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) trans-formation.*

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 $\hat{A}_x = U^{-1} A_x U$ 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

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

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

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 $\hat{A}_x = U^{-1} A_x U$. □

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 $A_1, A_2, \ldots, A_h$ □

$$MA_x = A_x M \ . \tag{2.28}$$

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

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

$$H_j A_x = A_x H_j \ , \quad \text{where} \quad j = 1, 2 \ . \tag{2.32}$$

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

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

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

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

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)

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

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

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

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*.