5. Representations of groups

Consider a set $\mathcal{W}$ of elements $\{x\}$. The set could be finite, infinite or even continuous. (For instance, it could be a vector space.) A *transformation* of the set is a one-to-one mapping of the set onto itself. Relative to multiplication defined as consecutive action of transformations, the set of transformations of $\mathcal{W}$ is a group.

> [The consecutive action of two transformations clearly produces a transformation. Consecutive action of mappings is associative. Mapping every element $x \in \mathcal{W}$ into itself is obviously an identity transformation. One-to-one onto mappings have inverses, so each transformation has an inverse.]

If the set $\mathcal{W}$ is finite, having n members, then a transformation is a permutation of the elements of $\mathcal{W}$ and the group of transformations is just the symmetric group on n elements, $\mathcal{S}_n$.

Now consider a group $\mathcal{G}$, its elements being $\{R\}$. The *action* of $\mathcal{G}$ on the set $\mathcal{W}$ is defined as a homomorphism of $\mathcal{G}$ to the group of transformations of $\mathcal{W}$. This means that each $R \in \mathcal{G}$ is mapped into a transformation T_R of $\mathcal{W}$ in such a way that $\mathsf{T}_{RS} = \mathsf{T}_R\mathsf{T}_S$ for all $R, S \in \mathcal{G}$.

> [**Note** that the transformations T act on elements of $\mathcal{W}$, mapping $x \to \mathsf{T}(x)$. Preservation of multiplication by the homomorphism means $\mathsf{T}_{RS}(x) = \mathsf{T}_R(\mathsf{T}_S(x))$. The elements of the group $\mathcal{G}$ do not generally act directly on $\mathcal{W}$. However, for reasons of economy, this distinction is frequently blurred, so one speaks of R and S acting on $\mathcal{W}$ and writes $RSx = R(Sx)$. Such statements should be understood as a shorthand for the more precise definition of the action of $\mathcal{G}$ on $\mathcal{W}$.]

Suppose the action of a group $\mathcal{G}$ on a set $\mathcal{W}$ has been defined. The *orbit* of the element $x \in \mathcal{W}$ is the subset $\{Rx,\ \text{for all } R \in \mathcal{G}\}$ of $\mathcal{W}$. It includes x itself. Let $y \neq x$ be an element of the orbit of x. By closure of $\mathcal{G}$ (and the product-preserving property of the homomorphism), the orbit of y is the same as the orbit of x. (Note again the role of the rearrangement theorem.) Belonging to a specified orbit is an equivalence relation, so the action of $\mathcal{G}$ partitions $\mathcal{W}$ into distinct orbits.

Consider further the set of elements of $\mathcal{G}$ which, acting on $\mathcal{W}$, map a given $x \in \mathcal{W}$ to itself. This set is closed under multiplication, so it is a subgroup of $\mathcal{G}$. It is called the *isotropy group* or *stabiliser* of x and denoted $\mathcal{G}_x$.

An interesting special case occurs when the set $\mathcal{W}$ is the group $\mathcal{G}$ itself. In this case, the statement that elements of $\mathcal{G}$ act on $\mathcal{W}$ is not just convenient shorthand, but is precise. Two specific actions are of particular interest. In the first, $\mathsf{T}_R S = RS$, i.e. the action is left multiplication of group elements. (The homomorphism property is ensured by associativity of group multiplication.) The whole group is a single orbit; the stabiliser of every element $R \in \mathcal{G}$ is the unit group $\{E\}$. In the second action, $\mathcal{T}_R S = RSR^{-1}$, i.e. the action is conjugation of group elements. The orbits of the action are the classes of $\mathcal{G}$; the stabiliser of an element $R \in \mathcal{G}$ is its normaliser $\mathcal{N}_R$.

