just the unit matrix. We list the results below:

$$D(e) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad D(a) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$D(b = a^2) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad D(c = a^3) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \tag{10.17}$$

This demonstrates the existence of a representation which is not one-dimensional; the reader can easily prove that these four matrices satisfy Table (10.4). Moreover, since each matrix is distinct, the mapping from the cyclic group is an isomorphism; such a representation is said to be *faithful*. Unfortunately, several problems seem to arise from this example. For one thing, we can use this regular representation to generate many more representations by using similarity transformations, or by using the Kronecker (or direct) product introduced in Section 3.11.

The situation regarding similarity transforms is as follows: If $D(U_i)$ is a representation of G, then for any nonsingular matrix, S, of the same dimensionality as $D(U_i)$, we define

$$\tilde{D}(U_i) \equiv S^{-1}D(U_i)S\,,$$

where the tilde does not indicate the transpose of a matrix, but is merely a label. We then have

$$\tilde{D}(U_i)\tilde{D}(U_j) = S^{-1}D(U_i)SS^{-1}D(U_j)S = S^{-1}D(U_i)D(U_j)S = S^{-1}D(U_iU_j)S\,,$$

since $D(U_i)$ is a representation. Thus

$$\tilde{D}(U_i)\tilde{D}(U_j) = \tilde{D}(U_iU_j)\,,$$

so $\tilde{D}(U_i)$ is also a representation.

In Section 3.11, we showed that if we denote the direct product of A and B by $A \otimes B$, then

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1A_2) \otimes (B_1B_2)\,, \tag{10.18}$$

if A_1 and A_2 (and B_1 and B_2) are of the same dimensionality. This suggests that given any two representations of a group G, say $D(U_i)$ and $D'(U_i)$, another representation can be obtained by the identification

$$U_i \longrightarrow D(U_i) \otimes D'(U_i)\,.$$

This preserves products, since by Eq. (10.18)

$$\begin{aligned}[D(U_i) \otimes D'(U_i)][D(U_j) \otimes D'(U_j)] &= [D(U_i)D(U_j)] \otimes [D'(U_i)D'(U_j)] \\ &= D(U_iU_j) \otimes D'(U_iU_j)\,.\end{aligned}$$

Thus we have indeed constructed another representation whose dimension will be the product of the dimensions of D and D'. Since we have already found one specific multidimensional representation, the regular representation, we can use direct products to generate an arbitrarily large number of other representations. Hence the problem of classifying representations would appear to be hopeless.

Before we accept such a pessimistic conclusion, however, it is a good idea to look at the problem from a different point of view. Let us start with some simple representations and then try to work up to more complicated ones. First, we must decide how to go about finding representations. In the previous section we have already seen that one way of constructing them is to find a collection of linearly independent functions which transform among themselves under the action of the group, that is, a collection of functions $\{f_i(x)\}$ which satisfy

$$\mathscr{U} f_i(x) = \sum_j D_{ji}(U)\, f_j(x)$$

for all $\mathscr{U} \in \mathscr{G}$.

Now the only problem is to decide how to find a set of functions which has this nice property of transforming into itself under the action of a group of operators. One possible way would be just to pick an arbitrary function $f(x)$ and operate on it with all the $\mathscr{U}_i \in \mathscr{G}$. In this way we find a set $\{f_i(x)\}$, where

$$f_i(x) \equiv \mathscr{U}_i f(x) .$$

By the group property, this set satisfies the basic requirement of closure:

$$\mathscr{U}_\nu f_j(x) = \mathscr{U}_\nu \mathscr{U}_j f(x) = \mathscr{U}_i f(x) = f_i(x) ,$$

where i is determined by the multiplication table. If all elements of the set $\{f_i(x)\}$ are linearly independent, then we have succeeded in our aims. In this case,

$$\mathscr{U}_\nu f_j(x) = \sum_k D_{kj}(U_\nu)\, f_k(x) ,$$

where

$$D_{kj}(U_\nu) = \delta_{ik} ,$$

i being determined by $\mathscr{U}_\nu \mathscr{U}_j = \mathscr{U}_i$. Looking back a few paragraphs, we find that this is just the regular representation. However, this is by no means the only possibility. It might well happen that the $f_i(x)$ are not linearly independent. Then we must pick out a linearly independent subset which will, of course, also be invariant under the group operations (Why?). Carrying out the above process, we obtain in this case a representation of smaller dimension than the regular representation. Of course, it is often not necessary to do this, for one may well be able to spot a set of functions which transform among themselves without recourse to this unintuitive formalism.

Example 10.9. A particularly straightforward example is the group of rotations about the z-axis. This is an infinite, continuous group, but that is no reason to be daunted. It is, in fact, very easy to analyze this group.

Now when one talks about angles and rotations, the functions which come to mind are sines and cosines. Denoting a rotation through α by R_α ($R_\alpha \phi = \phi - \alpha$), we consider $f_1(\phi) = \cos\phi$ and $f_2(\phi) = \sin\phi$. Then

$$f_1(R_\alpha^{-1}\phi) = \cos(\phi + \alpha) = \cos\alpha f_1(\phi) - \sin\alpha f_2(\phi) ,$$
$$f_2(R_\alpha^{-1}\phi) = \sin(\phi + \alpha) = \sin\alpha f_1(\phi) + \cos\alpha f_2(\phi) .$$

From this we conclude, using Eq. (10.15), that

$$D^{(1)}(R_\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} .$$

(We use the superscript to distinguish between various representations of this group.) A simple multiplication of matrices shows that

$$D^{(1)}(R_\alpha)D^{(1)}(R_\beta) = D^{(1)}(R_{\alpha+\beta}) = D^{(1)}(R_\beta)D^{(1)}(R_\alpha) = D^{(1)}(R_\alpha R_\beta) ,$$

so we have found a two-dimensional representation.

Of course, $\cos\phi$ and $\sin\phi$ are not the only choice of functions we could have made. For instance, $e^{i\phi} = f(\phi)$ is such that

$$f(R_\alpha^{-1}\phi) = e^{i(\phi+\alpha)} = e^{i\alpha}e^{i\phi} = e^{i\alpha}f(\phi) .$$

Thus we obtain a one-dimensional representation:

$$D^{(2)}(R_\alpha) = e^{i\alpha} .$$

Similarly, if we choose $g(\phi) = e^{-i\phi}$, we find that

$$D^{(3)}(R_\alpha) = e^{-i\alpha} .$$

In fact it is clear that the function $f(\phi) = e^{im\phi}$ leads directly to a representation $D(R_\alpha) = e^{im\alpha}$. We choose m to be a positive or negative integer (or zero) so that $D(R_0) = D(R_{2\pi})$. Now since the functions $f_1(\phi) = e^{i\phi}$ and $f_2(\phi) = e^{-i\phi}$ are linearly independent we can use them together to form a set $\{f_i\}$. This leads to the two-dimensional representation

$$D^{(4)}(R_\alpha) = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} .$$

This example brings us to the heart of the basic problem of representation theory. The two-dimensional representation $D^{(4)}(R_\alpha)$ is clearly just two one-dimensional representations joined together. In some sense, we would not want to call this a *basic* representation since it is compounded out of representations of lower dimensionality. The situation is complicated further if we look back at $D^{(1)}(R_\alpha)$. This does not, on the surface of things, look much like $D^{(4)}(R_\alpha)$, but if we diagonalize $D^{(1)}$, we find that

$$S^{-1}D^{(1)}(R_\alpha)S = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} = D^{(4)}(R_\alpha) ,$$

where

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

is a matrix independent of α. Thus we have a similarity transformation which will bring $D^{(1)}(R_\alpha)$ to exactly the form of $D^{(4)}(R_\alpha)$, and this can be done simultaneously for *all* R_α. Thus, even $D^{(1)}(R_\alpha)$ is essentially a sum of the two one-dimensional representations, $D^{(2)}$ and $D^{(3)}$.

This can also be the case for more complicated situations. For example, the four matrices composing the regular representation of the four-element cyclic group [Eq. (10.17)] can be diagonalized simultaneously by the matrix

$$S' = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & i & -i \\ 1 & 1 & -1 & -1 \\ -1 & 1 & -i & i \end{bmatrix}.$$

Therefore $D(e)$, $D(a)$, $D(b)$, and $D(c)$ are each sums of four one-dimensional representations.

We can generalize this as follows: If we find a representation $D(U_i)$ which can be brought to the form

$$\bar{D}(U_i) \equiv SD(U_i)S^{-1} = \left[\begin{array}{c|c|c} D^{(1)}(U_i) & & \mathrm{O} \\ \hline & U^{(2)}(D_i) & \\ \hline \mathrm{O} & & D^{(3)}(U_i) \end{array}\right], \tag{10.19}$$

where S is the same for all U_i, then we will not consider it as a basic representation, but rather as a sum of representations.

In terms of the functions $f_i(x)$, this means that although we have found a linearly independent set $\{f_i\}$, the elements of which transform among themselves, there are actually smaller subsets which also transform into themselves under the group in question.

Representations of the type of Eq. (10.19) are called *reducible* representations; those which cannot be reduced in this manner are called *irreducible* representations. It is clearly the irreducible representations of groups that we should study, since other representations are simply built up out of them. However, even among irreducible representations we do not want to count the similarity transformation of a representation as a separate representation. Two representations which are related by a similarity transform are said to be *equivalent*. Thus, in Eq. (10.19), assuming $D^{(1)}$, $D^{(2)}$, and $D^{(3)}$ to be irreducible and inequivalent, we would write symbolically

$$D(U_i) = D^{(1)}(U_i) \oplus D^{(2)}(U_i) \oplus D^{(3)}(U_i)$$

for all $U_i \in G$. In practice, when we reduce a representation to this form, a given irreducible representation may occur more than once. Thus we would write generally

$$D(U_i) = a^{(1)}D^{(1)}(U_i) \oplus a^{(2)}D^{(2)}(U_i) \oplus \cdots \oplus a^{(N)}D^{(N)}(U_i)\,, \tag{10.20}$$

where $a^{(\nu)}$ is the number of times the irreducible representation $D^{(\nu)}(U_i)$, or any representation equivalent to it, appears in the decomposition.

A very useful characterization of a representation which is the same for all equivalent representations, is given by the *trace* of the representative matrices. The trace of $D(U)$ is denoted by $\chi(U)$ and is called the *character* of U in the representation in question. Since

$$\operatorname{tr}[C^{-1}D(U)C] = \operatorname{tr}[D(U)] \equiv \chi(U) ,$$

the character of U is the same for all equivalent representations. When we are dealing with an irreducible representation, $D^{(\nu)}(U)$, we will write for the corresponding character $\chi^{(\nu)}(U)$. Now within a given representation, many of the characters will be the same. In fact, consider a class, $\mathscr{C}_i$, of conjugate elements of the group G. Then if U_ν, $U_\mu \in \mathscr{C}_i$, there exists some $U_\lambda \in G$ such that

$$U_\nu = U_\lambda U_\mu U_\lambda^{-1} .$$

This means that

$$D(U_\nu) = D(U_\lambda)D(U_\mu)D(U_\lambda^{-1}) ,$$

and hence

$$\begin{aligned}\chi(U_\nu) &\equiv \operatorname{tr}[D(U_\nu)] = \operatorname{tr}[D(U_\lambda)D(U_\mu)D(U_\lambda^{-1})] \\ &= \operatorname{tr}[D(U_\lambda^{-1})D(U_\lambda)D(U_\mu)] = \operatorname{tr}[D(U_\mu)] \equiv \chi(U_\mu) ,\end{aligned}$$

that is, all elements of $\mathscr{C}_i$ have the same character. For this reason one can, with no ambiguity, refer to $\chi_i^{(\nu)}$ as the character of the ith class in the νth irreducible representation, thereby removing the necessity of referring explicitly to the group elements. Thus in a group with k classes, for a given representation, ν, there are k characteristic numbers,

$$\chi_1^{(\nu)}, \chi_2^{(\nu)}, \cdots, \chi_k^{(\nu)}$$

to be determined. These k constants, in fact, provide a surprisingly large amount of information about the group, as we will see in the following sections.

10.6 UNITARY REPRESENTATIONS, SCHUR'S LEMMAS, AND ORTHOGONALITY RELATIONS

In this section we will develop the main results of representation theory, which will enable us to determine irreducible representations. To begin, we prove a theorem which restricts tremendously the possible forms which a representation can take.

Theorem 10.6. Every representation of a finite group is equivalent to a unitary representation.

Proof. The first step is to construct an inner product in the n-dimensional vector space, V, of the representation and show that the representation is unitary with respect to this inner product. For $x, y \in V$, we denote the usual inner product of x and y by the familiar (x, y); in terms of this inner product, we define a new inner product as follows:

$$\{x, y\} = \sum_{U \in G} (D(U)x, D(U)y) , \tag{10.21}$$

where D is any representation of G. It is easy to check that since (,) is an inner product and since $D^{-1}(U)$ exists for all $U \in G$, { , } will also satisfy all the inner-product axioms. Consider now any $D(U')$. We want to show that

$$\{D(U')x, D(U')y\} = \{x, y\} .$$

From Eq. (10.21), we have

$$\begin{aligned}\{D(U')x, D(U')y\} &= \sum_{U \in G} (D(U)D(U')x, D(U)D(U')y) \\ &= \sum_{U \in G} (D(UU')x, D(UU')y) .\end{aligned} \tag{10.22}$$

But as U runs through all the elements of G, so does UU', according to Theorem 10.1. Of course, the ordering may be different, but that is immaterial. Hence

$$\sum_{U \in G} (D(UU')x, D(UU')y) = \sum_{\tilde{U} \in G} (D(\tilde{U})x, D(\tilde{U})y) = \{x, y\} . \tag{10.23}$$

Combining Eqs. (10.22) and (10.23), we get

$$\{D(U')x, D(U')y\} = \{x, y\}$$

for any $U' \in G$, as required.

