Chapter 1

Group Representations

Definition 1.1 *A* representation α *of a group* G *in a vector space* V *over* k *is defined by a homomorphism*

$$\alpha : G \to \mathbf{GL}(V).$$

The degree *of the representation is the dimension of the vector space:*

$$\deg \alpha = \dim_k V.$$

Remarks:

1. Recall that $\mathbf{GL}(V)$—the general linear group on V—is the group of invertible (or non-singular) linear maps $t : V \to V$.

2. We shall be concerned almost exclusively with representations of *finite degree*, that is, in *finite-dimensional* vector spaces; and these will almost always be vector spaces over $\mathbb{R}$ *or* $\mathbb{C}$. Therefore, to avoid repetition, let us agree to use the term 'representation' to mean *representation of finite degree over* $\mathbb{R}$ *or* $\mathbb{C}$, unless the contrary is explicitly stated.

   Furthermore, in this first Part we shall be concerned almost exclusively with *finite* groups; so let us also agree that the term 'group' will mean *finite group* in this Part, unless the contrary is stated.

3. Suppose $\{e_1, \ldots, e_n\}$ is a basis for V. Then each linear map $t : V \to V$ is defined (with respect to this basis) by an $n \times n$-matrix T. Explicitly,

$$te_j = \sum_i T_{ij} e_i;$$

or in terms of coordinates,

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Thus a representation in V can be defined by a homomorphism

$$\alpha : G \to \mathbf{GL}(n, k),$$

where $\mathbf{GL}(n, k)$ denotes the group of non-singular $n \times n$-matrices over k. In other words, α is defined by giving matrices $A(g)$ for each $g \in G$, satisfying the conditions

$$A(gh) = A(g)A(h)$$

for all $g, h \in G$; and also

$$A(e) = I.$$

4. There is another way of looking at group representations which is almost always more fruitful than the rather abstract definition we started with.

   Recall that a group is said to *act* on the set X if we have a map

   $$G \times X \to X : (g, x) \mapsto gx$$

   satisfying

   (a) $(gh)x) = g(hx)$,

   (b) $ex = x$.

   Now suppose $X = V$ is a vector space. Then we can say that G acts linearly on V if in addition

   (c) $g(u + v) = gu + gv$,

   (d) $g(\rho v) = \rho(gv)$.

   Each representation α of G in V defines a linear action of G on V, by

   $$gv = \alpha(g)v;$$

   and every such action arises from a representation in this way.

   Thus the notions of *representation* and *linear action* are completely equivalent. We can use whichever we find more convenient in a given case.

5. There are two other ways of looking at group representations, completely equivalent to the definition but expressing slightly different points of view.

   Firstly, we may speak of the vector space V, with the action of G on it. as a G-*space*. For those familiar with category theory, this would be the categorical approach. Representation theory, from this point of view, is the study of the category of G-spaces and G-maps, where a G-map

   $$t : U \to V$$

   from one G-space to another is a linear map preserving the action of G, ie satisfying

   $$t(gu) = g(tu) \quad (g \in G, u \in U).$$

6. Secondly, and finally, mathematical physicists often speak—strikingly—of the vector space V *carrying* the representation α.

Examples:

1. Recall that the dihedral group D_4 is the symmetry group of a square $ABCD$

   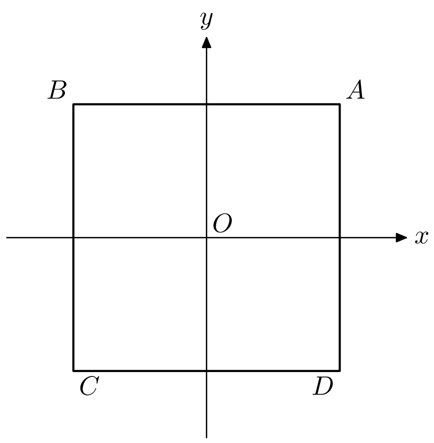

   Figure 1.1: The natural representation of D_4

   (Figure **??**). Let us take coordinates Ox, Oy as shown through the centre O of the square. Then

   $$D_4 = \{e, r, r^2, r^3, c, d, x, y\},$$

   where r is the rotation about O through $\pi/2$ (sending A to B), while c, d, x, y are the reflections in AC, BD, Ox, Oy respectively.

By definition a symmetry $g \in D_4$ is an isometry of the plane E^2 sending the square into itself. Evidently g must send O into itself, and so gives rise to a linear map

$$A(g) : \mathbb{R}^2 \to \mathbb{R}^2.$$

The map

$$g \mapsto A(g) \in \mathbf{GL}(2, \mathbb{R})$$

defines a 2-dimensional representation ρ of D_4 over $\mathbb{R}$. We may describe this as the *natural* 2-dimensional representation of D_4.

(Evidently the symmetry group G of any bounded subset $S \subset E^n$ will have a similar 'natural' representation in $\mathbb{R}^n$.)

The representation ρ is given in matrix terms by

$$e \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ r \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},\ r^2 \mapsto \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},\ r^3 \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$
$$c \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ d \mapsto \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},\ x \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ y \mapsto \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Each group relation is represented in a corresponding matrix equation, eg

$$cd = r^2 \Longrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The representation ρ is *faithful*, ie the homomorphism defining it is injective. Thus a relation holds in D_4 if and only if the corresponding matrix equation is true. However, representations are not necessarily faithful, and in general the implication is only one way.

Every finite-dimensional representation can be expressed in matrix form in this way, after choosing a basis for the vector space carrying the representation. However, while such matrix representations are reassuringly concrete, they are impractical except in the lowest dimensions. Better just to keep at the back of one's mind that a representation *could* be expressed in this way.

2. Suppose G acts on the set X:

$$(g, x) \mapsto gx.$$

Let

$$C(X) = C(X, k)$$

denote the space of maps

$$f : X \to k.$$

Then G acts linearly on $C(X)$—and so defines a representation ρ of G—by

$$gf(x) = f(g^{-1}x).$$

(We need g^{-1} rather than g on the right to satisfy the rule

$$g(hf) = (gh)f.$$

It is fortunate that the relation

$$(gh)^{-1} = h^{-1}g^{-1}$$

enables us to correct the order reversal. We shall often have occasion to take advantage of this, particularly when dealing—as here—with spaces of functions.)