Transformations of a set can be represented by matrices. First consider a finite set $\mathcal{W}$ with d elements and a transformation $\mathsf{T}x_i \to x_{T(i)}$, where $\{T(1), T(2), \ldots, T(d)\}$ is a permutation of $\{1, 2, \ldots, d\}$. The matrix $\mathcal{T}$ is defined by rewriting $x'_i = \sum_{j=1}^{d} \mathcal{T}_{ji} x_j$, where $\mathcal{T}_{ij} = \delta_{i,T(j)}$. Now apply two consecutive transformations, T and U, to $\mathcal{W}$, $\mathsf{UT}x_i \to \mathsf{U}x_{T(i)} \to x_{U(T(i))}$. Using the corresponding matrices, $\mathsf{UT}x_i = \sum_{j=1}^{d} \mathcal{T}_{ji} \mathsf{U} x_j = \sum_{j=1}^{d} \mathcal{T}_{ji} \sum_{k=1}^{d} \mathcal{U}_{kj} x_k = \sum_{k=1}^{d} (\mathcal{U}\mathcal{T})_{ki} x_k$, using standard matrix multiplication. So the product of transformations UT is represented by the product of matrices $\mathcal{U}\mathcal{T}$. Note also that $(\mathcal{U}\mathcal{T})_{ij} = \sum_{k=1}^{d} \delta_{i,U(k)} \delta_{k,T(j)} = \delta_{i,U(T(j))}$, or $(UT)(j) = U(T(j))$. The association of transformations with matrices preserves multiplication. The matrices constitute a group.

The action of $\mathcal{G}$ on $\mathcal{W}$, together with the association of a matrix with each transformation, defines a homomorphism of $\mathcal{G}$ to the group of representative matrices. This is called a *(matrix) representation* of the group. The *dimension* of the representation is the dimension of the matrices, which is just the number d of elements of the set $\mathcal{W}$.

Alternatively, and this will be the focus of attention from now on, the set $\mathcal{W}$ could be a vector space, to be denoted $\mathcal{V}$ henceforth. For application to physics, the vector space should be equipped with an inner product. The transformations of the vector space are required to be *linear* — $\mathsf{T}(a\psi_1 + b\psi_2) = a\mathsf{T}\psi_1 + b\mathsf{T}\psi_2$. It is still possible to associate a matrix with every transformation of the vector space, by choosing a basis $\{\psi_i\}$ for the space and expressing the effect of the transformation on the basis vectors with the aid of a matrix. For each ψ_i, the vector $\mathsf{T}\psi_i$ is expanded in terms of the basis vectors, $\mathsf{T}\psi_i = \sum_{j=1}^{d} \mathcal{T}_{ji} \psi_j$. For a vector space of dimension d, this defines the $d \times d$ matrix $\mathcal{T}$.

Now consider

$$\mathsf{TU}\psi_i = \sum_{j=1}^{d} \mathcal{U}_{ji} \mathsf{T}\psi_j = \sum_{j=1}^{d} \sum_{k=1}^{d} \mathcal{U}_{ji} \mathcal{T}_{kj} \psi_k = \sum_{k=1}^{d} (\mathcal{T}\mathcal{U})_{ki} \psi_k.$$

The matrix representing the product of transformations TU is the product of the matrices representing T and U. Once again, a matrix representation of the group is obtained, of dimension equal to the dimension of the vector space $\mathcal{V}$. The vector space is said to *carry* the representation and is called a *carrier space*.

> [**Important!** Pay special attention to the order of the indices i, j in the equations defining the matrices, for both the finite set $\mathcal{W}$ and the vector space $\mathcal{V}$. The indices of the matrix appear to be "backwards". This is not a mistake. If the more intuitive order of indices were chosen, the order of multiplication of matrices representing a product of transformations would be reversed. A simple mnemonic is provided by the Dirac notation, $\mathsf{T}|i\rangle = \sum_j |j\rangle\langle j|\mathsf{T}|i\rangle = \sum_j \mathcal{T}_{ji}|j\rangle$.]