But to say that $D(U')$ is unitary with respect to some special inner product is simply to say that we have written $D(U')$ in an inconvenient basis. By a similarity transformation we can easily correct this difficulty and make D unitary with respect to the original inner product. Let $\{\xi_i\}$ be a complete orthonormal basis with respect to the inner product (,) and let $\{\eta_i\}$ be a complete orthonormal basis with respect to the special inner product { , }. We define a linear operator S on V by

$$\eta_i = S\xi_i , \qquad i = 1, 2, \cdots, n . \tag{10.24}$$

Note that with this definition for any $x, y \in V$

$$\{Sx, Sy\} = (x, y) , \tag{10.25a}$$

or

$$(S^{-1}x, S^{-1}y) = \{x, y\} . \tag{10.25b}$$

This follows by expanding x and y in the $\{\xi_i\}$ basis:

$$x = \sum_{i=1}^{n} a_i\xi_i , \qquad y = \sum_{j=1}^{n} b_j\xi_j .$$

Then

$$\{Sx, Sy\} = \sum_{i,j=1}^{n} a_i^* b_j \{S\xi_i, S\xi_j\} = \sum_{i,j=1}^{n} a_i^* b_j \{\eta_i, \eta_j\} ,$$

by use of Eq. (10.24). Since the η_i form an orthonormal set with respect to $\{\ ,\ \}$, we have

$$\{Sx, Sy\} = \sum_{i,j=1}^{n} a_i^* b_j \delta_{ij} = \sum_{i=1}^{n} a_i^* b_i = (x, y) ,$$

which is just Eq. (10.25a). S is precisely the operator we need to transform the original representation.

For any $U \in G$, we define

$$\tilde{D}(U) \equiv S^{-1} D(U) S .$$

Consider

$$(\tilde{D}(U)x, \tilde{D}(U)y) = (S^{-1}D(U)Sx, S^{-1}D(U)Sy) .$$

By Eq. (10.25b), we have immediately

$$(\tilde{D}(U)x, \tilde{D}(U)y) = \{D(U)Sx, D(U)Sy\} .$$

But with respect to $\{\ ,\ \}$ $D(U)$ is unitary, so

$$(\tilde{D}(U)x, \tilde{D}(U)y) = \{Sx, Sy\} .$$

Finally, Eq. (10.25a) gives

$$(\tilde{D}(U)x, \tilde{D}(U)y) = (x, y) ,$$

so $\tilde{D}(U)$ is unitary with respect to the original inner product.

According to this theorem, we can restrict ourselves to unitary matrices in our search for irreducible representations. Note that although we have stated this theorem for finite groups, we have made very little use of the finite order of G, using it only in Eq. (10.21) to write the sum over all $U \in G$. This sum is finite since it contains only a finite number of terms. For many infinite groups, it is possible to make the transition from discrete sums to integrals over a continuous parameter; for example, in the case of continuous groups such a possibility is evident. However, even if it is possible to replace sums over group elements by integrals over group elements, it may be that the integral formulation of Eq. (10.21) will not be meaningful because the integrations run over an infinite range and are divergent. Nevertheless, there are many groups, such as the rotation groups, where this does not happen, and for them Theorem 10.6 is valid, the proof being identical to the one presented here, save for the presence of integrals instead of summations.

The fact that the matrices in a representation are equivalent to unitary representations is a very powerful result. From it follows the two main theorems of representation theory, the first and second lemmas of Schur. The proof of these two important group theoretical results hinges on the following lemma, which has nothing to do with group theory.

Lemma. If a matrix, M, commutes with a unitary matrix, U, and if we write $M = M_+ + iM_-$, where $M_+ = (M + M^\dagger)/2$ and $M_- = (M - M^\dagger)/2i$, then both M_+ and M_- commute with U.

Proof. If

$$UM = MU, \tag{10.26}$$

then

$$M^\dagger U^\dagger = U^\dagger M^\dagger .$$

Multiplying on the left by U, we get

$$UM^\dagger U^\dagger = UU^\dagger M^\dagger = M^\dagger ,$$

since U is unitary. Multiplying this last equation on the right by U, we find that

$$UM^\dagger U^\dagger U = M^\dagger U .$$

Thus, using the unitarity of U once more,

$$UM^\dagger = M^\dagger U . \tag{10.27}$$

Combining Eqs. (10.26) and (10.27), we readily obtain

$$U\left[\frac{M + M^\dagger}{2}\right] = \left[\frac{M + M^\dagger}{2}\right]U, \tag{10.28a}$$

$$U\left[\frac{M - M^\dagger}{2i}\right] = \left[\frac{M - M^\dagger}{2i}\right]U, \tag{10.28b}$$

or

$$UM_+ = M_+U, \qquad UM_- = M_-U,$$

as required.

With this lemma at our disposal, we now prove the main result of this section.

> **Lemma (Schur).** If $D(U)$ is an element of an irreducible representation of G and if $D(U)M = MD(U)$ for all $U \in G$, then M must be a multiple of the unit matrix.

Proof. According to the previous lemma, if $MD(U) = D(U)M$ for all $U \in G$ and if we write

$$M = M_+ + iM_- , \tag{10.29}$$

then since the representation can be taken to be unitary,

$$M_+D(U) = D(U)M_+ \qquad \text{and} \qquad M_-D(U) = D(U)M_-$$

for all $U \in G$. $M_+ = (M + M^\dagger)/2$ and $M_- = (M - M^\dagger)/2i$ are both self-adjoint, so let us begin by considering the eigenvalue problem associated with M_+:

$$M_+x_n^{(i)} = \lambda_n x_n^{(i)} , \qquad i = 1, 2, \cdots, m_n ,$$

where m_n is the multiplicity of λ_n. Since M_+ is Hermitian, the set of all $x_n^{(i)}$ spans the N-dimensional space of the representation. For all $U \in G$, we have

$$D(U)M_+x_n^{(i)} = \lambda_n D(U)x_n^{(i)} .$$

But $[D(U), M_+] = 0$, so this is just

$$M_+[D(U)x_n^{(i)}] = \lambda_n[D(U)x_n^{(i)}] .$$

Since $D(U)x_n^{(i)}$ is an eigenvector of M_+ belonging to λ_n, it must be a linear combination of the $x_n^{(i)}$ $(i = 1, 2, \cdots, m_n)$:

$$D(U)x_n^{(i)} = \sum_{j=1}^{m_n} a_{ij}(U)x_n^{(j)} \tag{10.30}$$

for all $U \in G$. Let us compute the matrix D in the space spanned by the eigenvectors of M_+. We take $n = 1$ first. Using Eq. (10.30), we find that

$$(x_1^{(l)}, D(U)x_1^{(i)}) = \sum_{j=1}^{m_1} a_{ij}(U)(x_1^{(l)}, x_1^{(j)}) = a_{il}(U) , \qquad i, l = 1, 2, \cdots, m_1 , \tag{10.31a}$$

for all $U \in G$. For any $n \neq 1$, Eq. (10.30) tells us directly that

$$(x_n^{(l)}, D(U)x_1^{(i)}) = 0 ; \qquad l = 1, 2, \cdots, m_n , \quad i = 1, 2, \cdots, m_1 , \tag{10.31b}$$

for all $U \in G$, since the eigenvectors belonging to different eigenvalues of an Hermitian operator are orthogonal. Similarly, for $n \neq 1$, we also have from Eq. (10.30)

$$(x_1^{(l)}, D(U)x_n^{(i)}) = 0; \qquad l = 1, 2, \cdots, m_1 , \quad i = 1, 2, \cdots, m_n . \tag{10.31c}$$

Equations (10.31a), (10.31b), and (10.31c) mean that the matrix $D(U)$ looks like

$$D(U) = \left(\begin{array}{c|c} m_1 \times m_1 & 0 \\ \hline 0 & (N - m_1) \times (N - m_1) \end{array}\right),$$

for all $U \in G$, where N is the dimension of the representation. The $m_1 \times m_1$ matrix in the upper left-hand corner has elements which are given by the $a_{ij}(U)$ defined in Eq. (10.30). But the representation is assumed to be irreducible, so this can only be the case if $m_1 = N$; that is, the multiplicity of λ_1 must be equal to N, the dimensionality of the representation. The only Hermitian matrix having an eigenvalue whose multiplicity equals the dimension of the matrix is a real constant times the unit matrix. Thus $M_+ = c_+ I$. Similarly, $M_- = c_- I$, so we obtain from Eq. (10.29) $M = (c_+ + ic_-)I$, and the proof is complete.

As an immediate corollary (really no more than a restatement) of this lemma, we have:

Corollary. If there exists a matrix which commutes with every element of a representation and is *not* a multiple of the unit matrix, then the representation is necessarily reducible.

At this point it is perhaps a good idea to say a word about a more abstract way of looking at irreducible representations which may clarify the situation, as well as boil down the preceding pages into a few words.

As a careful reading of the previous section shows, we can characterize the irreducible representations of a group, G, of order g in a more abstract manner as follows. Let

$$\mathscr{G} = \{D(U_i), \quad i = 1, 2, \cdots, g\}$$

be a group of linear operators on an n-dimensional vector space, V, and let this group of operators be homomorphic to G, that is, $\mathscr{G}$ is a representation of G. $\mathscr{G}$ is said to be an *irreducible* representation of G if no *proper* subspace of V is left invariant (i.e., mapped into itself) under all $D \in \mathscr{G}$. (The proper subspaces of V are all subspaces of V *except* for V itself.) It may of course happen that some *subset* of $\mathscr{G}$ leaves invariant some proper subspace of V, but if we have an irreducible representation, then $\mathscr{G}$ in its entirety must leave only V invariant. Obviously, if we look just at one unitary $D(U) \in \mathscr{G}$, then it will have n invariant subspaces, defined by the n orthonormal eigenvectors of $D(U)$.

In light of this, Schur's first lemma has a very simple meaning: If there is a matrix which commutes with $D(U)$ for all $U \in G$, the linearly independent eigenvectors of this matrix define a collection of subspaces of V which are invariant under $D(U)$ for all $U \in G$. The only admissible invariant subspace is V itself, and the only matrix which has every vector in V as an eigenvector is the unit matrix or a multiple thereof.

A simple, but very useful, consequence of Schur's first lemma is the fact that an abelian group can have only one-dimensional irreducible representations! This follows immediately from the statement that if G is abelian, then for any $U \in G$, $D(U)$ commutes with all elements of the representation D. Thus, by Schur's first lemma, for all $U \in G$,

$$D(U) = C(U)I,$$

where $C(U)$ is a constant which can depend on U. The unit matrix is irreducible only when it is the one-dimensional unit matrix. Thus, for an abelian group, all irreducible representations must be one-dimensional.

Schur's second lemma gives an analogous result for two *different* representations.

Lemma. Let $D(U)$ and $D'(U)$ be irreducible representations. If for all $U \in G$, $D(U)M = MD'(U)$, then either D and D' are equivalent or else $M = 0$.

Proof. In general, D and D' will have different dimensions, so that M need not be a square matrix. If D is $n \times n$ and D' is $m \times m$, then M must be $n \times m$. The proof starts out in the same direction as that of the first lemma. As before

$$D(U)M = MD'(U) \tag{10.32}$$

for all $U \in G$ implies that

$$M^\dagger D^\dagger(U) = D'^\dagger(U)M^\dagger$$

for all $U \in G$, where $M^\dagger$ is an $m \times n$ matrix, $(M^\dagger)_{ij} = M^*_{ji}$. Taking $D(U)$ to be unitary, we find that

$$M^\dagger D(U^{-1}) = D'(U^{-1})M^\dagger. \tag{10.33}$$

Multiplying Eq. (10.33) from the left by M, we get

$$MM^\dagger D(U^{-1}) = MD'(U^{-1})M^\dagger . \tag{10.34}$$

But by Eq. (10.32),

$$MD'(U^{-1})M^\dagger = D(U^{-1})MM^\dagger ,$$

so Eq. (10.34) becomes

$$MM^\dagger D(U^{-1}) = D(U^{-1})MM^\dagger$$

for all $U \in G$. The previous lemma (and the uniqueness of inverses) tells us straightaway that

$$MM^\dagger = cI .$$

We leave it to the reader to show in the same way that multiplying Eq. (10.33) on the right by M leads to

$$M^\dagger M = cI .$$

It should be observed that the matrices $MM^\dagger$ and $M^\dagger M$ have different dimensions, the former being $n \times n$, the latter $m \times m$.

We now consider three cases separately:

(i) $m = n$. Then if $c = 0$, we have $MM^\dagger = 0$, which means that for all i,

$$(MM^\dagger)_{ii} = 0 = \sum_{j=1}^{n} M_{ij}(M^\dagger)_{ji} = \sum_{j=1}^{n} |M_{ij}|^2 .$$

Thus, for all i,

$$\sum_{j=1}^{n} |M_{ij}|^2 = 0 ,$$

so we conclude that $M_{ij} = 0$, for all i and j, and $M = 0$. (For a slightly stronger result see Problem 3.14) If, on the other hand, $c \neq 0$, then

$$\det (MM^\dagger) \neq 0 ,$$

so $\det M \neq 0$, and M is invertible. Thus according to Eq. (10.32),

$$D(U) = MD'(U)M^{-1}$$

for all $U \in G$, therefore D and D' are equivalent representations.