Now suppose that X is finite; say

$$X = \{x_1, \ldots, x_n\}.$$

Then

$$\deg \rho = n = \|X\|,$$

the number of elements in X. For the functions

$$e_y(x) = \begin{cases} 1 \text{ if } x = y, \\ 0 \text{ otherwise.} \end{cases}$$

(ie the characteristic functions of the 1-point subsets) form a basis for $C(X)$. Also

$$ge_x = e_{gx},$$

since

$$\begin{aligned} ge_y(x) &= e_y(g^{-1}x) \\ &= \begin{cases} 1 & \text{if } g^{-1}x = y \\ 0 & \text{if } g^{-1}x \neq y \end{cases} \\ &= \begin{cases} 1 & \text{if } x = gy \\ 0 & \text{if } x \neq gy \end{cases} \end{aligned}$$

Thus

$$g \mapsto P(g)$$

where $P = P(g)$ is the matrix with entries

$$P_{xy} = \begin{cases} 1 \text{ if } y = gx, \\ 0 \text{ otherwise.} \end{cases}$$

Notice that P is a *permutation matrix*, ie there is just one 1 in each row and column, all other entries being 0. We call a representation that arises from the action of a group on a set in this way a *permutational representation*.

As an illustration, consider the natural action of $S(3)$ on the set $\{a, b, c\}$. This yields a 3-dimensional representation ρ of $S(3)$, under which

$$\begin{aligned} (abc) &\mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\ (ab) &\mapsto \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

(These two instances actually define the representation, since (abc) and (ab) generate $S(3)$.)

3. A *1-dimensional* representation α of a group G over $k = \mathbb{R}$ or $\mathbb{C}$ is just a homomorphism

$$\alpha : G \to k^\times,$$

where $k^\times$ denotes the multiplicative group on the set $k \setminus \{0\}$. For

$$\mathbf{GL}(1, k) = k^\times,$$

since we can identify the 1×1-matrix $[x]$ with its single entry x.

We call the 1-dimensional representation defined by the identity homomorphism

$$g \mapsto 1$$

(for all $g \in G$) the *trivial representation* of G, and denote it by 1.

In a 1-dimensional representation, each group element is represented by a number. Since these numbers commute, the study of 1-dimensional representations is much simpler than those of higher dimension.

In general, when investigating the representations of a group G, we start by determining all its 1-dimensional representations.

Recall that two elements $g, h \in G$ are said to be *conjugate* if

$$h = xgx^{-1}$$

for some third element $x \in G$. Suppose α is a 1-dimensional representation of G. Then

$$\begin{aligned}\alpha(h) &= \alpha(x)\alpha(g)\alpha(x^{-1})\\ &= \alpha(x)\alpha(g)\alpha(x)^{-1}\\ &= \alpha(g)\alpha(x)\alpha(x)^{-1}\\ &= \alpha(g),\end{aligned}$$

since the numbers $\alpha(x), \alpha(g)$ commute. It follows that a 1-dimensional representation *is constant on each conjugacy class of* G.

Consider the group S_3. This has 3 classes (we shall usually abbreviate 'conjugacy class' to *class*):

$$\{1\},\ \{(abc),(acb)\},\ \{(bc),(ca),(ab)\}.$$

Let us write

$$s = (abc),\ t = (bc).$$

Then (assuming $k = \mathbb{C}$)

$$\begin{aligned} s^3 = 1 &\implies \alpha(s)^3 = 1 \Longrightarrow \alpha(s) = 1, \omega \text{ or } \omega^2,\\ t^2 = 1 &\implies \alpha(s)^2 = 1 \Longrightarrow \alpha(t) = \pm 1.\end{aligned}$$

But

$$tst^{-1} = s^2.$$

It follows that

$$\alpha(t)\alpha(s)\alpha(t)^{-1} = \alpha(s)^2,$$

from which we deduce that

$$\alpha(s) = 1.$$

It follows that S_3 has just two 1-dimensional representations: the trivial representation

$$1 : g \mapsto 1,$$

and the *parity representation*

$$\epsilon : g \mapsto \begin{cases} 1 & \text{if } g \text{ is even,}\\ -1 & \text{if } g \text{ is odd.}\end{cases}$$

4. The corresponding result is true for all the symmetric groups S_n (for $n \geq 2$); S_n has just two 1-dimensional representations, the trivial representation 1 and the parity representation ϵ.

   To see this, let us recall two facts about S_n.

(a) The transpositions $\tau = (xy)$ *generate* S_n, ie each permutation $g \in S_n$ is expressible (not uniquely) as a product of transpositions

$$g = \tau_1 \cdots \tau_r.$$

(b) The transpositions are all conjugate.

(This is a particular case of the general fact that two permutations in S_n are conjugate if and only if they are of the same *cyclic type*, ie they have the same number of cycles of each length.)

It follows from (1) that a 1-dimensional representation of S_n is completely determined by its values on the transpositions. It follows from (2) that the representation is constant on the transpositions. Finally, since each transposition τ satisfies $\tau^2 = 1$ it follows that this constant value is ± 1. Thus there can only be two 1-dimensional representations of S_n; the first takes the value 1 on the transpositions, and so is 1 everywhere; the second takes the value -1 on the transpositions, and takes the value $(-1)^r$ on the permutation

$$g = \tau_1 \cdots \tau_r.$$

Thus S_n has just two 1-dimensional representations; the trivial representation 1 and the parity representation ϵ.

5. Let's look again at the dihedral group D_4, ie the symmetry group of the square $ABCD$. Let r denote the rotation through $\pi/2$, taking A into B; and let c denote the reflection in AC.

   It is readily verified that r and c generate D_4, ie each element $g \in D_4$ is expressible as a *word* in r and c (eg $g = r^2cr$). This follows for example from Lagrange's Theorem. The subgroup generated by r and c contains at least the 5 elements $1, r, r^2, r^3, c$, and so must be the whole group. (We shall sometimes denote the identity element in a group by 1, while at other times we shall use e or I.)