Formally, a representation of a group $\mathcal{G}$ is a homomorphism from $\mathcal{G}$ to a group $\mathsf{T}(\mathcal{G})$ of linear transformations on a vector space $\mathcal{V}$, called the *carrier space* of the representation. (If the homomorphism is an isomorphism, the representation is said to be *faithful*.) Each of the operators in $\mathsf{T}(\mathcal{G})$ may be represented as a matrix, in terms of a suitable basis for $\mathcal{V}$, and the resulting set of matrices, themselves forming a group, constitute a *matrix representation* of $\mathcal{G}$. The dimension of the carrier space, and hence of the matrices, is called the *dimension* of the representation. A group may have many different representations of many different dimensions. The representative operator corresponding to an element $R \in \mathcal{G}$ will henceforth be denoted D_R, its matrix $\mathcal{D}_R$. When no confusion can arise, D_R will be abbreviated to R.

A change in the basis of the vector space $\mathcal{V}$ will generally change the representative matrices $\mathcal{D}$, as would a one-to-one mapping of $\mathcal{V}$ to another vector space $\mathcal{V}'$ of the same dimension. Such changes are implemented by a similarity transformation $\mathcal{D} \to S\mathcal{D}S^{-1}$. However, these are clearly not essentially different representations. Representations which are related by a similarity transformation of the representative matrices are called *equivalent* representations.

Equivalent representations will generally have different representative matrices. There is a useful property of a representation matrix that is the same for all equivalent representations. This is the trace of the matrix, the sum of all its diagonal elements, which exhibits *cyclic invariance* — the trace of a product of matrices is invariant under cyclic permutation of the factors.

[$\mathrm{tr}(ABC\cdots W) = \sum_{i,j,k\ldots,r} A_{ij}B_{jk}C_{kl}\ldots W_{ri}$ is clearly unchanged under a cyclic permutation of its factors, since the dummy indices retain the form $ij\ jk\ kl\ l\ldots r\ ri$.]

As a consequence of its cyclic invariance, the trace is unchanged under similarity transformation: $\mathrm{tr}(S\mathcal{D}S^{-1}) = \mathrm{tr}(S^{-1}S\mathcal{D}) = \mathrm{tr}\mathcal{D}$.

The trace of the matrix representing an element R of the group is called the *character* of that element in the given representation and is denoted χ_R. It is the same in all equivalent representations. All elements of a given class have the same character, again because of the cyclic invariance of the trace.

[$\mathrm{tr}(\mathcal{D}_{BAB^{-1}} = \mathrm{tr}(\mathcal{D}_B\mathcal{D}_A\mathcal{D}_{B^{-1}}) = \mathrm{tr}(\mathcal{D}_A\mathcal{D}_{B^{-1}}\mathcal{D}_B) = \mathrm{tr}(\mathcal{D}_{AB^{-1}B}) = \mathrm{tr}(\mathcal{D}_A)$.]

The identity element is represented in any representation by the unit matrix, so its character is the dimension of the representation. The set of characters of the classes of a group, in a given representation, can be viewed formally as a vector $(\chi_1, \chi_2, \ldots, \chi_k)$ in a space of dimension equal to the number of classes in the group. It characterises the representation.

Given an arbitrary vector $\phi \in \mathcal{V}$, the elements $R \in \mathcal{G}$ act on it to produce a set of vectors $\{\mathsf{D}_R\phi\} \in \mathcal{V}$ which is closed under the action of $\mathcal{G}$. This set of vectors generates a subspace $\mathcal{V}' \subset \mathcal{V}$ which is invariant under the action of $\mathcal{G}$ (i.e. for any $\psi \in \mathcal{V}'$ and for any $R \in \mathcal{G}$, $\mathsf{D}_R\psi \in \mathcal{V}'$) and which carries a representation of $\mathcal{G}$. The dimension of this representation cannot exceed $|\mathcal{G}|$. Clearly any invariant subspace (relative to $\mathcal{G}$) carries a representation of $\mathcal{G}$.