(ii) $n > m$. In this case, we make M into a square matrix by filling in $n - m$ columns of zeros. Calling the new matrix N, we have

$$N = \begin{array}{cc} \begin{array}{cc} \xleftrightarrow{\quad m \quad} & \xleftrightarrow{\; n-m \;} \end{array} & \\ \left(\begin{array}{c:c} \quad M \quad & \quad 0 \quad \end{array}\right) & \updownarrow n \end{array} , \qquad N^\dagger = \begin{array}{cc} \xleftrightarrow{\qquad n \qquad} & \\ \left(\begin{array}{c} M^\dagger \\ \hdashline 0 \end{array}\right) & \begin{array}{c} \updownarrow m \\ \updownarrow n-m \end{array} \end{array} .$$

It is clear that $MM^\dagger = NN^\dagger$, so $MM^\dagger = cI$ implies that $NN^\dagger = cI$. Now

$$\det (NN^\dagger) = c^n = (\det N)(\det N^\dagger) ,$$

but from the form of N, we see that $\det N = 0$, so the constant, c, must vanish. Thus

$$NN^\dagger = 0 = MM^\dagger$$

and, just as in case (i), we conclude that $M = 0$.

(iii) $m > n$. In this case, we must add $m - n$ rows of zeros to M to turn it into a square matrix. Again denoting the new matrix by N, we have

$$N = \begin{array}{c} \longleftarrow m \longrightarrow \\ \left(\begin{array}{c} M \\ \hdashline 0 \end{array}\right) \end{array} \begin{array}{l} \updownarrow n \\ \updownarrow m - n \end{array}, \qquad N^\dagger = \begin{array}{cc} \longleftarrow n \longrightarrow & \longleftarrow m-n \longrightarrow \\ \multicolumn{2}{c}{\left(\begin{array}{c:c} M^\dagger & 0 \end{array}\right)} \end{array} \updownarrow m \; .$$

In this case, $NN^\dagger \neq MM^\dagger$, but we do have $N^\dagger N = M^\dagger M$. Thus, using the relation $M^\dagger M = cI$ instead of $MM^\dagger = cI$, we see that the new matrix N satisfies

$$N^\dagger N = cI \, ;$$

as in case (i), we conclude that $N = 0$. Hence $M = 0$ as well, and the proof of the lemma is complete.

Note that in proving these lemmas, we have not made direct reference to the number of elements in the group G. We have actually proved these results for any finite-dimensional *unitary* representation. Of course, Theorem 10.6 tells us that for a finite group the representations can always be chosen to be unitary. But Schur's two lemmas hold also for the important class of infinite groups which have finite-dimensional unitary representations.

The reason that these two lemmas are so important is that it turns out to be rather easy to find matrices, M, which satisfy the conditions of the lemmas. These matrices in turn provide us with a vast amount of information about the irreducible representations and their characters. One simple example of a matrix which satisfies the conditions of Schur's first lemma is defined by

$$M_i^{(\nu)} = \sum_{U \in \mathscr{C}_i} D^{(\nu)}(U) \; . \tag{10.35}$$

$M_i^{(\nu)}$ is just the sum of all representative matrices of the elements of the class $\mathscr{C}_i$ in the νth irreducible representation. Now let U' be any element of G and consider the quantity

$$D^{(\nu)}(U')M_i^{(\nu)}D^{(\nu)}(U')^{-1} \; .$$

We have

$$\begin{aligned} D^{(\nu)}(U')M_i^{(\nu)}D^{(\nu)}(U')^{-1} &= \sum_{U \in \mathscr{C}_i} D^{(\nu)}(U')D^{(\nu)}(U)D^{(\nu)}(U'^{-1}) \\ &= \sum_{U \in \mathscr{C}_i} D^{(\nu)}(U'UU'^{-1}) \; . \end{aligned}$$

But by the definition of a class, if $U \in \mathscr{C}_i$, then $U'UU'^{-1} \in \mathscr{C}_i$. Since a similarity transformation just rearranges the elements of $\mathscr{C}_i$, we have from the previous equation

$$D^{(\nu)}(U')M_i^{(\nu)}D^{(\nu)}(U')^{-1} = \sum_{\tilde{U}\in\mathscr{C}_i} D^{(\nu)}(\tilde{U}) = M_i^{(\nu)} .$$

Thus, for all $U' \in G$,

$$D^{(\nu)}(U')M_i^{(\nu)} = M_i^{(\nu)}D^{(\nu)}(U') ,$$

so $M_i^{(\nu)}$ satisfies the requirements of Schur's first lemma. We conclude that

$$M_i^{(\nu)} = c_i^{(\nu)}I .$$

The constant $c_i^{(\nu)}$ can be readily evaluated. We have, calling the dimension of the νth representation n_ν,

$$\operatorname{tr} M_i^{(\nu)} = c_i^{(\nu)}n_\nu . \tag{10.36}$$

But

$$\operatorname{tr} M_i^{(\nu)} = \sum_{U\in\mathscr{C}_i} \operatorname{tr} D^{(\nu)}(U) = \sum_{U\in\mathscr{C}_i} \chi^{(\nu)}(U) .$$

The trace of all the representative matrices in a given class is the same, as was pointed out at the end of the last section. Hence using the notation established there, we have

$$\operatorname{tr} M_i^{(\nu)} = g_i\chi_i^{(\nu)} , \tag{10.37}$$

where g_i is the number of elements in the class $\mathscr{C}_i$. Combining Eqs. (10.36) and (10.37), we obtain

$$c_i^{(\nu)} = \frac{g_i}{n_\nu}\chi_i^{(\nu)} .$$

Therefore

$$M_i^{(\nu)} = \frac{g_i}{n_\nu}\chi_i^{(\nu)}I , \tag{10.38}$$

that is, the sum of all the representative matrices in a given class is a simple multiple of the unit matrix. We shall soon have occasion to make use of this result.

A more fruitful example of a matrix which fulfills all the requirements of Schur's lemmas is the following:

$$M = \sum_{U\in G} D^{(\nu)}(U)XD^{(\mu)}(U^{-1}) = \sum_{U\in G} D^{(\nu)}(U)XD^{(\mu)}(U)^{-1} ,$$

where $D^{(\nu)}$ and $D^{(\mu)}$ are irreducible representations and X is an arbitrary matrix, although in order for this expression to have any meaning, X must have n_ν rows and n_μ columns. Multiplying on the left by $D^{(\nu)}(U')$ and on the right by $D^{(\mu)}(U')^{-1}$, we find that

$$\begin{aligned} D^{(\nu)}(U')MD^{(\mu)}(U')^{-1} &= \sum_{U\in G} D^{(\nu)}(U')D^{(\nu)}(U)XD^{(\mu)}(U)^{-1}D^{(\mu)}(U')^{-1} \\ &= \sum_{U\in G} D^{(\nu)}(U'U)XD^{(\mu)}(U^{-1}U'^{-1}) . \end{aligned}$$

Calling $\tilde{U} = U'U$ and remembering that $\tilde{U}^{-1} = U^{-1}U'^{-1}$, we have

$$D^{(\nu)}(U')MD^{(\mu)}(U')^{-1} = \sum_{U \in G} D^{(\nu)}(\tilde{U})XD^{(\mu)}(\tilde{U})^{-1} . \tag{10.39}$$

But as U runs through the entire group, so does $\tilde{U} = U'U$, although in a different order. Thus the right-hand side of Eq. (10.39) is just equal to M, and

$$D^{(\nu)}(U')M = MD^{(\mu)}(U') \tag{10.40}$$

for all $U' \in G$. Assuming that $D^{(\nu)}$ is not equivalent to $D^{(\mu)}$ (in the case $n_\nu = n_\mu$), we have by Schur's lemmas

$$M = c(X)I\delta_{\nu\mu} .$$

In other words, $M = 0$ if the representations have different dimensionality (Schur's second lemma) and $M = cI$ if they have the same dimensionality (Schur's first lemma). The value of the constant depends, of course, on the choice of X. Thus we arrive at the result

$$\sum_{U \in G} D^{(\nu)}(U)XD^{(\mu)}(U)^{-1} = c(X)I\delta_{\mu\nu} . \tag{10.41}$$

If $D^{(\nu)}$ and $D^{(\mu)}$ are chosen to be unitary, this can also be written as

$$\sum_{U \in G} D^{(\nu)}(U)XD^{(\mu)\dagger}(U) = c(X)I\delta_{\mu\nu} . \tag{10.42}$$

To exploit this equation, let us write it in component form:

$$\sum_{U \in G} \sum_{i',l'} D^{(\nu)}_{ii'}(U)\; X_{i'l'}\,[D^{(\mu)\dagger}(U)]_{l'l} = c(X)\delta_{il}\delta_{\mu\nu} .$$

Now suppose that we choose the matrix X so that all its elements are zero except for the jk-element. Denoting the related $c(X)$ by c_{jk}, we have

$$\sum_{U \in G} D^{(\nu)}_{ij}(U)[D^{(\mu)\dagger}(U)]_{kl} = c_{jk}\delta_{il}\delta_{\mu\nu} . \tag{10.43}$$

Since

$$(A^\dagger)_{kl} = A^*_{lk} ,$$

we obtain from Eq. (10.43)

$$\sum_{U \in G} D^{(\mu)*}_{lk}(U)D^{(\nu)}_{ij}(U) = c_{jk}\delta_{il}\delta_{\mu\nu} , \tag{10.44}$$

for all i, j, k, l, μ, and ν.

All that remains is to determine c_{jk}. To do this, we set $\mu = \nu$, $l = i$ and sum on i. Thus

$$\sum_{U \in G} \sum_{i} D^{(\nu)*}_{ik}(U)D^{(\nu)}_{ij}(U) = c_{jk} \sum_{i} \delta_{ii} = c_{jk}n_\nu .$$

But since $D^{(\nu)}$ is unitary, the sum on i on the left-hand side of this equation is just equal to δ_{kj}. Hence

$$c_{jk}n_\nu = \sum_{U \in G} \delta_{jk} = g\delta_{jk} ,$$

where g is the order of the group. Therefore

$$c_{jk} = \frac{g}{n_\nu}\delta_{jk} ,$$

and we have finally from Eq. (10.44) the remarkable orthogonality theorem,

$$\sum_{U \in G} D_{lk}^{(\mu)*}(U) D_{ij}^{(\nu)}(U) = \frac{g}{n_\nu}\delta_{jk}\delta_{il}\delta_{\mu\nu} , \tag{10.45}$$

for all i, j, k, l, μ, and ν.

Now let us see what this means in simple terms. The sum on the left-hand side of Eq. (10.45) looks very much like an inner product of two vectors, except that the sum is over the elements of a group rather than a conventional subscript. However, there is no law against indexing the components of a vector by using the elements of a group. We may say, therefore, that Eq. (10.45) is an orthogonality relationship for the inner product of two g-dimensional vectors, each vector being identified by three labels. Equation (10.45) says that if, from any irreducible representation we pick the ij-element of the representative matrices, then this gives us a g-dimensional vector which is orthogonal to the g-dimensional vector obtained from any other element, say the $i'j'$-element, and *also* to any vector obtained in this way from any *other* representation. Thus in the g-dimensional space we have $n_1^2 + n_2^2 + \cdots + n_N^2$ orthogonal vectors, where N is the number of inequivalent irreducible representations. Since the number of orthogonal vectors in a g-dimensional space cannot exceed g, we must have

$$\sum_{\nu=1}^{N} n_\nu^2 \leq g . \tag{10.46}$$

Thus we arrive at the important result that a finite group can have only a finite number of inequivalent irreducible representations, all of which must have dimension less than $\sqrt{g}$.

From Eq. (10.45) it is a simple matter to obtain an orthogonality relation for the characters of a representation. In Eq. (10.45), if one sets $k = l$ and sums on l, one finds that

$$\sum_{U \in G} \sum_{l=1}^{n_\mu} D_{ll}^{(\mu)*}(U) D_{ij}^{(\nu)}(U) = \frac{g}{n_\nu}\delta_{\mu\nu} \sum_{l=1}^{n_\mu} \delta_{il}\delta_{lj} = \frac{g}{n_\nu}\delta_{\mu\nu}\delta_{ij} .$$

But

$$\sum_{l} D_{ll}^{(\mu)*}(U) = \operatorname{tr} D^{(\mu)*}(U) = \chi^{(\mu)*}(U) , \tag{10.47}$$

so

$$\sum_{U \in G} \chi^{(\mu)*}(U) D_{ij}^{(\nu)}(U) = \frac{g}{n_\nu}\delta_{\mu\nu}\delta_{ij} .$$

Finally, setting $i = j$ and summing on j, we find that

$$\sum_{U \in G} \chi^{(\mu)*}(U) \sum_{j=1}^{n_\nu} D_{jj}^{(\nu)}(U) = \frac{g}{n_\nu}\delta_{\mu\nu} \sum_{j=1}^{n_\nu} \delta_{jj} = g\delta_{\mu\nu} .$$

Using Eq. (10.47) once more, we arrive at

$$\sum_{U\in G} \chi^{(\mu)*}(U)\chi^{(\nu)}(U) = g\delta_{\mu\nu} . \tag{10.48}$$

Since all the members of a given class have the same character, we can convert the sum on $U\in G$ in Eq. (10.48) into a sum on the classes. If g_i equals the number of elements in $\mathscr{C}_i$, and if k denotes the number of classes, Eq. (10.48) becomes

$$\sum_{i=1}^{k} g_i\chi_i^{(\mu)*}\chi_i^{(\nu)} = g\delta_{\mu\nu} . \tag{10.49}$$

Equation (10.49) tells us that the N nonequivalent irreducible representations provide us with N k-dimensional orthogonal vectors via the class characters. Since in a k-dimensional space there cannot be more than k orthogonal vectors, we conclude that

$$N \leq k . \tag{10.50}$$

In the case of abelian groups, Eq. (10.50) has the same content as Eq. (10.46). Since abelian groups have only one-dimensional representations, we have

$$\sum_{\nu=1}^{N} n_\nu^2 = \sum_{\nu=1}^{N} 1 = N .$$

So Eq. (10.46) becomes

$$N \leq g .$$

On the other hand, for an abelian group every element forms a separate class. Thus $k = g$, and Eq. (10.50) becomes

$$N \leq g .$$

Hence both Eqs. (10.46) and (10.50) lead to the same inequality for an abelian group.