   It is also easy to see that r and c satisfy the relations

$$r^4 = 1,\ c^2 = 1,\ rc = cr^3.$$

   In fact these are *defining relations* for D_4, ie every relation between r and c can be derived from these 3.

   We can express this in the form

$$D_4 = \langle r, c : r^4 = c^2 = 1,\ rc = cr^3 \rangle.$$

Now suppose α is a 1-dimensional representation of D_4. Then we must have

$$\alpha(r)^4 = \alpha(c)^2 = 1,\ \alpha(r)\alpha(c) = \alpha(c)\alpha(r)^3.$$

From the last relation

$$\alpha(r)^2 = 1.$$

Thus there are just 4 possibilities

$$\alpha(r) = \pm 1,\ \alpha(c) = \pm 1.$$

It is readily verified that all 4 of these satisfy the 3 defining relations for s and t. It follows that each defines a homomorphism

$$\alpha : D_4 \to k^\times.$$

We conclude that D_4 has just 4 1-dimensional representations.

6. We look now at some examples from chemistry and physics. It should be emphasized, firstly that the theory is completely independent of these examples, which can safely be ignored; and secondly, that *we are not on oath when speaking of physics*. It would be inappropriate to delve too deeply here into the physical basis for the examples we give.

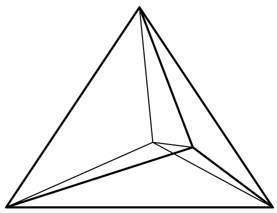

Figure 1.2: The methane molecule

First let us look at the methane molecule CH_4. In its stable state the 4 hydrogen atoms are situated at the vertices of a regular tetrahedron, with the single carbon atom at its centroid (Figure **??**).

The molecule evidently has symmetry group S_4, being invariant under permutations of the 4 hydrogen atoms.

Now suppose the molecule is vibrating about this stable position. We suppose that the carbon atom at the centroid remains fixed. (We shall return to this point later.) Thus the configuration of the molecule at any moment is defined by the displacement of the 4 hydrogen atoms, say

$$X_i = (x_{i1}, x_{i2}, x_{i3}) \quad (i = 1, 2, 3, 4).$$

Since the centroid remains fixed,

$$\sum_i x_{ij} = 0 \quad (j = 1, 2, 3).$$

This reduces the original 12 degrees of freedom to 9.

Now let us assume further that the *angular* momentum also remains 0, ie the molecule is not slowly rotating. This imposes a further 3 conditions on the x_{ij}, leaving 6 degrees of freedom for the 12 'coordinates' x_{ij}. Mathematically, the coordinates are constrained to lie in a 6-dimensional space. In other words we can find 6 'generalized coordinates' $q_1, \ldots, q_6$ — chosen so that $q_1 = q_2 = \cdots = q_6 = 0$ at the point of equilibrium — such that each of the x_{ij} is expressible in terms of the q_k:

$$x_{ij} = x_{ij}(q_1, \ldots, q_6).$$

The motion of the molecule is governed by the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial K}{\partial \dot{q}_k}\right) = -\frac{\partial V}{\partial q_k}$$

where K is the kinetic energy of the system, and V its potential energy. (These equations were developed for precisely this purpose, to express the motion of a system whose configuration is defined by generalized coordinates.)

The kinetic energy of the system is given in terms of the mass m of the hydrogen atom by

$$K = \frac{1}{2} m \sum_{i,j} \dot{x}_{ij}^2$$

On substituting

$$\dot{x_{ij}} = \frac{\partial x_{ij}}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial x_{ij}}{\partial q_6}\dot{q}_6,$$

we see that

$$K = K(\dot{q}_1, \ldots, \dot{q}_6),$$

where K is a positive-definite quadratic form. Although the coefficients of this quadratic form are actually functions of $q_1, \ldots, q_6$, we may suppose them constant since we are dealing with small vibrations.

The potential energy of the system, which we may take to have minimal value 0 at the stable position, is given to second order by some positive-definite quadratic form Q in the q_k:

$$V = Q(q_1, \ldots, q_6) + \ldots .$$

While we could explicitly choose the coordinates q_k, and determine the kinetic energy K, the potential energy form Q evidently depends on the forces holding the molecule together. Fortunately, we can say a great deal about the vibrational modes of the molecule without knowing anything about these forces.

Since these two forms are positive-definite, we can simultaneously diagonalize them, ie we can find new generalized coordinates $z_1, \ldots, z_6$ such that

$$\begin{aligned} K &= \dot{z}_1^2 + \cdots + \dot{z}_6^2, \\ V &= \omega_1^2 z_1^2 + \cdots + \omega_6^2 z_6^2. \end{aligned}$$

The Euler-Lagrange equations now give

$$\ddot{z}_i = -\omega_i^2 z_i \quad (i = 1, \ldots, 6).$$

Thus the motion is made up of 6 independent harmonic oscillations, with frequencies $\omega_1, \ldots, \omega_6$.

As usual when studying harmonic or wave motion, life is easier if we allow complex solutions (of which the 'real' solutions will be the real part). Each harmonic oscillation then has 1 degree of freedom:

$$z_j = C_j e^{i\omega_j t}.$$

The set of all solutions of these equations (ie all possible vibrations of the system) thus forms a 6-dimensional *solution-space* V.

So far we have made no use of the S_4-symmetry of the CH_4 molecule. But now we see that this symmetry group acts on the solution space V, which thus carries a representation, ρ say, of S_4. Explicitly, suppose $\pi \in S_4$ is a permutation of the 4 H atoms. This permutation is 'implemented' by a unique spatial isometry Π. (For example, the permutation $(123)(4)$ is effected by rotation through $1/3$ of a revolution about the axis joining the C atom to the 4th H atom.)

But now if we apply this isometry Π to any vibration $v(t)$ we obtain a new vibration $\Pi v(t)$. In this way the permutation π acts on the solution-space V.