[A one-dimensional representation is carried by a one-dimensional vector space, i.e. the set of all multiples of a single basis vector. The carrier space is invariant under the group — $\mathsf{D}_R\psi = \mathcal{D}_R\psi$ for $R \in \mathcal{G}$, where $\mathcal{D}_R$ is a number in the field over which the vector space is defined (in practice, a complex number), so that $\chi_R = \mathcal{D}_R$. Note that it is the vector space spanned by ψ which is invariant under $\mathcal{G}$, not ψ itself, which is a simultaneous eigenvector of all the D_R. Since, in any representation, $\mathcal{D}_{R^n} = (\mathcal{D}_R)^n$ for any integer n, in a one-dimensional representation $\chi_{R^n} = (\chi_R)^n$. Suppose the element $R \in \mathcal{G}$ is of order r, i.e. $R^r = \mathbf{1}$. Then, in a one-dimensional representation, $(\chi_R)^r = \chi(\mathbf{1}) = 1$, so χ_R is an r^{th} root of unity, $\chi_R = e^{2\pi im/r}$ for some integer m, and $|\chi_R| = 1$ for every $R \in \mathcal{G}$.]

Note, in passing, that the vectors $\{f\}$ carrying a representation of a group $\mathcal{G}$ with elements $\{R_i\}$ may be functions of arguments $\{\mathbf{r}\}$ on which the group elements act. The transformed function $[\mathsf{D}_R f](\mathbf{r})$ is defined to have the same value for argument $\mathbf{r}$ as the original function $f(\mathbf{r})$ had for the argument that was transformed into $\mathbf{r}$, i.e. $[\mathsf{D}_R f](\mathbf{r}) = f(R^{-1}\mathbf{r})$. Then $[\mathsf{D}_{R_2} f](\mathbf{r}) \equiv f'(\mathbf{r}) = f(R_2^{-1}\mathbf{r}) \Longrightarrow [\mathsf{D}_{R_1}\mathsf{D}_{R_2} f](\mathbf{r}) = [\mathsf{D}_{R_1} f'](\mathbf{r}) = f'(R_1^{-1}\mathbf{r}) = f(R_2^{-1}R_1^{-1}\mathbf{r}) = f((R_1R_2)^{-1}\mathbf{r}) = [\mathsf{D}_{R_1R_2} f](\mathbf{r})$, and multiplication is preserved.

Suppose $\mathcal{V}$ is a finite-dimensional vector space (with an inner product) which carries a representation D of the finite group $\mathcal{G}$. Denote the inner product by (x, y), for any pair of vectors $x, y \in \mathcal{V}$. If $(\mathsf{D}_R x, \mathsf{D}_R y) = (x, y)$ for every $R \in \mathcal{G}$ and every pair $x, y \in \mathcal{V}$, then the representation is *unitary* and the representative matrices satisfy $\mathcal{D}_R\mathcal{D}_R^\dagger = \mathbf{1} = \mathcal{D}_R^\dagger\mathcal{D}_R$. Not every representation is unitary. But consider the alternative inner product defined on $\mathcal{V}$ by $\langle x, y\rangle = \sum_{R\in\mathcal{G}}(\mathsf{D}_R x, \mathsf{D}_R y)/|\mathcal{G}|$. (It is not hard to show that this is an acceptable inner product.) Now $\langle \mathsf{D}_R x, \mathsf{D}_R y\rangle = \sum_{S\in\mathcal{G}}(\mathsf{D}_S\mathsf{D}_R x, \mathsf{D}_S\mathsf{D}_R y)/|\mathcal{G}| = \sum_{S\in\mathcal{G}}(\mathsf{D}_{SR} x, \mathsf{D}_{SR} y)/|\mathcal{G}| = \langle x, y\rangle$, by the rearrangement theorem, so the representation is unitary with respect to the new inner product. For any finite-dimensional representation of a finite group, there is an inner product which makes the representation unitary.