It is also worth noting that in deriving the orthogonality relations of Eqs. (10.45) and (10.49), we have been writing sums on group elements quite freely. However, there are, as mentioned before, cases involving infinite groups where the sum can be replaced by an integration. Furthermore, for certain groups, integrals over the entire group will be convergent, and the proofs of this section will hold virtually unchanged. As a simple example, we look briefly at the group of rotations about the z-axis. We have already found the one-dimensional representations

$$D^{(m)}(\alpha) = e^{im\alpha} . \tag{10.51}$$

The logical candidate for integration variable for this group is clearly α, with infinitesimal element $d\alpha$. For this case, we expect $\sum_{U\in G}$ to be replaced by $\int_0^{2\pi} d\alpha$. In particular, since

$$g = \sum_{U\in G} 1 ,$$

we expect g to be replaced by

$$\int_0^{2\pi} 1 d\alpha = 2\pi$$

in this case. For groups of this type we refer to g as the volume of the group. According to the above speculations, we would expect Eq. (10.45) to read

$$\int_0^{2\pi} D^{(m)*}(\alpha) D^{(n)}(\alpha) d\alpha = 2\pi\, \delta_{mn}\,, \tag{10.52}$$

since the dimension of all representations is 1. We see that if $D^{(m)}(\alpha)$ is given by Eq. (10.51), then Eq. (10.52) is obeyed. Similar considerations can be applied to other continuous groups, such as the rotation group and the unitary groups. However, there are exceptions to this simple-minded type of extension of the theory of finite groups, the most significant one being the Lorentz group.

10.7 THE DETERMINATION OF GROUP REPRESENTATIONS

In this section we want to apply the major theorems proved thus far to the problem of actually finding group representations. In doing so, we shall also be able to sharpen slightly the orthogonality relations obtained in the previous section so that they become completeness relations.

First, we note that the characters of a representation, D, which is *not* irreducible can be simply written in terms of those of the irreducible representations. If D is written in the form of Eq. (10.19)—as it always can be by use of a similarity transformation—then the character, χ_i, of a given class will have the form

$$\chi_i = a^{(1)}\chi_i^{(1)} + a^{(2)}\chi_i^{(2)} + \cdots + a^{(N)}\chi_i^{(N)} \equiv \sum_{\nu=1}^{N} a^{(\nu)}\chi_i^{(\nu)}\,. \tag{10.53}$$

This follows directly from the definition of the characters in terms of the trace of the representative matrices. $a^{(\nu)}$ is just the number of times the νth irreducible representation occurs in the decomposition of D given by Eq. (10.20). Multiplying Eq. (10.53) on both sides by $g_i\chi_i^{(\mu)*}$, summing on i, and using the orthogonality relation Eq. (10.49), we get

$$\sum_{i=1}^{k} g_i\chi_i^{(\mu)*}\chi_i = \sum_{\nu=1}^{N} a^{(\nu)} \sum_{i=1}^{k} g_i\chi_i^{(\mu)*}\chi_i^{(\nu)} = \sum_{\nu=1}^{N} g a^{(\nu)}\delta_{\mu\nu} = g a^{(\mu)}\,.$$

Hence

$$a^{(\mu)} = \frac{1}{g}\sum_{i=1}^{k} g_i\chi_i^{(\mu)*}\chi_i\,, \tag{10.54}$$

so the number of times a given irreducible representation occurs in an arbitrary representation is readily determined from the characters of that representation. Thus if two representations have the same characters they must be equivalent, since the $a^{(\mu)}$ will then be the same for all of them.

As an example, let us apply Eqs. (10.53) and (10.54) to the regular representation discussed in Section 10.5. This special case will lead to several useful results. The characters of the regular representation are easily determined. Since its dimensionality is g (the number of elements of G), the character χ_1 of the identity class will be equal to the trace of the g-dimensional unit matrix:

$$\chi_1 = g \, . \tag{10.55}$$

All other characters vanish, since all other representative matrices have only zeros on the diagonal, as was pointed out in Section 10.5. Hence

$$\chi_i = 0 \, , \qquad i \neq 1 \, . \tag{10.56}$$

Thus Eq. (10.53) becomes, for $i = 1$,

$$g = \sum_{\nu=1}^{N} a^{(\nu)} \chi_1^{(\nu)} \, .$$

But in any representation of dimension n_ν, the character of the identity class is just n_ν, the trace of the n_ν-dimensional unit matrix. Thus

$$g = \sum_{\nu=1}^{N} a^{(\nu)} n_\nu \, . \tag{10.57}$$

On the other hand, using Eqs. (10.55) and (10.56) together with the fact that $g_1 = 1$, we find from Eq. (10.54) that

$$a^{(\mu)} = \frac{1}{g} g_1 \chi_1^{(\mu)*} \chi_1 = \chi_1^{(\mu)*} = n_\mu \, , \tag{10.58}$$

so Eq. (10.57) becomes

$$g = \sum_{\nu=1}^{N} n_\nu^2 \, . \tag{10.59}$$

Equation (10.58) says that in the regular representation *every* irreducible representation occurs precisely as many times as its dimensionality. Equation (10.59) is a much stronger result than Eq. (10.46). It tells us that the $D_{ij}^{(\nu)}(U)$ provide us with exactly g orthogonal vectors (because $g = \sum n_\nu^2$) in the g-dimensional vector space generated by the g group elements via their irreducible representations. Thus the $D_{ij}^{(\nu)}(U)$ $(U \in G)$ are a *complete* set of vectors for this space. Normalizing in the manner required by Eq. (10.45), we may say that the "representation vectors" whose g components are

$$\sqrt{\frac{n_\nu}{g}}\, D_{ij}^{(\nu)}(U_1), \; \sqrt{\frac{n_\nu}{g}}\, D_{ij}^{(\nu)}(U_2), \; \cdots, \; \sqrt{\frac{n_\nu}{g}}\, D_{ij}^{(\nu)}(U_g) \tag{10.60}$$

for $\nu = 1, 2, \cdots, N$; $i, j = 1, 2, \cdots, n_\nu$ form a complete orthonormal set in the g-dimensional complex vector space. Using a cumbersome, but minimal notation we may express this completeness properly as

$$\sum_{\nu=1}^{N} \sum_{i=1}^{n_\nu} \sum_{j=1}^{n_\nu} \sqrt{\frac{n_\nu}{g}}\, D_{ij}^{(\nu)}(U) \sqrt{\frac{n_\nu}{g}}\, D_{ij}^{(\nu)*}(U') = \delta_{UU'} \, ,$$

where $\delta_{UU'}$ is a Kronecker δ-symbol which is equal to 1 if $U = U'$ and is zero otherwise. This may be written more simply as

$$\sum_{\nu=1}^{N} \sum_{i=1}^{n_\nu} \sum_{j=1}^{n_\nu} \frac{n_\nu}{g} D_{ij}^{(\nu)}(U) D_{ij}^{(\nu)*}(U') = \delta_{UU'} . \tag{10.61}$$

Having strengthened the orthogonality properties of the D's in this manner, we are led naturally to conjecture that a similar strengthening can be achieved for the characters. To see that this is indeed the case, let us sum both sides of Eq. (10.61) over all $U \in \mathscr{C}_l$ and all $U' \in \mathscr{C}_m$. We find that

$$\sum_{\nu=1}^{N} \sum_{i=1}^{n_\nu} \sum_{j=1}^{n_\nu} n_\nu \Big(\sum_{U \in \mathscr{C}_l} D_{ij}^{(\nu)}(U) \Big) \Big(\sum_{U' \in \mathscr{C}_m} D_{ij}^{(\nu)*}(U') \Big) = g \sum_{U \in \mathscr{C}_l} \sum_{U' \in \mathscr{C}_m} \delta_{UU'} . \tag{10.62}$$

The right-hand side of this equation vanishes unless $\mathscr{C}_l = \mathscr{C}_m$, for otherwise $U \neq U'$. If $\mathscr{C}_l = \mathscr{C}_m$, then each term with $U = U'$ will contribute 1 to the sum and all other terms will contribute zero. There are g_l contributing terms if g_l is the number of elements in $\mathscr{C}_l$, so in the case $\mathscr{C}_l = \mathscr{C}_m$, the right-hand side of Eq. (10.62) equals gg_l. But according to the definition of Eq. (10.35),

$$\sum_{U \in \mathscr{C}_l} D^{(\nu)}(U) = M_l^{(\nu)} .$$

Hence Eq. (10.62) becomes

$$\sum_{\nu=1}^{N} \sum_{i=1}^{n_\nu} \sum_{j=1}^{n_\nu} n_\nu [M_l^{(\nu)}]_{ij} [M_m^{(\nu)}]_{ij} = gg_l \delta_{lm} . \tag{10.63}$$

Now by Eq. (10.38),

$$[M_l^{(\nu)}]_{ij} = \frac{g_l}{n_\nu} \chi_l^{(\nu)} \delta_{ij} , \qquad [M_m^{(\nu)}]_{ij} = \frac{g_m}{n_\nu} \chi_m^{(\nu)} \delta_{ij} ,$$

so Eq. (10.63) reduces to

$$\sum_{\nu=1}^{N} \sum_{i=1}^{n_\nu} \sum_{j=1}^{n_\nu} \frac{1}{n_\nu} \chi_l^{(\nu)} \chi_m^{(\nu)*} \delta_{ij} = \frac{g}{g_m} \delta_{lm} ,$$

where we have made use of the fact that $\delta_{ij}^2 = \delta_{ij}$. Carrying out the sum on i and j, we obtain finally

$$\sum_{\nu=1}^{N} \chi_l^{(\nu)} \chi_m^{(\nu)*} = \frac{g}{g_m} \delta_{lm} . \tag{10.64}$$

Thus if we consider the N constants obtained from the characters of the lth class in each irreducible representation as components of a vector in N-dimensional space, then this vector is orthogonal to the vector obtained in a similar way by using the mth class. Therefore, we have k orthogonal vectors in an N-dimensional space, and we conclude that

$$k \leq N .$$

Combining this result with Eq. (10.50), we get

$$k = N. \tag{10.65}$$

Hence the number of irreducible representations is equal to the number of classes in the group! Combining Eqs. (10.49) and (10.64), we may say that the "character vectors" whose k elements are

$$\sqrt{\frac{g_1}{g}}\,\chi_1^{(\nu)},\ \sqrt{\frac{g_2}{g}}\,\chi_2^{(\nu)},\ \cdots,\ \sqrt{\frac{g_k}{g}}\,\chi_k^{(\nu)} \tag{10.66}$$

for $\nu = 1, 2, \cdots, N$ form a *complete* orthonormal set in k-dimensional space.

We now summarize all the basic results in their final form:

a) $$\sum_{\nu=1}^{N} n_\nu^2 = g\,, \tag{10.59}$$

b) $$k = N\,, \tag{10.65}$$

c) $$\sum_{U\in G} \sqrt{\frac{n_\mu}{g}}\,D_{lk}^{(\mu)*}(U)\,\sqrt{\frac{n_\nu}{g}}\,D_{ij}^{(\nu)}(U) = \delta_{il}\delta_{kj}\delta_{\mu\nu}\,, \tag{10.45}$$

d) $$\sum_{\nu=1}^{N}\sum_{i=1}^{n_\nu}\sum_{j=1}^{n_\nu} \sqrt{\frac{n_\nu}{g}}\,D_{ij}^{(\nu)}(U)\,\sqrt{\frac{n_\nu}{g}}\,D_{ij}^{(\nu)*}(U') = \delta_{UU'}\,, \tag{10.61}$$

e) $$\sum_{i=1}^{N} \sqrt{\frac{g_i}{g}}\,\chi_i^{(\mu)*}\,\sqrt{\frac{g_i}{g}}\,\chi_i^{(\nu)} = \delta_{\mu\nu}\,, \tag{10.49}$$

f) $$\sum_{\nu=1}^{N} \sqrt{\frac{g_i}{g}}\,\chi_i^{(\nu)}\,\sqrt{\frac{g_j}{g}}\,\chi_j^{(\nu)*} = \delta_{ij}\,. \tag{10.64}$$

Here g is the order of the group, g_i is the number of elements in the ith class, k is the number of classes, N is the number of irreducible representations and n_ν is the dimensionality of the νth representation.

With this arsenal at our disposal, we can easily determine the characters and representations of some of the basic finite groups. We start with the simplest group, namely, the one-element group consisting of the identity. Since there is only one class, there is only one representation which is one-dimensional. Thus the problem is solved:

$$D^{(1)}(e) = 1\,.$$

Consider now the two element group, $\{e, a\}$. The classes are

$$\mathscr{C}_1 = e\,, \qquad \mathscr{C}_2 = a\,,$$

so there are two representations. Since the group is abelian, these must be one-dimensional. Thus we expect to find

$$\sum_{\nu=1}^{2} n_\nu^2 = 2\,.$$

Since $n_\nu = 1$ for both $\nu = 1$ and $\nu = 2$, this relation is obeyed. For one-dimensional representations, the characters are the same as the representations, so we can concentrate on the characters. It is useful to make a table of the form

$\nu \backslash i$	1	2
1	1	α
2	1	β

(10.67)

where i labels the class and ν labels the representation. The $i\nu$th element of the square array is just $\chi_i^{(\nu)}$. Since the character of the identity class is equal to the dimensionality of the representation, we have filled in the first column in Eq. (10.67) accordingly (in what follows we shall always label the identity class by $i = 1$). Since in the case of one-dimensional representations the characters are equal to the representation matrices, they must obey the multiplication table of the group. Since $a^2 = e$, we have in this case $\alpha^2 = 1$, $\beta^2 = 1$, $\alpha = \pm 1$, $\beta = \pm 1$. Thus the two distinct representations form the character table

$\nu \backslash i$	1	2
1	1	1
2	1	-1

(10.68)

Note that the "character vector" $(1, 1)$ is orthogonal to the "character vector" $(1, -1)$, as it should be according to Eq. (10.49).

