2

Representation Theory and Basic Theorems

In this chapter we introduce the concept of a representation of an abstract group and prove a number of important theorems relating to irreducible representations, including the "Wonderful Orthogonality Theorem." This mathematical background is necessary for developing the group theoretical framework that is used for the applications of group theory to solid state physics.

2.1 Important Definitions

Definition 12. *Two groups are* isomorphic *or* homomorphic *if there exists a correspondence between their elements such that*

$$\begin{aligned} A &\rightarrow \hat{A} \\ B &\rightarrow \hat{B} \\ AB &\rightarrow \hat{A}\hat{B}\ , \end{aligned}$$

where the plain letters denote elements in one group and the letters with carets denote elements in the other group. If the two groups have the same order (same number of elements), then they are isomorphic *(one-to-one correspondence). Otherwise they are* homomorphic *(many-to-one correspondence).*

For example, the permutation group of three numbers $P(3)$ is *isomorphic* to the symmetry group of the equilateral triangle and *homomorphic* to its factor group, as shown in Table 2.1. Thus, the homomorphic representations in Table 2.1 are *unfaithful*. Isomorphic representations are *faithful*, because they maintain the one-to-one correspondence.

Definition 13. *A* representation *of an abstract group is a substitution group (matrix group with square matrices) such that the substitution group is homomorphic (or isomorphic) to the abstract group. We assign a matrix $D(A)$ to each element A of the abstract group such that $D(AB) = D(A)D(B)$.*

Table 2.1. Table of homomorphic mapping of $P(3)$ and its factor group

permutation group element		factor group
E, D, F	$\rightarrow$	$\mathcal{E}$
A, B, C	$\rightarrow$	$\mathcal{A}$

The matrices of (1.4) are an isomorphic representation of the permutation group $P(3)$. In considering the representation

$$\left.\begin{matrix} E \\ D \\ F \end{matrix}\right\} \rightarrow \quad (1) \qquad \left.\begin{matrix} A \\ B \\ C \end{matrix}\right\} \rightarrow \quad (-1)$$

the one-dimensional matrices (1) and (-1) are a homomorphic representation of $P(3)$ and an isomorphic representation of the factor group $\mathcal{E}, \mathcal{A}$ (see Sect. 1.7). The homomorphic one-dimensional representation (1) is a representation for any group, though an unfaithful one.

In quantum mechanics, the matrix representation of a group is important for several reasons. First of all, we will find that an eigenfunction for a quantum mechanical operator will transform under a symmetry operation similar to the application of the matrix representing the symmetry operation on the matrix for the wave function. Secondly, quantum mechanical operators are usually written in terms of a matrix representation, and thus it is convenient to write symmetry operations using the same kind of matrix representation. Finally, matrix algebra is often easier to manipulate than geometrical symmetry operations.

2.2 Matrices

Definition 14. *Hermitian matrices are defined by:* $\tilde{A} = A^*$, $\tilde{A}^* = A$, *or* $A^\dagger = A$ (*where the symbol* $*$ *denotes complex conjugation,* $\sim$ *denotes transposition, and* $\dagger$ *denotes taking the adjoint*)

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix} , \tag{2.1}$$

$$\tilde{A} = \begin{pmatrix} a_{11} & a_{21} & \cdots \\ a_{12} & a_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix} , \tag{2.2}$$

$$A^\dagger = \begin{pmatrix} a_{11}^* & a_{21}^* & \cdots \\ a_{12}^* & a_{22}^* & \cdots \\ \vdots & \vdots & \end{pmatrix} . \tag{2.3}$$

Unitary matrices are defined by: $\tilde{A}^* = A^\dagger = A^{-1}$*;*
Orthonormal matrices are defined by: $\tilde{A} = A^{-1}$*.*

Definition 15. *The dimensionality of a representation is equal to the dimensionality of each of its matrices, which is in turn equal to the number of rows or columns of the matrix.*

These representations are *not unique*. For example, by performing a similarity (or equivalence, or canonical) transformation $UD(A)U^{-1}$ we generate a new set of matrices which provides an equally good representation. A simple physical example for this transformation is the rotation of reference axes, such as (x, y, z) to (x', y', z'). We can also generate another representation by taking one or more representations and combining them according to

$$\begin{pmatrix} D(A) & \mathcal{O} \\ \mathcal{O} & D'(A) \end{pmatrix} , \tag{2.4}$$

where $\mathcal{O} = (m \times n)$ matrix of zeros, not necessarily a square zero matrix. The matrices $D(A)$ and $D'(A)$ can be either two distinct representations or they can be identical representations.