In general, *the symmetry group G of the physical configuration will act on the solution-space V.*

The fundamental result in the representation theory of a finite group G (as we shall establish) is that every representation ρ of G splits into parts, each corresponding to a 'simple' representation of G. Each finite group has a finite number of such simple representations, which thus serve as the 'atoms' out of which every representation of G is constructed. (There is a close analogy with the Fundamental Theorem of Arithmetic, that every natural number is uniquely expressible as a product of primes.)

The group S_4 (as we shall find) has 5 simple representations, of dimensions $1, 1, 2, 3, 3$. Our 6-dimensional representation must be the 'sum' of some of these.

It is not hard to see that there is just one 1-dimensional mode (up to a scalar multiple) corresponding to a 'pulsing' of the molecule in which the 4 H atoms move in and out (in 'sync') along the axes joining them to the central C atom. (Recall that S_4 has just two 1-dimensional representations: the trivial representation, under which each permutation leaves everything unchanged, and the parity representation, in which even permutations leave things unchanged, while odd permutations reverse them. In our case, the 4 atoms must move in the same way under the trivial representation, while their motion is reversed under an odd permutation. The latter is impossible. For by considering the odd permutation $(12)(3)(4)$ we deduce that the first atom is moving out while the second moves in; while under the action of the even permutation $(12)(34)$ the first and second atoms must move in and out together.)

We conclude (not rigorously, it should be emphasized!) that

$$\rho = 1 + \alpha + \beta$$

where 1 denotes the trivial representation of S_4, α is the unique 2-dimensional representation, and β is one of the two 3-dimensional representations.

Thus without any real work we've deduced quite a lot about the vibrations of CH_4.

Each of these 3 modes has a distinct frequency. To see that, note that our system — and in fact any similar non-relativistic system — has a *time symmetry* corresponding to the additive group $\mathbb{R}$. For if $(z_j(t) : 1 \leq j \leq 6)$ is one solution then $(z_j(t + c))$ is also a solution for any constant $c \in \mathbb{R}$.

The simple representations of $\mathbb{R}$ are just the 1-dimensional representations

$$t \mapsto e^{i\omega t}.$$

(We shall see that the simple representations of an *abelian* group are always 1-dimensional.) In effect, Fourier analysis — the splitting of a function or motion into parts corresponding to different frequencies — is just the representation theory of $\mathbb{R}$.

The actions of S_4 and $\mathbb{R}$ on the solution space commute, giving a representation of the product group $\mathcal{S}_4 \times \mathbb{R}$.

As we shall see, the simple representations of a product group $G \times H$ arise from simple representations of G and H: $\rho = \sigma \times \tau$. In the present case we must have

$$\rho = 1 \times E(\omega_1) + \alpha \times E(\omega_2) + \beta \times E(\omega_3),$$

where $\omega_1, \omega_2, \omega_3$ are the frequencies of the 3 modes.

If the symmetry is slightly broken, eg by placing the vibrating molecule in a magnetic field, these 'degenerate' frequencies will split, so that 6 frequencies will be seen: $\omega_1', \omega_2', \omega_2'', \omega_3', \omega_3'', \omega_3'''$, where eg ω_2' and ω_2'' are close to ω. This is the origin of 'multiple lines' in spectroscopy.

The concept of *broken symmetry* has become one of the corner-stones of mathematical physics. In 'grand unified theories' distinct particles are seen as identical (like our 4 H atoms) under some large symmetry group, whose action is 'broken' in our actual universe.

7. *Vibrations of a circular drum. [**?**].* Consider a circular elastic membrane. The motion of the membrane is determined by the function

$$z(x, y, t) \qquad (x^2 + y^2 \le r^2)$$

where z is the height of the point of the drum at position (x, y).

It is not hard to establish that under small vibrations this function will satisfy the wave equation

$$T\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right) = \rho \frac{\partial^2 z}{\partial t^2},$$

where T is the tension of the membrane and ρ its mass per unit area. This may be written

$$\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right) = \frac{1}{c^2}\frac{\partial^2 z}{\partial t^2},$$

where $c = (T/\rho)^{1/2}$ is the *speed* of the wave motion.

The configuration has $O(2)$ symmetry, where $O(2)$ is the group of 2-dimensional isometries leaving the centre O fixed, consisting of the rotations about O and the reflections in lines through O.

Although this group is not finite, it is *compact*. As we shall see, the representation theory of compact groups is essentially identical to the finite theory; the main difference being that a compact group has a countable infinity of simple representations.

For example, the group $O(2)$ has the trivial representation 1, and an infinity of representations $R(1), R(2), \ldots$, each of dimension 2.

The circular drum has corresponding modes $M(0), M(1), M(2), \ldots$, each with its characteristic frequency. As in our last example, after taking time symmetry into account, the solution-space carries a representation ρ of the product group $O(2) \times \mathbb{R}$, which splits into

$$1 \times E(\omega_0) + R(1) \times E(\omega_1) + R(2) \times E(\omega_2) + \cdots .$$

8. In the last example but one, we considered the 4 hydrogen atoms in the methane molecule as particles, or solid balls. But now let us consider a single hydrogen atom, consisting of an electron moving in the field of a massive central proton.

   According to classical non-relativistic quantum mechanics [**?**], the state of the electron (and so of the atom) is determined by a *wave function* $\psi(x, y, z, t)$, whose evolution is determined by *Schrödinger's equation*

   $$i\hbar \frac{\partial \psi}{\partial t} = H\psi.$$

   Here H is the *hamiltonian operator*, given by

   $$H\psi = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(r)\psi,$$

   where m is the mass of the electron, $V(t)$ is its potential energy, $\hbar$ is Planck's constant, and

   $$\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}.$$

   Thus Schrodinger's equation reads, in full,

   $$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right) - \frac{e^2}{r} \psi.$$

   The essential point is that this is a *linear* differential equation, whose solutions therefore form a vector space, the *solution-space*.