The three element group, $\{e, a, a^2\}$ with $a^3 = e$, has three classes

$$\mathscr{C}_1 = e\,, \qquad \mathscr{C}_2 = a\,, \qquad \mathscr{C}_3 = a^2\,,$$

and therefore three one-dimensional representations. Note that Eq. (10.59) is satisfied in this case. Using the multiplication table for the group, which says that the square of element two is equal to element three, we can write immediately for the character table

$\nu \backslash i$	1	2	3
1	1	α	α^2
2	1	β	β^2
3	1	γ	γ^2

and $\alpha^3 = \beta^3 = \gamma^3 = 1$. Thus α, β, and γ can take on three possible values, 1, $e^{2\pi i/3}$ and $e^{4\pi i/3}$. The three distinct representations are therefore

$\nu \backslash i$	1	2	3
1	1	1	1
2	1	ϵ	ϵ^2
3	1	ϵ^2	ϵ

(10.69)

with $\epsilon = {}^{2\pi i/3}$. The three resulting "character vectors," $(1, 1, 1)$, $(1, \epsilon, \epsilon^2)$ and $(1, \epsilon^2, \epsilon)$ are mutually orthogonal, as the reader may readily verify.

At the fourth order, we find two groups awaiting our attention, the cyclic group of order four and the four-group. For the cyclic group, we shall just remark that by our treatment of the previous three cyclic groups, it should be clear to the reader that the characters of the cyclic group of order n are generated in an obvious way by the n nth roots of unity. Thus the character table for the fourth-order cyclic group looks like

$\nu \backslash i$	1	2	3	4
1	1	1	1	1
2	1	δ	δ^2	δ^3
3	1	δ^2	1	δ^2
4	1	δ^3	δ^2	δ

where $\delta = e^{2\pi i/4} = i$ and we have used the fact that $\delta^4 = 1$ to simplify the array. The generalization to the cyclic group of order n is obvious.

For the four-group, if we denote the elements as usual by $\{e, a, b, c\}$, then the discussion in Section 10.3 makes it clear that $\{e, a\}$ is an invariant subgroup. The factor group is

$$E = \{e, a\}\,, \qquad A = \{b, c\}\,.$$

Thus representations of $\{E, A\}$ [see Eq. (10.68)] can be transferred to the four-group by assigning the representative of E to both e and a and the representative of A to both b and c. The character assignment $\chi(E) = 1, \chi(A) = 1$ gives $\chi(e) = 1$, $\chi(a) = 1$, $\chi(b) = 1$ and $\chi(c) = 1$. The assignment $\chi(E) = 1, \chi(A) = -1$ gives $\chi(e) = 1$, $\chi(a) = 1$, $\chi(b) = -1$ and $\chi(c) = -1$. Similarly, the invariant subgroup $\{e, b\}$ gives rise to two assignments for the four-group. One of these is the trivial representation (all characters equal unity) which we have already obtained from $\{e, a\}$; the other is $\chi(e) = 1$, $\chi(a) = -1$, $\chi(b) = 1$ and $\chi(c) =$

-1. Finally, the invariant subgroup $\{e, c\}$ gives rise to one new set of characters, $\chi(e) = 1$, $\chi(a) = -1$, $\chi(b) = -1$, and $\chi(c) = 1$. Thus the character table of the four-group looks like

$\nu \backslash i$	1	2	3	4
1	1	1	1	1
2	1	1	-1	-1
3	1	-1	1	-1
4	1	-1	-1	1

(10.70)

Once again, the orthogonality relations are all satisfied.

Since five is a prime number, the only group of order five is the cyclic group which is dealt with by extension of our discussion of the cyclic group of order four. The same remark applies to the cyclic group of order six, which brings us to S_3, the six-element group of permutations on three objects. This is the smallest nonabelian group. We know that it has three classes,

$$\mathscr{C}_1 = e\,, \qquad \mathscr{C}_2 = \{(12), (13), (23)\}\,, \qquad \mathscr{C}_3 = \{(123), (321)\}\,,$$

with corresponding characters χ_1, χ_2, and χ_3. The number of elements in each class is $g_1 = 1$, $g_2 = 3$, and $g_3 = 2$. Since there are three classes, there must be three irreducible representations. Hence by Eq. (10.59)

$$\sum_{\nu=1}^{3} n_\nu^2 = 6\,.$$

The only way to add three squares to obtain the value six is if $n_1 = 1$, $n_2 = 1$, and $n_3 = 2$. Thus for the first time we will have a two-dimensional representation. Using the facts above, we can start our character table as follows:

$\nu \backslash i$	1	2	3
1	1	1	1
2	1	a	b
3	2	c	d

We have put in the trivial representation directly and have used the fact that $\chi_1^{(\nu)} = n_\nu$. Now by Eq. (10.49)

$$\sum_{i=1}^{3} g_i \chi_i^{(1)*} \chi_i^{(2)} = 0\,,$$

that is

$$1 + 3a + 2b = 0\,. \tag{10.71}$$

Similarly, from

$$\sum_{i=1}^{3} g_i \chi_i^{(1)*} \chi_i^{(3)} = 0\,,$$

we conclude that

$$3c + 2d + 2 = 0\,. \tag{10.72}$$

Using Eqs. (10.71) and (10.72), we simplify the character table to

$\nu \backslash i$	1	2	3
1	1	1	1
2	1	a	$-\dfrac{3a+1}{2}$
3	2	c	$-\dfrac{3c+2}{2}$

We can now use other orthogonality properties of the characters [Eq. (10.61)] to write

$$\sum_{\nu=1}^{3} \chi_1^{(\nu)*} \chi_2^{(\nu)} = 0\,,$$

$$\sum_{\nu=1}^{3} \chi_1^{(\nu)*} \chi_3^{(\nu)} = 0\,.$$

These both lead directly to

$$a + 2c + 1 = 0\,,$$

so the character table now contains only one unknown:

$\nu \backslash i$	1	2	3
1	1	1	1
2	1	$-1-2c$	$1+3c$
3	2	c	$-1-\frac{3}{2}c$

A simple way of tying down the value of c is to note that $(123)^3 = e$ and $(123)^2 = (321)$. Therefore in the one-dimensional representation $\chi_3^3 = 1$ and also $\chi_3^2 = \chi_3$, since (123) and (321) both belong to the same class. Thus $\chi_3 = 1$, and this implies that $1 + 3c = 1$, or $c = 0$. The character table now takes its final form:

$\nu \backslash i$	1	2	3
1	1	1	1
2	1	-1	1
3	2	0	-1

The second irreducible representation is the so-called antisymmetric representation, obtained by representing odd permutations by -1 and even ones by $+1$. It clearly exists for all S_n. The reader should check that the six normalization conditions implied by Eqs. (10.49) and (10.61),

$$\sum_{i=1}^{3} g_i \chi_i^{(\nu)*} \chi_i^{(\nu)} = 6 \,, \qquad \nu = 1, 2, 3 \,,$$

and

$$\sum_{\nu=1}^{3} \chi_i^{(\nu)*} \chi_i^{(\nu)} = 6/g_i \,, \qquad i = 1, 2, 3 \,,$$

are also satisfied. It should be emphasized that this is by no means the fastest way of obtaining the character table for S_3. For example, the fact that $\mathscr{C}_1$ and $\mathscr{C}_3$ combine to form an invariant subgroup whose factor group is isomorphic to $\{e, a\}$ enables one to write the two one-dimensional representations immediately from Eq. (10.68). Then the character vector of the two-dimensional representation is uniquely determined by orthogonality requirements, as the reader can show.

The characters of the two one-dimensional representations are, of course, exactly equal to the corresponding one-dimensional representation matrices:

$$D^{(1)}(U) = 1 \,,$$

for all $U \in S_3$, and

$$D^{(2)}(e) = D^{(2)}(123) = D^{(2)}(321) = 1 \,,$$
$$D^{(2)}(12) = D^{(2)}(13) = D^{(2)}(23) = -1 \,.$$

However, the two-dimensional representation requires some additional work to obtain the representative matrices. The identity element is no problem. We have

$$D^{(3)}(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{10.74}$$

As for the matrix $D^{(3)}(12)$, let us choose it to be diagonal, since one can always pick one element of a representation to be diagonal in addition to the unit matrix (why?). Thus

$$D^{(3)}(12) = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

But $(12) \in \mathscr{C}_2$, so by the character table, $\chi_2^{(3)} = 0$, that is, the trace of $D^{(3)}(12)$ is zero. Hence $b = -a$. Also, since $(12)^2 = e$, we must have $a^2 = 1$. Thus choosing $a = 1$,

$$D^{(3)}(12) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{10.75}$$

Note that if we had taken $a = -1$, we would have found a different matrix, but one which is related to the matrix of Eq. (10.75) by a similarity transformation. Now what about $D^{(3)}(13)$ and $D^{(3)}(23)$? Clearly, since $(13)^2 = e$,

$$D^{(3)}(13) = D^{(3)}(13)^{-1}. \tag{10.76}$$

But since we need consider only unitary representations,

$$D^{(3)}(13)^\dagger = D^{(3)}(13)^{-1},$$

and combining this with Eq. (10.76), we get

$$D^{(3)}(13) = D^{(3)}(13)^\dagger,$$

that is, $D^{(3)}(13)$ is Hermitian. Similarly, $D^{(3)}(23)$ is Hermitian. Thus, using the fact that the trace of both $D^{(3)}(13)$ and $D^{(3)}(23)$ is zero, we may write in complete generality

$$D^{(3)}(13) = \begin{pmatrix} \alpha & \beta \\ \beta^* & -\alpha \end{pmatrix}, \qquad D^{(3)}(23) = \begin{pmatrix} \gamma & \delta \\ \delta^* & -\gamma \end{pmatrix} \tag{10.77}$$

where α and γ are real. But by Eqs. (10.35) and (10.38),

$$M_2^{(3)} = D^{(3)}(12) + D^{(3)}(13) + D^{(3)}(23) = \frac{g_2}{n_3}\chi_2^{(3)} I = 0,$$

since $\chi_2^{(3)} = 0$. Thus using Eqs. (10.75) and (10.77), we have

$$\begin{pmatrix} 1 + \alpha + \gamma & \beta + \delta \\ \beta^* + \delta^* & -1 - \alpha - \gamma \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

so that the matrices $D^{(3)}(13)$ and $D^{(3)}(23)$ take the form

$$D^{(3)}(13) = \begin{pmatrix} \alpha & \beta \\ \beta^* & -\alpha \end{pmatrix}, \qquad D^{(3)}(23) = \begin{pmatrix} -(\alpha + 1) & -\beta \\ -\beta^* & (\alpha + 1) \end{pmatrix}.$$

Finally, we use the unitarity conditions:

$$D^{(3)}(13)^\dagger D^{(3)}(13) = \begin{pmatrix} \alpha^2 + |\beta|^2 & 0 \\ 0 & \alpha^2 + |\beta|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$D^{(3)}(23)^\dagger D^{(3)}(23) = \begin{pmatrix} (\alpha + 1)^2 + |\beta|^2 & 0 \\ 0 & (\alpha + 1)^2 + |\beta|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

These yield the two equations

$$\alpha^2 + |\beta|^2 = 1\,, \qquad (\alpha + 1)^2 + |\beta|^2 = 1\,,$$

which are readily solved to give $\alpha = -\frac{1}{2}$, $\beta = (\sqrt{3}/2)e^{i\phi}$, where ϕ is arbitrary. Thus

$$D^{(3)}(13) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2}e^{i\phi} \\ \frac{\sqrt{3}}{2}e^{-i\phi} & \frac{1}{2} \end{pmatrix}, \tag{10.78}$$

$$D^{(3)}(23) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2}e^{i\phi} \\ -\frac{\sqrt{3}}{2}e^{-i\phi} & \frac{1}{2} \end{pmatrix}. \tag{10.79}$$

Now $(123) = (13)(12)$, so

$$\begin{aligned} D^{(3)}(123) = D^{(3)}(13)D^{(3)}(12) &= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2}e^{i\phi} \\ \frac{\sqrt{3}}{2}e^{-i\phi} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2}e^{i\phi} \\ \frac{\sqrt{3}}{2}e^{-i\phi} & -\frac{1}{2} \end{pmatrix}. \end{aligned} \tag{10.80}$$

Also, $(321) = (123)^{-1}$, so

$$D^{(3)}(321) = D^{(3)}(123)^{-1} = D^{(3)}(123)^{\dagger} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2}e^{i\phi} \\ -\frac{\sqrt{3}}{2}e^{-i\phi} & -\frac{1}{2} \end{pmatrix}, \tag{10.81}$$

which completes the determination of the representative matrices except for ϕ. However, it is always possible, without loss of generality, to choose $\phi = 0$, since a simple multiplication shows that any matrix of the form

$$\begin{pmatrix} a & be^{i\phi} \\ ce^{-i\phi} & d \end{pmatrix}$$

is reduced to the form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

when similarity transformed by the unitary matrix

$$\begin{pmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{pmatrix}.$$

This fact, along with Eqs. (10.74), (10.75), (10.78), (10.79), (10.80), and (10.81), enables us to write finally

$$\begin{gathered} D^{(3)}(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D^{(3)}(12) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad D^{(3)}(13) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}, \\ D^{(3)}(23) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}, \quad D^{(3)}(123) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \\ D^{(3)}(321) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \end{gathered} \tag{10.82}$$

We have already written these matrices in Eq. (3.5), where we pointed out that they were isomorphic to the group of permutations of three objects. Here we have started with S_3 and relentlessly derived these isomorphic matrices.