To overcome the difficulty of non-uniqueness of a representation with regard to a similarity transformation, we often just deal with the *traces* of the matrices which are invariant under similarity transformations, as discussed in Chap. 3. The *trace* of a matrix is defined as the sum of the diagonal matrix elements. To overcome the difficulty of the ambiguity of representations in general, we introduce the concept of *irreducible* representations.

2.3 Irreducible Representations

Consider the representation made up of two distinct or identical representations for every element in the group

$$\begin{pmatrix} D(A) & \mathcal{O} \\ \mathcal{O} & D'(A) \end{pmatrix} .$$

This is a reducible representation because the matrix corresponding to each and every element of the group is in the same block form. We could now carry out a similarity transformation which would mix up all the elements so that the matrices are no longer in block form. But still the representation is reducible. Hence the definition:

Definition 16. *If by one and the same equivalence transformation, all the matrices in the representation of a group can be made to acquire the same block form, then the representation is said to be* reducible*; otherwise it is* irreducible*. Thus, an irreducible representation cannot be expressed in terms of representations of lower dimensionality.*

We will now consider three irreducible representations for the permutation group $P(3)$:

$$\begin{array}{cccc}
 & E & A & B \\
\Gamma_1 : & (1) & (1) & (1) \\
\Gamma_{1'} : & (1) & (-1) & (-1) \\
\Gamma_2 : & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \\
 & C & D & F \\
\Gamma_1 : & (1) & (1) & (1) \\
\Gamma_{1'} : & (-1) & (1) & (1) \\
\Gamma_2 : & \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} & \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} & \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} .
\end{array} \tag{2.5}$$

A reducible representation containing these three irreducible representations is

$$\begin{array}{cccc}
 & E & A & B \\
\Gamma_\mathrm{R} : & \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix} & \begin{pmatrix} 1&0&0&0 \\ 0&-1&0&0 \\ 0&0&-1&0 \\ 0&0&0&1 \end{pmatrix} & \begin{pmatrix} 1&0&0&0 \\ 0&-1&0&0 \\ 0&0&\frac{1}{2}&-\frac{\sqrt{3}}{2} \\ 0&0&-\frac{\sqrt{3}}{2}&-\frac{1}{2} \end{pmatrix} \cdots ,
\end{array} \tag{2.6}$$

where Γ_R is of the form

$$\left(\begin{array}{c|c|c} \Gamma_1 & 0 & \mathcal{O} \\ \hline 0 & \Gamma_{1'} & \mathcal{O} \\ \hline \mathcal{O} & \mathcal{O} & \Gamma_2 \end{array}\right) . \tag{2.7}$$

It is customary to list the irreducible representations contained in a reducible representation Γ_R as

$$\Gamma_\mathrm{R} = \Gamma_1 + \Gamma_{1'} + \Gamma_2 . \tag{2.8}$$

In working out problems of physical interest, each irreducible representation describes the transformation properties of a set of eigenfunctions and corresponds to a distinct energy eigenvalue. Assume Γ_R is a reducible representation for some group G but an irreducible representation for some other group G'. If Γ_R contains the irreducible representations $\Gamma_1 + \Gamma_{1'} + \Gamma_2$ as illustrated earlier for the group $P(3)$, this indicates that some interaction is breaking up a fourfold degenerate level in group G' into three energy levels in group G: two nondegenerate ones and a doubly degenerate one. Group theory does not tell us what these energies are, nor their ordering. Group theory only specifies the symmetries and degeneracies of the energy levels. In general, the higher the symmetry, meaning the larger the number of symmetry operations in the group, the higher the degeneracy of the energy levels. Thus when a perturbation is applied to lower the symmetry, the degeneracy of the energy levels tends to be reduced. Group theory provides a systematic method for determining exactly how the degeneracy is lowered.

Representation theory is useful for the treatment of physical problems because of certain orthogonality theorems which we will now discuss. To prove the orthogonality theorems we need first to prove some other theorems (including the unitarity of representations in Sect. 2.4 and the two Schur lemmas in Sects. 2.5 and 2.6).

2.4 The Unitarity of Representations

The following theorem shows that in most physical cases, the elements of a group can be represented by unitary matrices, which have the property of preserving length scales. This theorem is then used to prove lemmas leading to the proof of the "Wonderful Orthogonality Theorem," which is a central theorem of this chapter.

Theorem. *Every representation with matrices having nonvanishing determinants can be brought into unitary form by an equivalence (similarity) transformation.*

Proof. By unitary form we mean that the matrix elements obey the relation $(A^{-1})_{ij} = A^{\dagger}_{ij} = A^{*}_{ji}$, where A is an arbitrary matrix of the representation. The proof is carried out by actually finding the corresponding unitary matrices if the A_{ij} matrices are not already unitary matrices.