Assume bases $\{u_i\}$ and $\{v_i\}$ have been chosen for the carrier space which are orthonormal relative to the original inner product and relative to the new inner product, respectively. An arbitrary vector $z \in \mathcal{V}$ is written $z = \sum_i z_i u_i$, in terms of its components in the u basis. Define the transformation T on $\mathcal{V}$ by $v_i = Tu_i$ for all i. Then $Tz = \sum_i z_i Tu_i = \sum_i z_i v_i$ has the same components in the v basis as z has in the u basis. This implies $\langle Tx, Ty\rangle = \sum_i x_i^* y_i = (x, y)$. Introduce the equivalent representation $\mathsf{D}'_R = T^{-1}\mathsf{D}_R T$, for all $R \in \mathcal{G}$. It satisfies $(\mathsf{D}'_R x, \mathsf{D}'_R y) = (T^{-1}\mathsf{D}_R Tx, T^{-1}\mathsf{D}_R Ty) = \langle \mathsf{D}_R Tx, \mathsf{D}_R Ty\rangle = \langle Tx, Ty\rangle = (x, y)$ for any $R \in \mathcal{G}$, so it is unitary with respect to the original inner product. Any finite-dimensional representation of a finite group (on a carrier space with a given inner product) is equivalent to a unitary representation (with respect to the same inner product). Unless otherwise stated, it will henceforth be assumed that the representations of finite groups under discussion are unitary.

Consider a representation in a space $\mathcal{V}$ of dimension d. If it is possible, by an appropriate transformation of the basis of $\mathcal{V}$, to bring all the representation matrices into the form $\begin{pmatrix} \mathcal{D}_R^{(1)} & \mathcal{A}_R \\ 0 & \mathcal{D}_R^{(2)} \end{pmatrix}$, where $\mathcal{D}^{(1)}, \mathcal{D}^{(2)}$ are square matrices of dimensions $d_1, d_2 < d$, then the representation is said to be *reducible*. Since products of matrices of this form retain the same form,

$\mathsf{D}^{(1)}, \mathsf{D}^{(2)}$ are separately representations of the group, of lower dimension than the original representation. The subspace of vectors of $\mathcal{V}$ whose last d_2 components vanish is invariant under the group. In general, if the carrier space of a representation contains a subspace invariant under the action of the group, then the representation is reducible.

If the off-diagonal submatrices $\mathcal{A}$ of a reducible representation can be made to vanish, the representation is said to be *fully reducible*. In this case, the complementary space to the invariant subspace is also invariant. The carrier space has been divided into invariant subspaces. This is always possible when the initial representation is unitary.

> [Let $\mathcal{V}$ be the carrier space of a unitary representation D of the group $\mathcal{G}$ and let $\mathcal{W}$ be a subspace of $\mathcal{V}$ invariant under the action of the group. Denote by $\mathcal{W}_\perp$ the orthogonal complement of $\mathcal{W}$, so that $v \in \mathcal{W}, w \in \mathcal{W}_\perp \implies (v, w) = 0$, in terms of the inner product of $\mathcal{V}$. Given $w \in \mathcal{W}_\perp$, then for any $R \in \mathcal{G}$ and for every $v \in \mathcal{W}$ $(\mathsf{D}_R w, v) = (\mathsf{D}_R w, \mathsf{D}_R \mathsf{D}_{R^{-1}} v) = (w, \mathsf{D}_{R^{-1}} v) = 0$, since D is unitary and $\mathcal{W}$ is invariant. So $\mathsf{D}_R w \in \mathcal{W}_\perp$ for all $R \in \mathcal{G}$ and all $w \in \mathcal{W}_\perp$, i.e. $\mathcal{W}_\perp$ is an invariant subspace.]

Every reducible representation of a finite group is fully reducible, and the carrier space is divided into two distinct invariant subspaces, each carrying one of the reduced representations. This procedure can be continued until it is no longer possible to break any of the carrier subspaces into invariant subspaces. The resulting "smallest" representations are then said to be *irreducible*. The original representation takes block diagonal form, with only irreducible representations occurring along the diagonal. These irreducible representations need not all be different from one another — an irreducible representation $\mathsf{D}^{(\nu)}$ may occur a_ν times (where equivalent representations are counted as multiple occurrences). This is expressed by writing $\mathsf{D} = \sum_\nu a_\nu \mathsf{D}^{(\nu)}$, with non-negative integers a_ν.