It is interesting to note that in this case, when the representation matrices are not all equal to characters, Eq. (10.45) is not equivalent to Eq. (10.49) as it is for purely one-dimensional cases. For example, the $i = 1$, $j = 1$ element of all the matrices of the two-dimensional representation [see Eq. (10.82)] gives a "representation vector" which is, according to Eq. (10.60),

$$\frac{1}{\sqrt{3}}\left(1,\ 1,\ -\frac{1}{2},\ -\frac{1}{2},\ -\frac{1}{2},\ -\frac{1}{2}\right).$$

On the other hand, the $i = 1$, $j = 2$ element gives

$$\frac{1}{\sqrt{3}}\left(0,\ 0,\ \frac{\sqrt{3}}{2},\ -\frac{\sqrt{3}}{2},\ -\frac{\sqrt{3}}{2},\ \frac{\sqrt{3}}{2}\right).$$

These two vectors are normalized to unity and are mutually orthogonal, as they should be according to Eq. (10.45). The nontrivial one-dimensional representation gives the "representation vector"

$$\frac{1}{\sqrt{6}}\,(1,\ -1,\ -1,\ -1,\ 1,\ 1)\ ,$$

which is normalized correctly to unity and is orthogonal to each of the two "representation vectors" written above. The reader may want to check the remaining orthonormality relations implied by Eqs. (10.45) and (10.61).

10.8 GROUP THEORY IN PHYSICAL PROBLEMS

Now that we know how to obtain the characters and representations of groups, we can indicate some of the ways in which group theory can be used to increase our understanding of certain basic physical processes.

As a particularly simple illustration of what we have in mind, consider the problem of a quantum-mechanical system which has various eigenstates which are degenerate because of the existence of some symmetry of Schrödinger's equation. As we mentioned in Section 10.4, the collection of eigenfunctions belonging to some degenerate eigenvalue will transform into itself under any transformation belonging to the symmetry group in question. According to Eq. (10.15), if $\mathscr{U} \in \mathscr{G}$, and $U \in G$, then

$$\mathscr{U}\phi_m(\mathbf{r}) = \sum_{m'} D_{m'm}(U)\phi_{m'}(\mathbf{r})\ ,$$

where $D(U)$ $(U \in G)$ is a representation of the group G. Now suppose that we apply to the system a weak perturbation having a smaller symmetry group than the original system. For example, we might have an atom, which in its natural environment in free space is a rotationally invariant system, but which when put in, say, the lattice of a cubic symmetric crystal feels a potential which has the symmetry of the cube. The symmetry of the cube is obviously much smaller than that of complete rotational invariance; in fact, the infinite-element symmetry group of all rotations in three dimensions is replaced by the 24-element symmetry group of the cube (this group is isomorphic to the group, S_4, of permutations on four objects).

When the atom finds itself in this new cubic potential, the energy levels will be modified in such a way that the degenerate eigenvectors belonging to the new energy levels will transform among themselves according to the representations of the group, $\mathcal{O}$, of symmetries of the cube. In other words, a μ-fold degenerate level of the rotationally invariant system must break up into M sets of degenerate levels belonging to $\mathcal{O}$, with multiplicities $\mu_1, \mu_2, \cdots, \mu_M$. Clearly, $\mu_1 + \mu_2 + \cdots + \mu_M = \mu$.

The proper framework for a discussion of these questions is provided by degenerate perturbation theory (Section 4.12). If the reader will look back at that section, he will see that the important objects in the discussion of perturbation theory are the so-called matrix elements of the perturbing operator between states of the unperturbed system. To be precise, let us suppose as in Section 4.12, that A_0 is the linear operator governing the unperturbed system:

$$A_0\phi_{n,\nu,i} = \lambda_{n,\nu}\,\phi_{n,\nu,i}\ .$$

Here we have chosen the indices to conform to the group theoretical facts which we have learned in this chapter. ν labels the irreducible representation which tells us how the eigenfunction transforms under the symmetry group, G_0, of A_0, and i is the column of the representation to which $\phi_{n,\nu,i}$ belongs. That is, in standard terminology,

$$\mathscr{V}\phi_{n,\nu,i} = \sum_{j=1}^{n_\nu} D_{ji}^{(\nu)}(V)\phi_{n,\nu,j}\ ; \tag{10.83a}$$

V is any element of G_0 ($\mathscr{V}$ is the corresponding element of $\mathscr{G}_0$) and n denotes any additional labels which may be necessary to specify uniquely the eigenvalue under consideration. For example, the νth representation may be of relevance more than once as we run through the entire spectrum of A_0.

We now imagine that the perturbation A_1 is invariant under each element, U, of the symmetry group G_1 and that $\psi_{n,\nu,i}$ is a set of functions which transform according to the νth representation of G_1:

$$\mathscr{U}\psi_{n,\nu,i} = \sum_{j=1}^{n_\nu} D_{ji}^{(\nu)}(U)\psi_{n,\nu,j} \tag{10.83b}$$

where $\mathscr{U} \in \mathscr{G}_1$. We wish to consider matrix elements of the form

$$M_{n,\nu,i;n',\mu,j} \equiv (\psi_{n,\nu,i},\ A_1\psi_{n',\mu,j})\ . \tag{10.84}$$

Since $\mathscr{U}^\dagger\mathscr{U} = \mathscr{U}\mathscr{U}^\dagger = I$, we may write Eq. (10.84) as

$$\begin{aligned} M_{n,\nu,i;n',\mu,j} &= (\psi_{n,\nu,i},\ \mathscr{U}^\dagger\mathscr{U}A_1\mathscr{U}^\dagger\mathscr{U}\psi_{n',\mu,j}) \\ &= (\mathscr{U}\psi_{n,\nu,i},\ \mathscr{U}A_1\mathscr{U}^{-1}\mathscr{U}\psi_{n,\mu,j})\ . \end{aligned}$$

Using the transformation properties of the $\psi_{n,\nu,i}$ we may write this as

$$M_{n,\nu,i;n',\mu,j} = \left(\sum_{k=1}^{n_\nu} D_{ki}^{(\nu)}(U)\psi_{n,\nu,k}, [\mathscr{U} A_1 \mathscr{U}^{-1}] \sum_{l=1}^{n_\mu} D_{lj}^{(\mu)}(U)\psi_{n',\mu,l} \right),$$

or

$$M_{n,\nu,i,n',\mu,j} = \sum_{k=1}^{n_\nu} \sum_{l=1}^{n_\mu} D_{ki}^{(\nu)*}(U) D_{lj}^{(\mu)}(U)\,(\psi_{n,\nu,k}, [\mathscr{U} A_1 \mathscr{U}^{-1}]\psi_{n',\mu,l}) . \quad (10.85)$$

But since we have assumed A_1 to be invariant under G_1,

$$\mathscr{U} A_1 \mathscr{U}^{-1} = A_1 ,$$

so Eq. (10.85) becomes

$$M_{n,\nu,i;n',\mu,j} = \sum_{k=1}^{n_\nu} \sum_{l=1}^{n_\mu} D_{ki}^{(\nu)*}(U) D_{lj}^{(\mu)}(U)\,(\psi_{n,\nu,k}, A_1\psi_{n',\mu,l}) . \quad (10.86)$$

Now the crucial step is to sum both sides of Eq. (10.86) over all $U \in G_1$, so we will be able to take advantage of the orthogonality relation of Eq. (10.45) on the right-hand side. On the left-hand side, since M does *not* depend on U [see Eq. (10.84)], we get just gM, where g is the order of G_1. Thus

$$gM_{n,\nu,i;n',\mu,j} = \sum_{k=1}^{n_\nu} \sum_{l=1}^{n_\mu} \sum_{U\in G_1} D_{ki}^{(\nu)*}(U) D_{lj}^{(\mu)}(U) M_{n,\nu,k;n',\mu,l} . \quad (10.87)$$

But the main part of this sum involving the representation elements can be evaluated immediately by using the orthogonality theorem [Eq. (10.45)] which tells us directly that

$$\sum_{U\in G_1} D_{ki}^{(\nu)*}(U) D_{lj}^{(\mu)}(U) = \frac{g}{n_\nu} \delta_{kl}\delta_{ij}\delta_{\mu\nu} .$$

Inserting this into Eq. (10.87), we obtain

$$M_{n,\nu,i;n',\mu,j} = \frac{1}{n_\nu} \delta_{\mu\nu}\delta_{ij} \sum_{l=1}^{n_\nu} M_{n,\nu,l;n',\nu,l} . \quad (10.88)$$

Thus we see that matrix elements of this type vanish unless the eigenfunctions belong to the same representation and, furthermore, to the same column of this representation. This result can often be helpful in perturbation calculations by severely limiting the number of eigenfunctions which can occur in the perturbation-theory sums [see, e.g., Eq. (4.60)].

On the most elementary level, Eq. (10.88) tells us that a perturbation, A_1, with the *same* symmetry group as the dominant operator, A_0, cannot split degeneracies in first order, since, in computing the matrix of A_1 between the degenerate states belonging to some eigenvalue of A_0, we will find only diagonal elements [$i = j$ in Eq. (10.88)], and all the diagonal elements are equal. Hence by Eq. (4.65b), all the first-order energies will be equal, and there is no splitting of the degeneracy. This is not surprising since if A_1 has the same symmetry as A_0, then the operator $A = A_0 + \epsilon A_1$ will also have the same symmetry as A_0.

Therefore, according to the general arguments of Section 10.4, the degenerate levels of A_0 belonging to some irreducible representation of G_0 can never split since these new levels of A would have a smaller multiplicity than the original level. This in turn would mean that there exist several linear combinations of eigenfunctions of the original degenerate level which transform among themselves under G_0. The original level must therefore give rise to a reducible representation of G_0, contrary to assumption. Thus no splitting can occur.

A more useful application of Eq. (10.88) is found by considering the case where A_1 has a symmetry lower than that of A_0 and where the group G_1, under which A_1 is invariant is a subgroup of the larger group G_0, under which A_0 is invariant. The key fact is that in such a case the representations $D^{(\nu)}(V)$ ($V \in G_0$) of G_0 will give rise to representations of G_1 by associating $D^{(\nu)}(V)$ with the $U \in G_1$ which corresponds to $V \in G_0$. In general, this representation will be reducible. For example, the group $\{e, (12)\}$ is a subgroup of $S_3 = \{e, (12), (13), (23), (123), (321)\}$. Now, in the two-dimensional representation of S_3, e is represented by the unit matrix while (12) is represented by

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Thus for the group $\{e, (12)\}$ the representation

$$e \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad (12) \to \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

is perfectly acceptable, but it is obviously reducible into the sum of two one-dimensional representations. In a more complicated case, where the decomposition is not so obvious, the number of times an *irreducible* representation of G_1 is contained in a representation of G_1 *induced* by G_0 can be calculated simply by knowing the characters of the irreducible representations of G_1 and using Eq. (10.54).

With this point of view clearly in mind, it is easy to see what happens when we split a degeneracy by applying a perturbation A_1, which is invariant under G_1, to a system A_0, which is invariant under G_0. A given degenerate level of A_0 belongs to an *irreducible* representation $D^{(\nu)}$ of G_0, i.e., the eigenfunctions of this level transform among themselves *under the action of* G_0 according to $D^{(\nu)}$. However, we can also consider this level as belonging to a *reducible* representation of G_1; that is, certain sets of linear combinations of wave-functions belonging to this level transform into themselves *under the action of* G_1.

We can see this as follows. Suppose $D(V)$ is some *irreducible* representation of G_0. If we restrict ourselves to the subset of V's which are also elements of G_1 and denote a typical element of this subset by U, then $D(U)$ is in general a *reducible* representation of G_1. Therefore there exists some matrix S (independent of U) such that for all $U \in G_1$,

$$S^{-1}D(U)S = D'(U) \qquad \text{or} \qquad D(U) = SD'(U)S^{-1}.$$

$D'(U)$ has the same form as the matrix shown in Eq. (10.19), that is, it has square matrices (*irreducible* representations of G_1) on its diagonal, with all other elements equal to zero. Suppose that $\{f_i\}$ is a set of functions which transform among themselves under G_0 according to the irreducible representation $D(V)$. Then if we restrict our attention to those elements, U, which also belong to G_1, we may write

$$\begin{aligned}\mathscr{U}f_i &= \sum_j D_{ji}(U)\, f_j \\ &= \sum_j [SD'(U)S^{-1}]_{ji}\, f_j \\ &= \sum_{j,k,m} S_{jk}\, D'_{km}(U)[S^{-1}]_{mi}\, f_j .\end{aligned}$$

This last relation can be written as

$$\mathscr{U}\sum_i S_{il}\, f_i = \sum_{j,k} S_{jk}\, D'_{kl}(U) f_j \,.$$

If we write

$$g_l = \sum_i S_{il}\, f_i \qquad \text{then} \qquad \mathscr{U}g_l = \sum_k D'_{kl}(U)\, g_k \,.$$

Looking at the form of $D'(U)$ [see Eq. (10.19)] we can pick out by inspection the sets of g's which transform among themselves according to the irreducible representations of G_1 on the diagonal of $D'(U)$.