Let $A_1, A_2, \cdots, A_h$ denote matrices of the representation. We start by forming the matrix sum

$$H = \sum_{x=1}^{h} A_x A_x^{\dagger} \,, \tag{2.9}$$

where the sum is over all the elements in the group and where the adjoint of a matrix is the transposed complex conjugate matrix $(A_x^{\dagger})_{ij} = (A_x)^{*}_{ji}$. The matrix H is Hermitian because

$$H^{\dagger} = \sum_x (A_x A_x^{\dagger})^{\dagger} = \sum_x A_x A_x^{\dagger} \,. \tag{2.10}$$

Any Hermitian matrix can be diagonalized by a suitable unitary transformation. Let U be a unitary matrix made up of the orthonormal eigenvectors which diagonalize H to give the diagonal matrix d:

$$\begin{aligned} d &= U^{-1} H U \\ &= \sum_x U^{-1} A_x A_x^{\dagger} U \\ &= \sum_x U^{-1} A_x U U^{-1} A_x^{\dagger} U \\ &= \sum_x \hat{A}_x \hat{A}_x^{\dagger} \,, \end{aligned} \tag{2.11}$$

where we define $\hat{A}_x = U^{-1}A_xU$ for all x. The diagonal matrix d is a *special* kind of matrix and contains only real, positive diagonal elements since

$$
\begin{aligned}
d_{kk} &= \sum_x \sum_j (\hat{A}_x)_{kj}(\hat{A}_x^\dagger)_{jk} \\
&= \sum_x \sum_j (\hat{A}_x)_{kj}(\hat{A}_x)^*_{kj} \\
&= \sum_x \sum_j |(\hat{A}_x)_{kj}|^2 \, .
\end{aligned} \tag{2.12}
$$

Out of the diagonal matrix d, one can form two matrices ($d^{1/2}$ and $d^{-1/2}$) such that

$$
d^{1/2} \equiv \begin{pmatrix} \sqrt{d_{11}} & & \mathcal{O} \\ & \sqrt{d_{22}} & \\ \mathcal{O} & & \ddots \end{pmatrix} \tag{2.13}
$$

and

$$
d^{-1/2} \equiv \begin{pmatrix} \frac{1}{\sqrt{d_{11}}} & & \mathcal{O} \\ & \frac{1}{\sqrt{d_{22}}} & \\ \mathcal{O} & & \ddots \end{pmatrix} , \tag{2.14}
$$

where $d^{1/2}$ and $d^{-1/2}$ are real, diagonal matrices. We note that the generation of $d^{-1/2}$ from $d^{1/2}$ requires that none of the d_{kk} vanish. These matrices clearly obey the relations

$$(d^{1/2})^\dagger = d^{1/2} \tag{2.15}$$

$$(d^{-1/2})^\dagger = d^{-1/2} \tag{2.16}$$

$$(d^{1/2})(d^{1/2}) = d \tag{2.17}$$

so that

$$d^{1/2}d^{-1/2} = d^{-1/2}d^{1/2} = \hat{1} = \text{unit matrix} \, . \tag{2.18}$$

From (2.11) we can also write

$$d = d^{1/2}d^{1/2} = \sum_x \hat{A}_x \hat{A}_x^\dagger \, . \tag{2.19}$$

We now define a new set of matrices

$$\hat{\hat{A}}_x \equiv d^{-1/2}\hat{A}_x d^{1/2} \tag{2.20}$$

and

$$\hat{A}_x^\dagger = (U^{-1}A_xU)^\dagger = U^{-1}A_x^\dagger U \tag{2.21}$$

$$\hat{\hat{A}}_x^\dagger = (d^{-1/2}\hat{A}_x d^{1/2})^\dagger = d^{1/2}\hat{A}_x^\dagger d^{-1/2} \, . \tag{2.22}$$