We regard the central proton as fixed at O. (A more accurate account might take O to be the centre of mass of the system.) The system is invariant under the orthogonal group $\mathbf{O}(3)$, consisting of all isometries — that is, distance-preserving transformations — which leave O fixed. Thus the solution space carries a representation of the compact group $\mathbf{O}(3)$.

This group is a product-group:

$$\mathbf{O}(3) = \mathbf{SO}(3) \times C_2,$$

where $C_2 = \{I, J\}$ (J denoting reflection in O), while $\mathbf{SO}(3)$ is the subgroup of orientation-preserving isometries. In fact, each such isometry is a rotation about some axis, so $\mathbf{SO}(3)$ is group of rotations in 3 dimensions.

The rotation group $\mathbf{SO}(3)$ has simple representations $D_0, D_1, D_2, \ldots$ of dimensions $1, 3, 5, \ldots$. To each of these corresponds a *mode* of the hydrogen atom, with a particular frequency ω and corresponding energy level $E = \hbar\omega$.

These energy levels are seen in the spectroscope, although the spectral lines of hydrogen actually correspond to *differences* between energy levels, since they arise from photons given off when the energy level changes.

This idea — considering the space of atomic wave functions as a representation of $\mathbf{SO}(3)$ gave the first explanation of the periodic table of the elements, proposed many years before by Mendeleev on purely empirical grounds [**?**].

The discussion above ignores the *spin* of the electron. In fact representation theory hints strongly at the existence of spin, since the 'double-covering' $\mathbf{SU}(2)$ of $\mathbf{SO}(3)$ adds the 'spin representations' $D_{1/2}, D_{3/2}, \ldots$ of dimensions $2, 4, \ldots$ to the sequence above, as we shall see.

Finally, it is worth noting that quantum theory (as also electrodynamics) are *linear* theories, where the Principle of Superposition rules. Thus the application of representation theory is exact, and not an approximation restricted to small vibrations, as in classical mechanical systems like the methane molecule, or the drum.

9. *The classification of elementary particles.* [**?**]. Consider an elementary particle E, eg an electron, in relativistic quantum theory. The possible states of E again correspond to the points of a vector space V. More precisely, they correspond to the points of the *projective space* $P(V)$ formed by the *rays*, or 1-dimensional subspaces, of V. For the wave functions ψ and $\rho\psi$ correspond to the same state of E.

The state space V is now acted on by the *Poincaré group* $E(1,3)$ formed by the isometries of Minkowski space-time. It follows that V carries a representation of $E(1,3)$.

Each elementary particle corresponds to a simple representation of the Poincaré group $E(1,3)$. This group is not compact. It is however a *Lie group*; and — as we shall see — a different approach to representation theory, based on *Lie algebras*, allows much of the theory to be extended to this case.

A last remark. One might suppose, from its reliance on linearity, that representation theory would have no rôle to play in curved space-time. But that is far from true. Even if the underlying topological space is curved, the vector and tensor *fields* on such a space preserve their linear structure. (So one can, for example, superpose vector fields on a sphere.) Thus representation theory can still be applied; and in fact, the so-called *gauge theories* introduced in the search for a unified 'theory of everything' are of precisely this kind.

Bibliography

[1] C. A. Coulson. *Waves*. Oliver and Boyd, 1955.

[2] Richard Feynman. *Lecture Notes on Physics III*. Addison-Wesley, 196.

[3] D. J. Simms. *Lie Groups and Quantum Mechanics*. Springer, 1968.

[4] Hermann Weyl. *The Theory of Groups and Quantum Mechanics*. Dover, 1950.

Exercises

All representations are over $\mathbb{C}$, unless the contrary is stated.

In Exercises 01–11 determine all 1-dimensional representations of the given group.

1 $*$ C_2	2 $**$ C_3	3 $**$ C_n	4 $**$ D_2	5 $**$ D_3
6 $***$ D_n	7 $***$ Q_8	8 $***$ A_4	9 $****$ A_n	10 $**$ $\mathbb{Z}$
11 $****$ $D_\infty = \langle r, s : s^2 = 1, rsr = s\rangle$				

Suppose G is a group; and suppose $g, h \in G$. The element $[g, h] = ghg^{-1}h^{-1}$ is called the *commutator* of g and h. The subgroup $G' \equiv [G, G]$ is generated by all commutators in G is called the commutator subgroup, or *derived group* of G.

12 $***$ Show that G' lies in the kernel of any 1-dimensional representation ρ of G, ie $\rho(g)$ acts trivially if $g \in G'$.

13 $***$ Show that G' is a normal subgroup of G, and that G/G' is abelian. Show moreover that if K is a normal subgroup of G then G/K is abelian if and only if $G' \subset K$. [In other words, G' is the smallest normal subgroup such that G/G' is abelian.)

14 $**$ Show that the 1-dimensional representations of G form an abelian group G^* under multiplication. [Nb: this notation G^* is normally only used when G is abelian.]

15 $**$ Show that $C_n^* \cong C_n$.

16 $***$ Show that for any 2 groups G, H

$$(G \times H)^* = G^* \times H^*.$$

17 $****$ By using the Structure Theorem on Finite Abelian Groups (stating that each such group is expressible as a product of cyclic groups) or otherwise, show that

$$A^* \cong A$$

for any finite abelian group A.

18 $**$ Suppose $\Theta : G \to H$ is a homomorphism of groups. Then each representation α of H defines a representation $\Theta\alpha$ of G.

19 $***$ Show that the 1-dimensional representations of G and of G/G' are in one-one correspondence.

In Exercises 20–24 determine the derived group G' of the given group G.

20 $***$ C_n	21 $****$ D_n	22 $**$ $\mathbb{Z}$	23 $****$ D_∞
24 $***$ Q_8	25 $***$ S_n	26 $***$ A_4	27 $****$ A_n

Chapter 2

Equivalent Representations

> Every mathematical theory starts from some notion of equivalence—an agreement not to distinguish between objects that 'look the same' in some sense.