If we now apply degenerate perturbation theory to the perturbing operator A_1, using the linear combinations of degenerate eigenfunctions appropriate to G_1, then according to Eq. (10.88) there will be no nonzero matrix elements of A_1 which connect different irreducible representations of G_1, and within an irreducible representation of G, the diagonal elements will give the only nonvanishing contributions. These diagonal elements are all equal within a given irreducible representation. Thus the matrix of A_1, calculated between appropriately chosen linear combinations of eigenfunctions belonging to a degenerate level of A_0, might appear as follows:

$$\begin{pmatrix}
a \\
& a \\
& & a \\
& & & a & & & & & \mathrm{O} \\
& & & & b \\
& & & & & b \\
& & & & & & b \\
& & & & & & & c \\
& & & & & & & & c \\
& & \mathrm{O} & & & & & & & c \\
& & & & & & & & & & c \\
& & & & & & & & & & & c \\
& & & & & & & & & & & & c
\end{pmatrix}$$

In this case, a thirteenfold degenerate level of A_0, belonging to a thirteen-dimensional representation of G_0, has been split into three levels—fourfold, threefold, and sixfold degenerate—belonging to three different irreducible representations of G_1. (The reader should note that we are assuming that a given irreducible representation of G_1 occurs only once in the decomposition. What would the matrix of A_1 look like if some irreducible representation occurred twice?) Of course, one would not often know in advance the correct linear combination of wave-functions, so the above diagonal array would in practice look like a complicated 13×13 matrix, which would take the above form after being similarity transformed (diagonalized). The similarity transform would also give the correct linear combinations of the original wave-functions in the manner explained in Section 4.12.

To illustrate the ideas contained in the above discussion, we consider the example mentioned at the beginning of this section, namely, a rotationally invariant system (an atom in free space) placed in a weak, cubic-symmetric potential (in a crystal with cubic symmetry). Before we can solve the problem, we must first obtain the characters of the irreducible representations in question. Let us take the rotation group first. This is a continuous group, and therefore strictly speaking lies outside the framework developed in the preceding sections. However, if we are interested only in characters, we may proceed as follows. We know from elementary considerations (see Section 5.8) that the $(2l + 1)$ spherical harmonics $Y_{lm}(\theta, \phi)$ $(m = -l, -l + 1, \cdots, l - 1, l)$ are eigenfunctions of the rotationally invariant operator L^2 and all belong to the single eigenvalue $l(l + 1)$ $(l = 0, 1, 2, \cdots)$. These degenerate functions must therefore transform among themselves according to a representation of the rotation group (which can be shown to be irreducible).

We know that a general rotation can be brought by a unitary transformation to the form of a simple rotation about an axis. Let us call the angle of rotation about this axis Φ. Since all rotations through an angle Φ are unitarily equivalent, such rotations form a class which can be labeled by Φ. To obtain the character of such a class, we need consider only a simple rotation through Φ about the z-axis. Such a rotation takes $Y_{lm}(\theta, \phi)$ into $e^{im\Phi} Y_{lm}(\theta, \phi)$ (see Section 5.8). Thus the matrix representing this rotation will be

$$\begin{pmatrix} e^{-il\Phi} & & & & & O \\ & e^{-i(l-1)\Phi} & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & e^{i(l-1)\Phi} & \\ O & & & & & e^{il\Phi} \end{pmatrix}.$$

The trace is readily computed; we find that

$$\chi^{(l)}(\Phi) = \sum_{m=-l}^{l} e^{im\Phi} = \frac{\sin (l + \frac{1}{2})\Phi}{\sin (\frac{\Phi}{2})}, \tag{10.89}$$

where we have evaluated the sum by using the formula for summing a geometric series. Note that the dimension, n_l, of the lth representation is

$$\lim_{\Phi\to 0} \frac{\sin (l + \frac{1}{2})\Phi}{\sin (\frac{\Phi}{2})} = 2l + 1 ,$$

as we should expect since $\Phi = 0$ corresponds to the identity transformation. When $l = 0$, we find the trivial one-dimensional transformation in which every rotation is represented by 1. When $l = 1$, we obtain the three-dimensional representation which should be familiar from classical mechanics; it is given, as a function of the Euler angles, by Eq. (1.25).

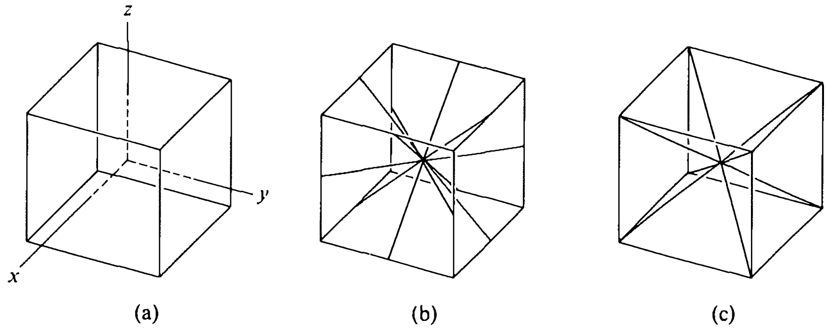

Fig. 10.2 The axes of symmetry of a cube.

Having found the class characters for all the representations of the group of three-dimensional rotations, let us now move to more familiar ground and find the representations and characters of the group $\mathcal{O}$ of symmetries of the cube. This group contains 24 elements in five classes, being isomorphic to S_4. The class $\mathscr{C}_1$ is the identity class. Figure 10.2(a) shows the axes of rotation which are relevant for the classes $\mathscr{C}_2$ and $\mathscr{C}_4$. $\mathscr{C}_2$ consists of the six rotations through $\pm\pi/2$ radians about each of the x, y, and z axes; the class $\mathscr{C}_4$ consists of the three elements corresponding to rotation through π radians about the x-, y-, and z-axes. Figure 10.2(b) shows the axes relevant to the class $\mathscr{C}_5$, which consists of rotations through π radians about each of these six axes. Finally, the class $\mathscr{C}_3$ consists of rotations through $\pm 2\pi/3$ radians about each of the four axes shown in Fig. 10.2(c). We have labeled these classes in this particular manner so they will correspond to the class assignments for S_4 given in Section 10.3. Just as with S_4, the classes $\mathscr{C}_1$ and $\mathscr{C}_4$ combine together to give an invariant subgroup V, the factor group of which is isomorphic to S_3. Thus, in the manner illustrated in the previous section, we can lift representations of S_3 onto S_4 (and hence onto $\mathcal{O}$). Since $\mathcal{O}$ has 24 elements and five classes, we must have

$$n_1^2 + n_2^2 + n_3^2 + n_4^2 + n_5^2 = 24 ,$$

where n_ν is the dimension of the νth representation. The only possible collection of n_ν which can satisfy this relation is $n_1 = 1$, $n_2 = 1$, $n_3 = 2$, $n_4 = 3$, and $n_5 = 3$.

According to our previous remarks, the first, second, and third representations (with $n_1 = 1$, $n_2 = 1$, and $n_3 = 2$) will correspond to those of $S_3 = S_4/V$. Thus, in usual manner, we can fill in a large part of the character table for $\mathcal{O}$:

$\nu \backslash i$	1	2	3	4	5
1	1	1	1	1	1
2	1	-1	1	1	-1
3	2	0	-1	2	0
4	3	a	b	c	d
5	3	α	β	γ	δ

(10.90)

Since $g_1 = 1$, $g_2 = 6$, $g_3 = 8$, $g_4 = 3$, and $g_5 = 6$, we can use Eq. (10.49) to obtain

$$6a + 8b + 3c + 6d = -3\,,$$
$$-6a + 8b + 3c - 6d = -3\,,$$
$$-8b + 6c = -6\,.$$

The equations for α, β, γ, and δ are identical to the above three. These equations are readily solved to give $b = 0$, $c = -1$, $d = -a$; similarly $\beta = 0$, $\gamma = -1$, $\delta = -\alpha$. Thus Table (10.90) becomes

$\nu \backslash i$	1	2	3	4	5
1	1	1	1	1	1
2	1	-1	1	1	-1
3	2	0	-1	2	0
4	3	a	0	-1	$-a$
5	3	α	0	-1	$-\alpha$

The normalization condition implied by Eq. (10.49) gives $a^2 = 1 = \alpha^2$, that is, $a = \pm 1$ and $\alpha = \pm 1$. The two distinct choices are $a = 1$, $\alpha = -1$ or $a =$

-1, $\alpha = 1$. These lead to the same character table except for an interchange of the last two rows. Thus the character table takes its final form:

$\nu \backslash i$	1	2	3	4	5
1	1	1	1	1	1
2	1	-1	1	1	-1
3	2	0	-1	2	0
4	3	1	0	-1	-1
5	3	-1	0	-1	1

(10.91)

In this case, the representation matrices can also be readily obtained. For the first three representations, the matrices are just those previously obtained for S_3. The three-dimensional representations are obtained in a simple manner from the three-dimensional representation of the rotation group. Since we will not need these matrices here, we will omit a detailed analysis of them.

With all the routine work out of the way, we can see what happens when a rotationally invariant atomic system is put into a weak cubic-symmetric environment. First, we note that the various representations of the rotation group induce representations of $\mathcal{O}$. Consider, for example, the $l = 1$ representation of the rotation group. The classes of the cubic group correspond to the following angles of rotations (in radians):

$$\mathscr{C}_1 \to 0\,, \qquad \mathscr{C}_2 \to \pm\pi/2\,, \qquad \mathscr{C}_3 \to \pm 2\pi/3\,, \qquad \mathscr{C}_4 \to \pi\,, \qquad \mathscr{C}_5 \to \pi\,.$$

The characters corresponding to these rotations are obtained from Eq. (10.89). We find that

$$\chi_1 = 3\,, \qquad \chi_2 = 1\,, \qquad \chi_3 = 0\,, \qquad \chi_4 = -1\,, \qquad \chi_5 = -1\,.$$

By looking at the character table of $\mathcal{O}$ [Table (10.91)], we see that this has given us an *irreducible* representation of $\mathcal{O}$, corresponding to the fourth row of the character table. Thus an eigenvalue of the rotationally invariant system belonging to the $l = 1$ representation of the rotation group will not be split when subjected to a perturbation of cubic symmetry.

Moving to the $l = 2$ representation and using Eq. (10.89), we find that

$$\chi_1 = 5\,, \qquad \chi_2 = -1\,, \qquad \chi_3 = -1\,, \qquad \chi_4 = 1\,, \qquad \chi_5 = 1$$

for the characters of the representation of $\mathcal{O}$ induced by the rotation group. Now, to determine the number of times a given *irreducible* representation of $\mathcal{O}$

is contained in this reducible representation, we use Eq. (10.54). We find that

$$a^{(1)} = \frac{1}{24}\sum_{i=1}^{5} g_i\chi_i^{(1)}\chi_i = \frac{1}{24}[5 - 6 - 8 + 3 + 6] = 0\,,$$

$$a^{(2)} = \frac{1}{24}\sum_{i=1}^{5} g_i\chi_i^{(2)}\chi_i = \frac{1}{24}[5 + 6 - 8 + 3 - 6] = 0\,,$$

$$a^{(3)} = \frac{1}{24}\sum_{i=1}^{5} g_i\chi_i^{(3)}\chi_i = \frac{1}{24}[10 + 0 + 8 + 6 + 0] = 1\,,$$

$$a^{(4)} = \frac{1}{24}\sum_{i=1}^{5} g_i\chi_i^{(4)}\chi_i = \frac{1}{24}[15 - 6 + 0 - 3 - 6] = 0\,,$$

$$a^{(5)} = \frac{1}{24}\sum_{i=1}^{5} g_i\chi_i^{(5)}\chi_i = \frac{1}{24}[15 + 6 + 0 - 3 + 6] = 1\,.$$

Thus the third and fifth *irreducible* representations of $\mathcal{O}$ occur in the *reducible* representation induced by the $l = 2$ representation of the rotation group. Therefore the fivefold degenerate state corresponding to an $l = 2$ representation of the rotation group will split in a cubic symmetric field into a twofold degenerate state and a threefold degenerate state.

The lifting of degeneracy in the problem follows *completely* from the symmetry of the problem without any knowledge of the explicit form of the interactions! Such effects have been verified in innumerable experiments on the spectra of atoms in solids, providing one of the most beautiful examples of the power of group theoretical techniques in physics. In recent years, similar group theoretical methods have been used in elementary particle physics to obtain information about the relationship between the masses of various elementary particles, using only certain assumptions about the symmetry of the "unperturbed" elementary particle Hamiltonian and about the nature of the perturbation term which breaks the symmetry.