We now show that the matrices $\hat{\hat{A}}_x$ are unitary:

$$
\begin{aligned}
\hat{\hat{A}}_x \hat{\hat{A}}_x^\dagger &= (d^{-1/2} \hat{A}_x d^{1/2})(d^{1/2} \hat{A}_x^\dagger d^{-1/2}) \\
&= d^{-1/2} \hat{A}_x d \hat{A}_x^\dagger d^{-1/2} \\
&= d^{-1/2} \sum_y \hat{A}_x \hat{A}_y \hat{A}_y^\dagger \hat{A}_x^\dagger d^{-1/2} \\
&= d^{-1/2} \sum_y (\hat{A}_x \hat{A}_y)(\hat{A}_x \hat{A}_y)^\dagger d^{-1/2} \\
&= d^{-1/2} \sum_z \hat{A}_z \hat{A}_z^\dagger \; d^{-1/2}
\end{aligned}
\tag{2.23}
$$

by the rearrangement theorem (Sect. 1.4). But from the relation

$$
d = \sum_z \hat{A}_z \hat{A}_z^\dagger \tag{2.24}
$$

it follows that $\hat{\hat{A}}_x \hat{\hat{A}}_x^\dagger = \hat{1}$, so that $\hat{\hat{A}}_x$ is unitary.

Therefore we have demonstrated how we can always construct a unitary representation by the transformation:

$$
\hat{\hat{A}}_x = d^{-1/2} U^{-1} A_x U d^{1/2} \;, \tag{2.25}
$$

where

$$
H = \sum_{x=1}^{h} A_x A_x^\dagger \tag{2.26}
$$

$$
d = \sum_{x=1}^{h} \hat{A}_x \hat{A}_x^\dagger \;, \tag{2.27}
$$

and where U is the unitary matrix that diagonalizes the Hermitian matrix H and $\hat{A}_x = U^{-1} A_x U$. $\square$

Note: On the other hand, not all symmetry operations can be represented by a unitary matrix; an example of an operation which cannot be represented by a unitary matrix is the time inversion operator (see Chap. 16). Time inversion symmetry is represented by an antiunitary matrix rather than a unitary matrix. It is thus not possible to represent all symmetry operations by a unitary matrix.

2.5 Schur's Lemma (Part 1)

Schur's lemmas (Parts 1 and 2) on irreducible representations are proved in order to prove the "Wonderful Orthogonality Theorem" in Sect. 2.7. We next prove Schur's lemma Part 1.

Lemma. *A matrix which commutes with all matrices of an irreducible representation is a constant matrix, i.e., a constant times the unit matrix. Therefore, if a non-constant commuting matrix exists, the representation is reducible; if none exists, the representation is irreducible.*

Proof. Let M be a matrix which commutes with all the matrices of the representation $A_1, A_2, \ldots, A_h$ □

$$MA_x = A_xM \ . \tag{2.28}$$

Take the adjoint of both sides of (2.28) to obtain

$$A_x^\dagger M^\dagger = M^\dagger A_x^\dagger \ . \tag{2.29}$$

Since A_x can in all generality be taken to be unitary (see Sect. 2.4), multiply on the right and left of (2.29) by A_x to yield

$$M^\dagger A_x = A_x M^\dagger \ , \tag{2.30}$$

so that if M commutes with A_x so does $M^\dagger$, and so do the Hermitian matrices H_1 and H_2 defined by

$$\begin{aligned} H_1 &= M + M^\dagger \\ H_2 &= i(M - M^\dagger) \ , \end{aligned} \tag{2.31}$$

$$H_jA_x = A_xH_j \ , \quad \text{where} \quad j = 1, 2 \ . \tag{2.32}$$

We will now show that a commuting Hermitian matrix is a constant matrix from which it follows that $M = H_1 - iH_2$ is also a constant matrix.

Since H_j $(j = 1, 2)$ is a Hermitian matrix, it can be diagonalized. Let U be the matrix that diagonalizes H_j (for example H_1) to give the diagonal matrix d

$$d = U^{-1}H_jU \ . \tag{2.33}$$

We now perform the unitary transformation on the matrices A_x of the representation $\hat{A}_x = U^{-1}A_xU$. From the commutation relations (2.28), (2.29), and (2.32), a unitary transformation on all matrices $H_jA_x = A_xH_j$ yields

$$\underbrace{(U^{-1}H_jU)}_{d}\underbrace{(U^{-1}A_xU)}_{\hat{A}_x} = \underbrace{(U^{-1}A_xU)}_{\hat{A}_x}\underbrace{(U^{-1}H_jU)}_{d} \ . \tag{2.34}$$

So now we have a diagonal matrix d which commutes with all the matrices of the representation. We now show that this diagonal matrix d is a constant matrix, if all the $\hat{A}_x$ matrices (and thus also the A_x matrices) form an irreducible representation. Thus, starting with (2.34)

$$d\hat{A}_x = \hat{A}_xd \tag{2.35}$$

we take the ij element of both sides of (2.35)

$$d_{ii}(\hat{A}_x)_{ij} = (\hat{A}_x)_{ij} d_{jj} \ , \tag{2.36}$$

so that

$$(\hat{A}_x)_{ij}(d_{ii} - d_{jj}) = 0 \tag{2.37}$$

for all the matrices A_x.