Definition 2.1 *Suppose α, β are two representations of G in the vector spaces U, V over k. We say that α and β are* equivalent, *and we write $\alpha = \beta$, if U and V are isomorphic G-spaces.*

In other words, we can find a linear map

$$t : U \to V$$

which preserves the action of G, ie

$$t(gu) = g(tu) \quad \text{for all } g \in G, u \in U.$$

Remarks:

1. Suppose α and β are given in matrix form:

   $$\alpha : g \mapsto A(g), \quad \beta : g \mapsto B(g).$$

   If $\alpha = \beta$, then U and V are isomorphic, and so in particular $\dim \alpha = \dim \beta$, ie the matrices $A(g)$ and $B(g)$ are of the same size.

   Suppose the linear map $t : U \to V$ is given by the matrix P. Then the condition $t(gu) = g(tu)$ gives

   $$B(g) = PA(g)P^{-1}$$

   for each $g \in G$. This is the condition in matrix terms for two representations to be equivalent.

2. Recall that two $n \times n$ matrices S, T are said to be *similar* if there exists a non-singular (invertible) matrix P such that

$$T = PSP^{-1}.$$

A necessary condition for this is that A, B have the same eigenvalues. For the characteristic equations of two similar matrices are identical:

$$\begin{aligned} \det\left(PSP^{-1} - \lambda I\right) &= \det P \det(S - \lambda I) \det P^{-1} \\ &= \det(S - \lambda I). \end{aligned}$$

3. In general this condition is necessary but not sufficient. For example, the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

have the same eigenvalues 1,1, but are not similar. (No matrix is similar to the identity matrix I except I itself.)

However, there is one important case, or particular relevance to us, where the converse is true. Let us recall a result from linear algebra.

An $n \times n$ complex matrix A is diagonalisable if and only if it satisfies a separable polynomial equation, ie one without repeated roots.

It is easy to see that if A is diagonalisable then it satisfies a separable equation. For if

$$A \sim \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_1 & & \\ & & & \lambda_2 & \\ & & & & \ddots \end{pmatrix}$$

then A satisfies the separable equation

$$m(x) \equiv (x - \lambda_1)(x - \lambda_2) \cdots = 0.$$

The converse is less obvious. Suppose A satisfies the polynomial equation

$$p(x) \equiv (x - \lambda_1) \cdots (x - \lambda_r) = 0$$

with $\lambda_1, \ldots, \lambda_r$ distinct. Consider the expression of $1/p(x)$ as a sum of partial fractions:

$$\frac{1}{p(x)} = \frac{a_1}{x - \lambda_1} + \cdots + \frac{a_r}{x - \lambda_r}.$$

Multiplying across,

$$1 = a_1 Q_1(x) + \cdots + a_r Q_r(x),$$

where

$$Q_i(x) = \prod_{j \neq i} (x - \lambda_j) = \frac{p(x)}{x - \lambda_i}.$$

Substituting $x = A$,

$$I = a_1 Q_1(A) + \cdots + a_r Q_r(A).$$

Applying each side to the vector $v \in V$,

$$\begin{aligned} v &= a_1 Q_1(A) v + \cdots + a_r Q_r(A) v \\ &= v_1 + \cdots + v_r, \end{aligned}$$

say. The vector v_i is an eigenvector of A with eigenvalue λ_i, since

$$(A - \lambda_i) v_i = a_i p(A) v = 0.$$

Thus every vector is expressible as a sum of eigenvectors. In other words the eigenvectors of A span the space.

But that is precisely the condition for A to be diagonalisable. For we can find a basis for V consisting of eigenvectors, and with respect to this basis A will be diagonal.

4. It is important to note that while each matrix $A(g)$ is diagonalisable *separately*, we cannot in general diagonalise all the $A(g)$ *simultaneously*. That would imply that the $A(g)$ *commuted*, which is certainly not the case in general.

5. However, we can show that *if* $A_1, A_2, \ldots$ *is a set of commuting matrices then they can be diagonalised simultaneously if and only if they can be diagonalised separately.*

   To see this, suppose λ is an eigenvalue of A_1. Let

   $$E = \{v : A_1 v = \lambda v\}$$

   be the corresponding eigenspace. Then E is stable under all the A_i, since

   $$v \in E \Longrightarrow A_1(A_i v) = A_i A_1 v = \lambda A_i v \Longrightarrow A_i v \in E.$$

Thus we have reduced the problem to the simultaneous diagonalisation of the restrictions of $A_2, A_3, \ldots$ to the eigenspaces of A_1. A simple inductive argument on the degree of the A_i yields the result.

In our case, this means that we can diagonalise some (or all) of our representation matrices

$$A(g_1), A(g_2), \ldots$$

if and only it these matrices commute.

This is perhaps best seen as a result on the representations of abelian groups, which we shall meet later.

6. To summarise, two representations α, β are certainly *not* equivalent if $A(g), B(g)$ have different eigenvalues for some $g \in G$.

   Suppose to the contrary that $A(g), B(g)$ have the same eigenvalues for all $g \in G$. Then as we have seen

   $$A(g) \sim B(g)$$

   for all g, ie

   $$B(g) = P(g)A(g)P(g)^{-1}$$

   for some invertible matrix $P(g)$.

   Remarkably, we shall see that if this is so for all $g \in G$, then in fact α and β are equivalent. In other words, if such a matrix $P(g)$ exists for all g then we can find a matrix P *independent of* g such that

   $$B(g) = PA(g)P^{-1}$$

   for all $g \in G$.

7. Suppose $A \sim B$, ie

   $$B = PAP^{-1}.$$

   We can interpret this as meaning that A and B represent the same linear transformation, under the change of basis defined by P.

   Thus we can think of two equivalent representations as being, if effect, the *same* representation looked at from two points of view, that is, taking two different bases for the representation-space.

Example: Let us look again at the natural 2-dimensional real representation ρ of the symmetry group D_4 of the square $ABCD$. Recall that when we took coordinates with respect to axes Ox, Oy bisecting DA, AB, ρ took the matrix form

$$s \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad c \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

where s is the rotation through a right-angle (sending A to B), and c is the reflection in AC.