As a final illustration of the utility of group theoretical techniques in physics, we shall modify Eq. (10.85) slightly to obtain an interesting result concerning the transitions between quantum-mechanical states. We have already considered the case where A_1 is invariant under all $\mathscr{U} \in \mathscr{G}_1$. Suppose now that A_1 transforms under $\mathscr{G}_1$ according to some irreducible representation of $\mathscr{G}_1$. This would be the case if A_1 were one of a collection of operators, which we might write as $\{A_1^{(\lambda,m)}, m = 1, 2, \cdots, n_\lambda\}$, where λ indicates the irreducible representation under which the set of $A_1^{(\lambda,m)}$ transform; i.e.,

$$\mathscr{U} A_1^{(\lambda,m)} \mathscr{U}^{-1} = \sum_{m'=1}^{n_\lambda} D_{m'm}^{(\lambda)}(U) A_1^{(\lambda,m')}\,.$$

Thus we would rewrite Eq. (10.85) as

$$M_{n,\nu,i;\,n',\mu,j}^{(\lambda,m)} = \sum_{k=1}^{n_\nu}\sum_{l=1}^{n_\mu}\sum_{m'=1}^{n_\lambda} D_{ki}^{(\nu)*}(U) D_{lj}^{(\mu)}(U) D_{m'm}^{(\lambda)}(U)\,(\psi_{n,\nu,k},\, A_1^{(\lambda,m')}\psi_{n',\mu,l})\,, \quad (10.92)$$

where we have defined

$$M^{(\lambda,m)}_{n,\nu,i;\,n',\mu,j} \equiv (\psi_{n,\nu,i},\, A^{(\lambda,m)}_1 \psi_{n',\mu,j}) \,.$$

Thus Eq. (10.92) takes the form

$$M^{(\lambda,m)}_{n,\nu,i,\,n',\mu,j} = \sum_{k=1}^{n_\nu} \sum_{l=1}^{n_\mu} \sum_{m'=1}^{n_\lambda} D^{(\nu)*}_{ki}(U) D^{(\mu)}_{lj}(U) D^{(\lambda)}_{m'm}(U) M^{(\lambda,m')}_{n,\nu,k,\,n',\mu,l} \,. \tag{10.93}$$

What can we say about the product $D^{(\mu)}_{lj}(U) D^{(\lambda)}_{m'm}(U)$? According to Eq. (3.60) this is equal to the lm', jm-element of the direct product of $D^{(\mu)}$ and $D^{(\lambda)}$:

$$[D^{(\mu)}(U) \otimes D^{(\lambda)}(U)]_{lm',jm} = D^{(\mu)}_{lj}(U) D^{(\lambda)}_{m'm}(U) \,.$$

Thus Eq. (10.93) becomes

$$M^{(\lambda,m)}_{n,\nu,i;\,n',\mu,j} = \sum_{k=1}^{n_\nu} \sum_{l=1}^{n_\mu} \sum_{m'=1}^{n_\lambda} D^{(\nu)*}_{ki}(U) [D^{(\mu)}(U) \otimes D^{(\lambda)}(U)]_{lm',jm} M^{(\lambda,m')}_{n,\nu,k;\,n',\mu,l} \,. \tag{10.94}$$

Now since the direct product of two representations is again a representation, the quantity in square brackets in Eq. (10.94) will be a representation matrix element, but even though $D^{(\lambda)}$ and $D^{(\mu)}$ are *irreducible*, their direct product will usually be *reducible*. The number of times that an irreducible representation of G occurs in the direct product is easily determined [by Eq. (10.54)] if we know the characters of the irreducible representations, since the character of U corresponding to $D^{(\mu)}(U) \otimes D^{(\lambda)}(U)$ is just $\chi(U) = \chi^{(\mu)}(U)\chi^{(\lambda)}(U)$, according to Eq. (3.61). Let us label the irreducible representations occurring in $D^{(\mu)} \otimes D^{(\lambda)}$ by the elements ρ, of a set $R_{\lambda\mu}$ of integers. Then we can write

$$[D^{(\mu)}(U) \otimes D^{(\lambda)}(U)]_{lm',jm} = \sum_{\rho \in R_{\lambda\mu}} \sum_{\sigma,\tau=1}^{n_\rho} a^{(\lambda\mu\rho)}_{lm',jm;\,\sigma\tau} D^{(\rho)}_{\sigma\tau}(U) \,. \tag{10.95}$$

Putting Eq. (10.95) into Eq. (10.94) and, as before, summing both sides of Eq. (10.94) over all $U \in G$, we find that

$$M^{(\lambda,m)}_{n,\nu,i;\,n',\mu,j} = \frac{1}{n_\nu} \sum_{\rho \in R_{\mu\lambda}} \delta_{\nu\rho} \sum_{k,l,m'} a^{(\lambda\mu\nu)}_{lm',jm;\,ik} M^{(\lambda,m')}_{n,\nu,k;\,n',\mu,l} \,. \tag{10.96}$$

Without any further work, we can see that unless $\nu \in R_{\lambda\mu}$ the expression in Eq. (10.96) vanishes. In other words, unless the direct product of $D^{(\lambda)}$ and $D^{(\mu)}$ contains the irreducible representation $D^{(\nu)}$, the quantity $M^{(\lambda,m)}_{n,\nu,i,n',\mu,j}$ must equal zero for any choice of i, j, and m.

The usefulness of this result can most easily be seen by considering the rate of transitions between two quantum-mechanical states of a system which is invariant under some symmetry group. Suppose the levels belong respectively to the νth and μth representations of the symmetry group. It is known from quantum mechanics that if A_1 is the operator which induces the transition, then the transition rate is proportional to

$$|(\psi_{n,\nu,i},\, A_1\psi_{n',\mu,j})|^2 \,,$$

where we label the wave functions in the usual manner. If A_1 transforms according to some representation λ of the symmetry group in question, we see that unless the νth representation is contained in the direct product of the μth representation times the λth representation, the transition cannot occur. This restriction follows purely from the symmetry of the problem.

The theory of selection rules of quantum mechanics follow from these considerations. For example, a rotationally invariant atomic (or nuclear) system has energy levels which, as we have seen, can be classified according to the irreducible representations of the rotation group. Now the operator which induces electric dipole transitions between atomic levels is proportional to the electronic position vector $\mathbf{r}$ which transforms according to the $l = 1$ representation of the rotation group, as we saw in Chapter 1. Thus for a transition to take place from a level labeled by l to a level labeled by l', the l' representation of the rotation group must occur in the direct product of the l-representation times the 1-representation. The study of the decomposition of the direct product of representations is thus very important for the applications of group theory to physics, since such a study leads to a direct determination of the selection rules which play a central role in atomic, nuclear, and elementary particle physics.

PROBLEMS

1. Consider the plane of Fig. 10.3, where the objects at A, B, and C are fixed to the circular loop. This figure is clearly invariant under rotation through 120° and 240° and under reflection through the lines a, b, and c. Show that these five operations plus the identity form a group. Write out the group multiplication table. What are the classes into which the various group elements fall?

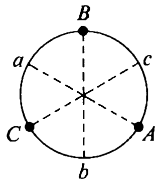

Fig. 10.3

2. Consider the group built up from the unit matrix and the Pauli spin matrices. The elements of this group are $\pm I$, $\pm i\sigma_x$, $\pm i\sigma_y$, and $\pm i\sigma_z$, where
$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$
What is the multiplication table? Find all subgroups and classes. Are there any invariant subgroups?
3. Find the symmetry group, including only rotations, of a square, rectangular parallelepiped (i.e., square on the top and bottom and rectangular on the sides). Show that this group is isomorphic to the symmetry group of the square which

includes both rotations and reflection. What are the classes? Show that there is an invariant subgroup. Find the factor group associated with this invariant subgroup.

4. Determine all the elements of the group of rotational symmetries of the regular tetrahedron (i.e., a tetrahedron whose sides are equilateral triangles). Put these elements into the classes to which they belong. Find all the subgroups. Are there any invariant subgroups?
5. Consider an arbitrary group, G, with a subgroup, H. Show that if H is of index two, then H is invariant. Remember that the index of a subgroup is the ratio of the order of the group to the order of the subgroup.
6. Find the subgroup of S_5 to which the cyclic group of order five is isomorphic. Demonstrate the isomorphism explicitly by performing all the relevant multiplications of the elements of S_5.
7. Let $(123\cdots l)$ be a cycle of length l. Prove that if $\lambda < l$ then $(123\cdots l)^\lambda \neq e$, while $(123\cdots l)^l = e$.
8. Let G and G' be two groups. Show that in a homomorphism of G onto G' the elements of G which are mapped into the identity of G' form an invariant subgroup, H, of G.
9. Show that the group G' of the previous problem is isomorphic to the factor group G/H, where H is as defined in the previous problem.
10. a) By direct calculation, show that

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

is invariant under rotation through the Euler angles α, β and γ. To refresh your memory about Euler angles, see Section 1.4.

[*Hint*: Note that the most general rotation is built up from three simpler rotations.]

b) Determine how the three components of the quantum mechanical angular momentum [see Eqs. (5.86)] transform under the group of rotations about the z-axis.

11. Show that the set $P_n = [1, x, x^2, \cdots, x^{n-1}]$ transforms into itself under the action of the translation group in one dimension. Denoting a translation through a distance a by T_a (that is, $T_a x = x - a$)· find the n-dimensional representation of the group of all one-dimensional translations which is provided by the set of functions P_n. Is this representation equivalent to a unitary representation? Show explicitly for the case of the three-dimensional representation that $T_a T_b = T_{a+b}$.
12. Show that the following three sets of functions, $[\cos 3\theta]$, $[\sin 3\theta]$ and $[\cos 4\theta, \sin 4\theta]$, transform among themselves under the elements of the group discussed in Problem 10.1. What are the representations and characters which are obtained from these three sets of functions? Are they irreducible representations? What can you say about the representation provided by $[\cos 5\theta, \sin 5\theta]$?
13. Show that if C_i is a class containing $S_1, S_2, \cdots, S_{g_i}$ and $C_{i'}$ is the collection of elements consisting of $S_1^{-1}, S_2^{-1}, \cdots, S_{g_i}^{-1}$, then $C_{i'}$ is also a class. Show that for a general irreducible representation, i.e., one which is not necessarily unitary,

$$\sum_i g_i \chi_i^{(\mu)} \chi_{i'}^{(\nu)} = g\delta_{\mu\nu}, \qquad a_\mu = \frac{1}{g} \sum_i g_i \chi_i \chi_{i'}^{(\mu)},$$

where a_μ is the number of times the μth irreducible representation occurs in the representation (not necessarily irreducible) whose characters are χ_i.

14. Show that any 2×2 matrix which commutes with the three Pauli matrices given in Problem 10.2 must be a multiple of the unit matrix.
[*Hint*: Show that any 2×2 matrix can be written as a linear combination of the unit matrix and the Pauli matrices.]

15. Show that the most general 2×2 unitary matrix whose eigenvalues are 1 and -1 can be written as

$$U = \begin{pmatrix} \cos\theta & \sin\theta e^{i\phi} \\ \sin\theta e^{-i\phi} & -\cos\theta \end{pmatrix}.$$

16. Find the characters and irreducible representations of the group discussed in Problem 10.3.
[*Hint*: For the two-dimensional representation, the result of Problem 10.15 may be useful.]

17. Using the three-dimensional representation of the rotation group, given, for example, by Eq. (1.25), find the matrices for a three-dimensional representation of the symmetry group, $\mathcal{O}$, of the cube. We saw in Section 10.8 that there were *two* three-dimensional representations of $\mathcal{O}$. Give a simple prescription for finding the matrices belonging to the second three-dimensional representation (i.e., the one not given by the rotation group).

18. Using the results of Problem 10.4, find the characters and representations of the group of rotational symmetries of the regular tetrahedron.

19. Suppose that we have a linear operator with cubic symmetry, and we apply a small perturbation having the symmetry of a rectangular parallelepiped. Making use of the results of Problem 10.16 and the work of Section 10.8, find how the various possible degenerate states of cubic symmetry break up when this perturbation is applied.

20. Consider a linear operator, A_0, with rotational symmetry. Suppose that a small perturbation having the symmetry of a regular tetrahedron is applied. Using the results of Problem 10.18 and of Section 10.8, find how the four states of A_0 with the lowest degeneracies (onefold, threefold, fivefold, and sevenfold) break up when this perturbation is applied.

21. Consider all possible representations of the cubic group, $\mathcal{O}$, formed by taking the direct product of the various irreducible representations of $\mathcal{O}$. Find out how many times the irreducible representations of $\mathcal{O}$ occur in these product representations.

22. In quantum mechanics one defines an electron spin wave function using α for spin "up" and β for spin "down"; α and β are orthonormal. For a system of three electrons we can form product wave functions of the form $\alpha(1)\alpha(2)\alpha(3)$, $\alpha(1)\alpha(2)\beta(3)$, etc., there being $2^3 = 8$ such products. Clearly, these eight functions transform among themselves under the action of the elements of S_3, and they are mutually orthogonal. They therefore form a set of basis functions for a representation of S_3. Find how many times the various irreducible representations of S_3 occur in this representation. Is any representation missing? Why? Can you relate some of the objects you're dealing with to physics?

23. Suppose that all the eight product spin functions of the previous problem are degenerate eigenfunctions of some linear operator, A_0. Let A_0 be perturbed by an operator of the form $A_1 = AP_{12} + BP_{13} + CP_{23}$, where P_{ij} transforms coordinate i into coordinate j and vice versa. For example,

$$P_{12}\alpha(1)\beta(2)\alpha(3) = \alpha(2)\beta(1)\alpha(3) .$$

Show that the matrix which must be diagonalized to obtain the first-order corrections to the eigenvalue of A_0 to which the product functions belong can be written in the form

$$\begin{bmatrix} E & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & E & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & E & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & E & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & E_{11} & E_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 & E_{12} & E_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & E_{11} & E_{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & E_{12} & E_{22} \end{bmatrix},$$

where $E = A + B + C$, $E_{11} = A - (B + C)/2$, $E_{12} = \sqrt{3}(B - C)/2$ and $E_{22} = -E_{11}$, if we use the representation of S_3 given in Section 10.7.

24. If in the previous problem we had used a different (but unitarily equivalent) two-dimensional representation of S_3, we would have found different values for E_{11}, E_{12}, and E_{22}. Would the eigenvalues of the matrix of Problem 10.23 be different? Why?

FURTHER READINGS

DIXON, J. D., *Problems in Group Theory.* Waltham, Mass.: Blaisdell, 1967. Contains solutions.

LEDERMANN, W., *The Theory of Finite Groups.* New York: Interscience, 1964.

HAMERMESH, M., *Group Theory.* Reading, Mass.: Addison-Wesley, 1962.

TINKHAM, M., *Group Theory and Quantum Mechanics.* New York: McGraw-Hill, 1964.

WEYL, H., *Theory of Groups and Quantum Mechanics.* London: Methuen, 1931

WIGNER, E. P., *Group Theory.* New York: Academic Press, 1959.