If $d_{ii} \neq d_{jj}$, so that the matrix d is not a constant diagonal matrix, then $(\hat{A}_x)_{ij}$ must be 0 for all the $\hat{A}_x$. This means that the similarity or unitary transformation $U^{-1}A_xU$ has brought all the matrices of the representation $\hat{A}_x$ into the same block form, since any time $d_{ii} \neq d_{jj}$ all the matrices $(\hat{A}_x)_{ij}$ are null matrices. Thus by definition the representation A_x is reducible. But we have assumed the A_x to be an irreducible representation. Therefore $(\hat{A}_x)_{ij} \neq 0$ for all $\hat{A}_x$, so that it is necessary that $d_{ii} = d_{jj}$, and Schur's lemma *Part 1* is proved.

2.6 Schur's Lemma (Part 2)

Lemma. *If the matrix representations $D^{(1)}(A_1), D^{(1)}(A_2), \ldots, D^{(1)}(A_h)$ and $D^{(2)}(A_1), D^{(2)}(A_2), \ldots, D^{(2)}(A_h)$ are two irreducible representations of a given group of dimensionality ℓ_1 and ℓ_2, respectively, then, if there is a matrix of ℓ_1 columns and ℓ_2 rows M such that*

$$MD^{(1)}(A_x) = D^{(2)}(A_x)M \tag{2.38}$$

for all A_x, then M must be the null matrix ($M = \mathcal{O}$) if $\ell_1 \neq \ell_2$. If $\ell_1 = \ell_2$, then either $M = \mathcal{O}$ or the representations $D^{(1)}(A_x)$ and $D^{(2)}(A_x)$ differ from each other by an equivalence (or similarity) transformation.

Proof. Since the matrices which form the representation can always be transformed into unitary form, we can in all generality assume that the matrices of both representations $D^{(1)}(A_x)$ and $D^{(2)}(A_x)$ have already been brought into unitary form. □

Assume $\ell_1 \leq \ell_2$, and take the adjoint of (2.38)

$$[D^{(1)}(A_x)]^\dagger M^\dagger = M^\dagger [D^{(2)}(A_x)]^\dagger \ . \tag{2.39}$$

The unitary property of the representation implies $[D(A_x)]^\dagger = [D(A_x)]^{-1} = D(A_x^{-1})$, since the matrices form a substitution group for the elements A_x of the group. Therefore we can write (2.39) as

$$D^{(1)}(A_x^{-1})M^\dagger = M^\dagger D^{(2)}(A_x^{-1}) \ . \tag{2.40}$$

Then multiplying (2.40) on the left by M yields

$$MD^{(1)}(A_x^{-1})M^\dagger = MM^\dagger D^{(2)}(A_x^{-1}) = D^{(2)}(A_x^{-1})MM^\dagger \ , \tag{2.41}$$

which follows from applying (2.38) to the element A_x^{-1} which is also an element of the group

$$MD^{(1)}(A_x^{-1}) = D^{(2)}(A_x^{-1})M \ . \tag{2.42}$$

We have now shown that if $MD^{(1)}(A_x) = D^{(2)}(A_x)M$ then $MM^\dagger$ commutes with all the matrices of representation (2) and $M^\dagger M$ commutes with all matrices of representation (1). But if $MM^\dagger$ commutes with all matrices of a representation, then by Schur's lemma (Part 1), $MM^\dagger$ is a constant matrix of dimensionality $(\ell_2 \times \ell_2)$:

$$MM^\dagger = c\,\hat{1} \ , \tag{2.43}$$

where $\hat{1}$ is the unit matrix.

First we consider the case $\ell_1 = \ell_2$. Then M is a square matrix, with an inverse

$$M^{-1} = \frac{M^\dagger}{c} \ , \quad c \neq 0 \ . \tag{2.44}$$

Then if $M^{-1} \neq \mathcal{O}$, multiplying (2.38) by M^{-1} on the left yields

$$D^{(1)}(A_x) = M^{-1}D^{(2)}(A_x)M \tag{2.45}$$

and the two representations differ by an equivalence transformation.

However, if $c = 0$ then we cannot write (2.44), but instead we have to consider $MM^\dagger = 0$

$$\sum_k M_{ik}M^\dagger_{kj} = 0 = \sum_k M_{ik}M^*_{jk} \tag{2.46}$$

for all ij elements. In particular, for $i = j$ we can write

$$\sum_k M_{ik}M^*_{ik} = \sum_k |M_{ik}|^2 = 0 \ . \tag{2.47}$$

Therefore each element $M_{ik} = 0$ so that M is a null matrix. This completes proof of the case $\ell_1 = \ell_2$ and $M = \mathcal{O}$.