> [**Note:** Since the term "irreducible representation" occurs very often in discussions of algebraic structures, it is frequently abbreviated as "irrep". This convention will be adopted from now on.]

Since the set $\{\mathsf{D}_R \psi\}$, for all $R \in \mathcal{G}$ and for any vector ψ on which the group acts, spans an invariant subspace, it follows that no irrep of $\mathcal{G}$ can have a dimension larger than $|\mathcal{G}|$. It is evident that every group has a trivial one-dimensional irrep, referred to as the *unit irrep*, in which every element of the group is mapped into unity and all characters are 1. Every group of permutations containing odd elements has an additional one-dimensional irrep, the *alternating irrep*, in which every permutation is mapped into its alternating character and all characters are ± 1.

> [The basis vector of the unit irrep satisfies $\mathsf{D}_R\psi = \psi$, for all $R \in \mathcal{G}$. It is therefore invariant under the action of the group. Conversely, any invariant vector is a basis for the unit irrep. The operator $\sum_{R\in\mathcal{G}} \mathsf{D}_R$, an element of the group algebra, acting on an arbitrary vector ψ, produces an invariant vector (by the rearrangement theorem), unless the result vanishes.]

Assume $\mathsf{D}(\mathcal{G})$ is a representation of dimension d on the vector space $\mathcal{V}$ with basis $\{\psi_i\}$. If D is reducible, there is an invariant subspace $\mathcal{V}'$ with basis $\{\phi_i\}$ carrying a representation D' of dimension $d' < d$. Since $\mathcal{V}' \subset \mathcal{V}$, $\phi_i = \sum_{j=1}^{d} a_{ij}\psi_j$. Acting with an element $R \in \mathcal{G}$ on the left side of this equation produces $\sum_{j=1}^{d'}(\mathcal{D}'_R)_{ji}\phi_j = \sum_{j=1}^{d'}(\mathcal{D}'_R)_{ji}\sum_{k=1}^{d} a_{jk}\psi_k$, while acting on the right side produces $\sum_{j=1}^{d} a_{ij}\sum_{k=1}^{d}(\mathcal{D}_R)_{kj}\psi_k$. Since the $\{\psi_i\}$ are linearly independent, it may be concluded that there exists a $d \times d'$ matrix $\mathcal{A}$, the transpose of the matrix of coefficients a_{ij}, such that $\mathcal{D}_R\mathcal{A} = \mathcal{A}\mathcal{D}'_R$ for all $R \in \mathcal{G}$. The converse is also true — if, for a given representation D of dimension d, there exists a $d\times d'$ matrix $\mathcal{A}$ with $d' < d$ such that $\mathcal{D}_R\mathcal{A} = \mathcal{A}\mathcal{D}'_R$ for all $R \in \mathcal{G}$, then the representation D is reducible.

> [Given the basis $\{\psi_i\}$ for D, define the d' functions $\phi_i = \sum_{j=1}^{d}\mathcal{A}_{ji}\psi_j$. Acting with R, $R\phi_i = \sum_{j=1}^{d}\mathcal{A}_{ji}\sum_{k=1}^{d}(\mathcal{D}_R)_{kj}\psi_k = \sum_{k=1}^{d}(\mathcal{D}_R\mathcal{A})_{ki}\psi_k = \sum_{k=1}^{d}(\mathcal{A}\mathcal{D}'_R)_{ki}\psi_k = \sum_{k=1}^{d}\sum_{j=1}^{d'}\mathcal{A}_{kj}(\mathcal{D}'_R)_{ji}\psi_k = \sum_{j=1}^{d'}(\mathcal{D}'_R)_{ji}\phi_j$. The $d' < d$ functions $\{\phi_i\}$ span an invariant subspace of the carrier space of D, which is thus reducible.]

So a representation D is reducible if and only if there exists a matrix $\mathcal{A}$ such that $\mathcal{D}_R\mathcal{A} = \mathcal{A}\mathcal{D}'_R$ for all $R \in \mathcal{G}$ and the dimension of D' is less than that of D. This is a criterion for reducibility of a representation.