Now suppose we choose instead the axes OA, OB. Then we obtain the equivalent representation

$$s \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad c \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We observe that c has the same eigenvalues, ± 1, in both cases.

Since we have identified equivalent representations, it makes sense to ask for *all the representations of a given group G of dimension d,* say. What we have to do in such a case is to give a list of d-dimensional representations, prove that every d-dimensional representation is equivalent to one of them, and show also that no two of the representations are equivalent.

It isn't at all obvious that the number of such representations is finite, even after we have identified equivalent representations. We shall see later that this is so: *a finite group G has only a finite number of representations of each dimension.*

Example: Let us find all the 2-dimensional representations over $\mathbb{C}$ of

$$S_3 = \langle s, t : s^3 = t^2 = 1, st = ts^2 \rangle,$$

that is, all 2-dimensional representations *up to equivalence*.

Suppose α is a representation of $S(3)$ in the 2-dimensional vector space V. Consider the eigenvectors of s. There are 2 possibilities:

1. s has an eigenvector e with eigenvalue $\lambda \neq 1$. Since $s^3 = 1$, it follows that $\lambda^3 = 1$, ie $\lambda = \omega$ or ω^2.

   Now let $f = te$. Then

   $$sf = ste = ts^2 e = \lambda^2 te = \lambda^2 f.$$

   Thus f is also an eigenvector of s, although now with eigenvector λ^2.

   Since e and f are eigenvectors corresponding to different eigenvalues, they must be linearly independent, and therefore span (and in fact form a basis for) V:

   $$V = \langle e, f \rangle.$$

   Since $se = \lambda e,\ sf = \lambda^2 f$, we see that s is represented with respect to this basis by the matrix

   $$s \mapsto \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^2 \end{pmatrix}.$$

On the other hand, $te = f,\ tf = t^2 e = e$, and so

$$t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The 2 cases $\lambda = \omega, \omega^2$ give the representations

$$\begin{aligned} \alpha: \quad & s \mapsto \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \\ \beta: \quad & s \mapsto \begin{pmatrix} \omega^2 & 0 \\ 0 & \omega \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \end{aligned}$$

In fact these 2 representations are equivalent,

$$\alpha = \beta,$$

since one is got from the other by the swapping the basis elements: $e, f \mapsto f, e$.

2. The alternative possibility is that both eigenvalues of s are equal to 1. In that case, since s is diagonalisable, it follows that

$$s \mapsto I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

with respect to some basis. But then it follows that this remains the case with respect to every basis: s is always represented by the matrix I.

In particular, s is always diagonal. So if we diagonalise c—as we know we can—then we will simultaneously diagonalise s and c, and so too all the elements of D_4.

Suppose

$$s \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

Then it is evident that

$$s \mapsto 1,\ t \mapsto \lambda$$

and

$$s \mapsto 1,\ t \mapsto \mu$$

will define two 1-dimensional representations of S_3. But we know these representations; there are just 2 of them. In combination, these will give 4

2-dimensional representations of S_3. However, two of these will be equivalent. The 1-dimensional representations 1 and ϵ give the 2-dimensional representation

$$s \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(Later we shall denote this representation by $1 + \epsilon$, and call it the *sum* of 1 and ϵ.)

On the other hand, ϵ and 1 in the opposite order give the representation

$$s \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This is equivalent to the previous case, one being taken into the other by the change of coordinates $(x, y) \mapsto (y, x)$. (In other words, $\epsilon + 1 = 1 + \epsilon$.)

We see from this that we obtain just 3 2-dimensional representations of S_3 in this way (in the notation above they will be $1 + 1$, $1 + \epsilon$ and $\epsilon + \epsilon$).

Adding the single 2-dimensional representation from the first case, we conclude that S_3 has just 4 2-dimensional representations.

It is easy to see that no 2 of these 4 representations are equivalent, by considering the eigenvalues of s and c in the 4 cases.

Exercises

All representations are over $\mathbb{C}$, unless the contrary is stated.

In Exercises 01–15 determine all 2-dimensional representations (up to equivalence) of the given group.

1 $**$ C_2	2 $**$ C_3	3 $**$ C_n	4 $***$ D_2	5 $***$ D_4
6 $***$ D_5	7 $****$ D_n	8 $***$ S_3	9 $****$ S_4	10 $*****$ S_n
11 $****$ A_4	12 $*****$ A_n	13 $***$ Q_8	14 $**$ $\mathbb{Z}$	15 $****$ D_∞

16 $***$ Show that a real matrix $A \in \mathbf{Mat}(n, \mathbb{R})$ is diagonalisable over $\mathbb{R}$ if and only if its minimal polynomial has distinct roots, all of which are real.

17 $***$ Show that a rational matrix $A \in \mathbf{Mat}(n, \mathbb{Q})$ is diagonalisable over $\mathbb{Q}$ if and only if its minimal polynomial has distinct roots, all of which are rational.

18 $****$ If 2 real matrices $A, B \in \mathbf{Mat}(n, \mathbb{R})$ are similar over $\mathbb{C}$, are they necessarily similar over $\mathbb{R}$, ie can we find a matrix $P \in \mathbf{GL}(n, \mathbb{R})$ such that $B = PAP^{-1}$?

19 $****$ If 2 rational matrices $A, B \in \mathbf{Mat}(n, \mathbb{Q})$ are similar over $\mathbb{C}$, are they necessarily similar over $\mathbb{Q}$?

20 $*****$ If 2 integral matrices $A, B \in \mathbf{Mat}(n, \mathbb{Z})$ are similar over $\mathbb{C}$, are they necessarily similar over $\mathbb{Z}$, ie can we find an integral matrix $P \in \mathbf{GL}(n, \mathbb{Z})$ with integral inverse, such that $B = PAP^{-1}$?

The matrix $A \in \mathbf{Mat}(n, k)$ is said to be *semisimple* if its minimal polynomial has distinct roots. It is said to be *nilpotent* if $A^r = 0$ for some $r > 0$.

21 $***$ Show that a matrix $A \in \mathbf{Mat}(n, k)$ cannot be both semisimple and nilpotent, unless $A = 0$.