Finally we prove that for $\ell_1 \neq \ell_2$, then $M = \mathcal{O}$. Suppose that $\ell_1 \neq \ell_2$, then we can arbitrarily take $\ell_1 < \ell_2$. Then M has ℓ_1 columns and ℓ_2 rows. We can make a square $(\ell_2 \times \ell_2)$ matrix out of M by adding $(\ell_2 - \ell_1)$ columns of zeros

$$\begin{array}{cc} & \ell_1 \text{ columns} \\ \ell_2 \text{ rows} & \begin{pmatrix} & 0\,0 \cdots 0 \\ & 0\,0 \cdots 0 \\ M & 0\,0 \cdots 0 \\ & \vdots\,\vdots \quad \vdots \\ & 0\,0 \cdots 0 \end{pmatrix} \end{array} = N = \text{square } (\ell_2 \times \ell_2) \text{ matrix} \ . \tag{2.48}$$

The adjoint of (2.48) is then written as

$$\begin{pmatrix} & & M^{\dagger} & & \\ & & & & \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} = N^{\dagger} \tag{2.49}$$

so that

$$NN^{\dagger} = MM^{\dagger} = c\,\hat{1} \quad \text{dimension } (\ell_2 \times \ell_2)\,. \tag{2.50}$$

$$\sum_k N_{ik} N^{\dagger}_{ki} = \sum_k N_{ik} N^{*}_{ik} = c\,\hat{1}$$

$$\sum_{ik} N_{ik} N^{*}_{ik} = c\ell_2\,.$$

But if we carry out the sum over i we see by direct computation that some of the diagonal terms of $\sum_{k,i} N_{ik} N^{*}_{ik}$ are 0, so that c must be zero. But this implies that for every element we have $N_{ik} = 0$ and therefore also $M_{ik} = 0$, so that M is a null matrix, completing the proof of Schur's lemma *Part 2*.

2.7 Wonderful Orthogonality Theorem

The orthogonality theorem which we now prove is so central to the application of group theory to quantum mechanical problems that it was named the "Wonderful Orthogonality Theorem" by Van Vleck, and is widely known by this name. The theorem is in actuality an orthonormality theorem.

Theorem. *The orthonormality relation*

$$\sum_R D^{(\Gamma_j)}_{\mu\nu}(R) D^{(\Gamma_{j'})}_{\nu'\mu'}(R^{-1}) = \frac{h}{\ell_j} \delta_{\Gamma_j,\Gamma_{j'}} \delta_{\mu\mu'} \delta_{\nu\nu'} \tag{2.51}$$

is obeyed for all the inequivalent, irreducible representations of a group, where the summation is over all h group elements $A_1, A_2, \ldots, A_h$ and ℓ_j and $\ell_{j'}$ are, respectively, the dimensionalities of representations Γ_j and $\Gamma_{j'}$. If the representations are unitary, the orthonormality relation becomes

$$\sum_R D^{(\Gamma_j)}_{\mu\nu}(R) \left[D^{(\Gamma'_j)}_{\mu'\nu'}(R) \right]^{*} = \frac{h}{\ell_j} \delta_{\Gamma_j,\Gamma_{j'}} \delta_{\mu\mu'} \delta_{\nu\nu'}\,. \tag{2.52}$$

Example. To illustrate the meaning of the mathematical symbols of this theorem, consider the orthogonality between the Γ_1 and $\Gamma_{1'}$ irreducible representations for the $P(3)$ group in Sect. 2.5 using the statements of the theorem (2.52):

$$\sum_R D_{\mu\nu}^{(\Gamma_1)}(R) D_{\mu'\nu'}^{(\Gamma_{1'})*}(R) = [(1)\cdot(1)] + [(1)\cdot(1)] + [(1)\cdot(1)] \\ +[(1)\cdot(-1)] + [(1)\cdot(-1)] + [(1)\cdot(-1)] = 0\,. \tag{2.53}$$

Proof. Consider the $\ell_{j'} \times \ell_j$ matrix

$$M = \sum_R D^{(\Gamma_{j'})}(R) X D^{(\Gamma_j)}(R^{-1})\,, \tag{2.54}$$

where X is an arbitrary matrix with $\ell_{j'}$ rows and ℓ_j columns so that M is a rectangular matrix of dimensionality $(\ell_{j'} \times \ell_j)$. Multiply M by $D^{(\Gamma_{j'})}(S)$ for some element S in the group:

$$\underbrace{D^{(\Gamma_{j'})}(S)M}_{\ell_{j'}\times\ell_j} = \sum_R D^{(\Gamma_{j'})}(S) D^{(\Gamma_{j'})}(R)\; X\; D^{(\Gamma_j)}(R^{-1})\,. \tag{2.55}$$

We then carry out the multiplication of two elements in a group

$$\underbrace{D^{(\Gamma_{j'})}(S)M}_{\ell_{j'}\times\ell_j} = \sum_R D^{(\Gamma_{j'})}(SR)\; X\; D^{(\Gamma_j)}(R^{-1}S^{-1}) D^{(\Gamma_j)}(S)\,, \tag{2.56}$$

where we have used the group properties (1.3) of the representations Γ_j and $\Gamma_{j'}$. By the rearrangement theorem, (2.56) can be rewritten

$$D^{(\Gamma_{j'})}(S)M = \underbrace{\sum_R D^{(\Gamma_{j'})}(R)\; X\; D^{(\Gamma_j)}(R^{-1})}_{M} D^{(\Gamma_j)}(S) = M\; D^{(\Gamma_j)}(S)\,. \tag{2.57}$$

Now apply Schur's lemma Part 2 for the various cases. □

Case 1. $\ell_j \neq \ell_{j'}$ or if $\ell_j = \ell_{j'}$, and the representations are not equivalent.

Since $D^{(\Gamma_{j'})}(S)M = MD^{(\Gamma_j)}(S)$, then by Schur's lemma Part 2, M must be a null matrix. From the definition of M we have

$$0 = M_{\mu\mu'} = \sum_R \sum_{\gamma,\lambda} D_{\mu\gamma}^{(\Gamma_{j'})}(R) X_{\gamma\lambda} D_{\lambda\mu'}^{(\Gamma_j)}(R^{-1})\,. \tag{2.58}$$

But X is an arbitrary matrix. By choosing X to have an entry 1 in the $\nu\nu'$ position and 0 everywhere else, we write

$$X = \begin{pmatrix} 0\;0\;0\;0\;0\;0\cdots \\ 0\;0\;0\;1\;0\;0\cdots \\ 0\;0\;0\;0\;0\;0\cdots \\ 0\;0\;0\;0\;0\;0\cdots \\ \vdots\;\vdots\;\vdots\;\vdots\;\vdots\;\vdots \end{pmatrix}, \quad X_{\gamma\lambda} = \delta_{\gamma\nu}\delta_{\lambda\nu'}\,. \tag{2.59}$$

It then follows by substituting (2.59) into (2.58) that

$$0 = \sum_R D^{(\Gamma_{j'})}_{\mu\nu}(R) D^{(\Gamma_j)}_{\nu'\mu'}(R^{-1}) \; . \tag{2.60}$$

Case 2. $\ell_j = \ell_{j'}$ and the representations Γ_j and $\Gamma_{j'}$ are equivalent.

If the representations Γ_j and $\Gamma_{j'}$ are equivalent, then $\ell_j = \ell_{j'}$ and Schur's lemma part 1 tells us that $M = c\hat{1}$. The definition for M in (2.54) gives

$$M_{\mu\nu'} = c\delta_{\mu\mu'} = \sum_R \sum_{\gamma,\lambda} D^{(\Gamma_{j'})}_{\mu\gamma}(R) X_{\gamma\lambda} D^{(\Gamma_{j'})}_{\lambda\mu'}(R^{-1}) \; . \tag{2.61}$$

Choose X in (2.59) as above to have a nonzero entry at $\nu\nu'$ and 0 everywhere else. Then $X_{\gamma\lambda} = c'_{\nu\nu'}\delta_{\gamma\nu}\delta_{\lambda\nu'}$, so that

$$c''_{\nu\nu'}\delta_{\mu\mu'} = \sum_R D^{(\Gamma_{j'})}_{\mu\nu}(R) \; D^{(\Gamma_{j'})}_{\nu'\mu'}(R^{-1}) \; , \tag{2.62}$$

where $c''_{\nu\nu'} = c/c'_{\nu\nu'}$. To evaluate $c''_{\nu\nu'}$ choose $\mu = \mu'$ in (2.62) and sum on μ:

$$c''_{\nu\nu'} \underbrace{\sum_\mu \delta_{\mu\mu}}_{\ell_{j'}} = \sum_R \sum_\mu D^{(\Gamma_{j'})}_{\mu\nu}(R) \; D^{(\Gamma_{j'})}_{\nu'\mu}(R^{-1}) = \sum_R D^{(\Gamma_{j'})}_{\nu'\nu}(R^{-1}R) \; . \tag{2.63}$$

since $D^{(\Gamma_{j'})}(R)$ is a representation of the group and follows the multiplication table for the group. Therefore we can write

$$c''_{\nu\nu'}\ell_{j'} = \sum_R D^{(\Gamma_{j'})}_{\nu'\nu}(R^{-1}R) = \sum_R D^{(\Gamma_{j'})}_{\nu'\nu}(E) = D^{(\Gamma_{j'})}_{\nu'\nu}(E) \sum_R 1 \; . \tag{2.64}$$

But $D^{(\Gamma_{j'})}_{\nu'\nu}(E)$ is a unit $(\ell_{j'} \times \ell_{j'})$ matrix and the $\nu'\nu$ matrix element is $\delta_{\nu'\nu}$. The sum of unity over all the group elements is h. Therefore we obtain

$$c''_{\nu\nu'} = \frac{h}{\ell_{j'}}\delta_{\nu\nu'} \; . \tag{2.65}$$

Substituting (2.65) into (2.62) gives:

$$\frac{h}{\ell_{j'}}\delta_{\mu\mu'}\delta_{\nu\nu'} = \sum_R D^{(\Gamma_{j'})}_{\mu\nu}(R) \; D^{(\Gamma_{j'})}_{\nu'\mu'}(R^{-1}) \; . \tag{2.66}$$

We can write the results of Cases 1 and 2 in compact form

$$\sum_R D^{(\Gamma_j)}_{\mu\nu}(R) \; D^{(\Gamma_{j'})}_{\nu'\mu'}(R^{-1}) = \frac{h}{\ell_j}\delta_{\Gamma_j,\Gamma_{j'}}\delta_{\mu\mu'}\delta_{\nu\nu'} \; . \tag{2.67}$$

For a unitary representation (2.67) can also be written as

$$\sum_R D^{(\Gamma_j)}_{\mu\nu}(R) \; D^{(\Gamma_{j'})*}_{\mu'\nu'}(R) = \frac{h}{\ell_j}\delta_{\Gamma_j,\Gamma_{j'}}\delta_{\mu\mu'}\delta_{\nu\nu'} \; . \tag{2.68}$$

This completes the proof of the wonderful orthogonality theorem, and we see explicitly that this theorem is an orthonormality theorem.

2.8 Representations and Vector Spaces

Let us spend a moment and consider what the representations in (2.68) mean as an orthonormality relation in a vector space of dimensionality h. Here h is the order of the group which equals the number of group elements. In this space, the representations $D_{\mu\nu}^{(\Gamma_j)}(R)$ can be considered as elements in this h-dimensional space:

$$V_{\mu,\nu}^{(\Gamma_j)} = \left[D_{\mu\nu}^{(\Gamma_j)}(A_1), D_{\mu\nu}^{(\Gamma_j)}(A_2), \ldots, D_{\mu\nu}^{(\Gamma_j)}(A_h) \right] . \tag{2.69}$$

The three indices Γ_j, μ, ν label a particular vector. All distinct vectors in this space are orthogonal. Thus two representations are orthogonal if any one of their three indices is different. But in an h-dimensional vector space, the maximum number of orthogonal vectors is h. We now ask how many vectors $V_{\mu,\nu}^{(\Gamma_j)}$ can we make? For each representation, we have ℓ_j choices for μ and ν so that the total number of vectors we can have is $\sum_j \ell_j^2$ where we are now summing over representations Γ_j. This argument yields the important result

$$\sum_j \ell_j^2 \leq h . \tag{2.70}$$

We will see later (Sect. 3.7) that it is the *equality* that holds in (2.70). The result in (2.70) is extremely helpful in finding the totality of irreducible (non-equivalent) representations (see Problem 2.2).

Selected Problems

2.1. Show that every symmetry operator for every group can be represented by the (1×1) unit matrix. Is it also true that every symmetry operator for every group can be represented by the (2×2) unit matrix? If so, does such a representation satisfy the Wonderful Orthogonality Theorem? Why?

2.2. Consider the example of the group $P(3)$ which has six elements. Using the irreducible representations of Sect. 2.3, find the sum of ℓ_j^2. Does the equality or inequality in (2.70) hold? Can $P(3)$ have an irreducible representation with $\ell_j = 3$? Group $P(4)$ has 24 elements and 5 irreducible representations. Using (2.70) as an equality, what are the dimensionalities of these 5 irreducible representations (see Problem 1.4)?