22 $***$ Show that a polynomial $p(x)$ has distinct roots if and only if

$$\gcd\left(p(x), p'(x)\right) = 1.$$

23 $****$ Show that every matrix $A \in \mathbf{Mat}(n, \mathbb{C})$ is uniquely expressible in the form

$$A = S + N,$$

where S is semisimple, N is nilpotent, and

$$SN = NS.$$

(We call S and N the semisimple and nilpotent parts of A.)

24 $****$ Show that S and N are expressible as polynomials in A.

25 ∗∗∗∗ Suppose the matrix $B \in \mathbf{Mat}(n, \mathbb{C})$ commutes with all matrices that commute with A, ie

$$AX = XA \Longrightarrow BX = XB.$$

Show that B is expressible as a polynomial in A.

Chapter 3

Simple Representations

Definition 3.1 *The representation α of G in the vector space V over k is said to be* simple *if no proper subspace of V is stable under G.*

In other words, α is simple if it has the following property: if U is a subspace of V such that

$$g \in G, u \in U \Longrightarrow gu \in U$$

then either $U = 0$ or $U = V$.

Proposition 3.1 *1. A 1-dimensional representation over k is necessarily simple.*

2. If α is a simple representation of G over k then

$$\dim \alpha \leq \|G\|.$$

Proof ► (1) is evident since a 1-dimensional space has no proper subspaces, stable or otherwise.

For (2), suppose α is a simple representation of G in V. Take any $v \neq 0$ in V, and consider the set of all *transforms* gv of V. Let U be the subspace spanned by these:

$$U = \langle gv : g \in G \rangle.$$

Each $g \in G$ permutes the transforms of v, since

$$g(hv) = (gh)v.$$

It follows that g sends U into itself. Thus U is stable under G. Since α is simple, by hypothesis,

$$V = U.$$

But since U is spanned by the $\|G\|$ transforms of v,

$$\dim V = \dim U \le \|G\|.$$

◀

Remark: This result can be greatly improved, as we shall see. If $k = \mathbb{C}$—the case of greatest interest to us—then we shall prove that

$$\dim \alpha \le \|G\|^{\frac{1}{2}}$$

for any simple representation α.

We may as well announce now the full result. Suppose G is a finite group. Then we shall show (in due course) that

1. The number of simple representations of G over $\mathbb{C}$ is equal to the number s of conjugacy classes in G;

2. The dimensions of the simple representations $\sigma_1, \dots, \sigma_s$ of G over $\mathbb{C}$ satisfy the relation
$$\dim^2 \sigma_1 + \cdots + \dim^2 \sigma_s = \|G\|.$$

3. The dimension each simple representation σ_i divides the order of the group:
$$\dim \sigma_i \mid \|G\|.$$

Of course we cannot use these results in any proof; and in fact we will not even use them in examples. But at least they provide a useful check on our work.

Examples:

1. The first stage in studying the representation theory of a group G is to determine the simple representations of G.

   Let us agree henceforth to adopt the convention that if the scalar field k is not explicitly mentioned, then we may take it that $k = \mathbb{C}$.

   We normally start our search for simple representations by listing the 1-dimensional representations. In this case we know that S_3 has just 2 1-dimensional representations, the trivial representation 1, and the parity representation ϵ.

   Now suppose that α is a simple representation of S_3 of dimension > 1. Recall that
$$S_3 = \langle s, t : s^3 = t^2 = 1, /; st = ts^2 \rangle,$$
   where $s = (abc), /; t = (ab)$.

Let e be an eigenvector of s. Thus

$$se = \lambda e,$$

where

$$s^3 = 1 \Longrightarrow \lambda^3 = 1 \Longrightarrow \lambda = 1,\ \omega,\ \text{or}\ \omega^2.$$

Let

$$f = te.$$

Then

$$sf = ste = ts^2e = \lambda^2 te = \lambda^2 f.$$

Thus f is also an eigenvector of s, but with eigenvalue λ^2.

Now consider the subspace

$$U = \langle e, f \rangle$$

spanned by e and f. Then U is stable under s and t, and so under S_3. For

$$se = \lambda e,\ sf = \lambda^2 f,\ te = f,\ tf = t^2 e = e.$$

It follows, since α is simple, that

$$V = U.$$

So we have shown, in particular, that the simple representations of S_3 can only have dimension 1 or 2.

Let us consider the 3 possible values for λ:

(a) $\lambda = \omega$. In this case the representation takes the matrix form

$$s \mapsto \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

(b) $\lambda = \omega^2$. In this case the representation takes the matrix form

$$s \mapsto \begin{pmatrix} \omega^2 & 0 \\ 0 & \omega \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

But *this is the same representation as the first,* since the coordinate swap $(x, y) \mapsto (y, x)$ takes one into the other.

(c) $\lambda = 1$. In this case

$$se = e,\ sf = f \Longrightarrow sv = v \text{ for all } v \in V.$$

In other words s acts as the identity on V. It follows that s is represented by the matrix I with respect to *any* basis of V.

(More generally, is $g \in G$ is represented by a scalar multiple ρI of the identity with respect to one basis, then it is represented by ρI with respect to every basis; because

$$P(\rho I)P^{-1} = \rho I,$$

if you like.)

So in this case we can turn to t, leaving s to 'look after itself'. Let e be an eigenvector of t. Then the 1-dimensional space

$$U = \langle e \rangle$$

is stable under S_3, since

$$se = e, /; te = \pm e.$$

Since α is simple, it follows that $V = U$, ie V is 1-dimensional, contrary to hypothesis.

We conclude that S_3 has just 3 simple representations

$$1,\ \epsilon \text{ and } \alpha,$$

of dimensions 1, 1 and 2, given by

$$\begin{array}{ll}
1: & s \mapsto 1, /; t \mapsto 1 \\
\epsilon: & s \mapsto 1, /; t \mapsto -1 \\
\alpha: & s \mapsto \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
\end{array}$$

2. Now let us determine the simple representations (over $\mathbb{C}$) of the quaternion group

$$Q_8 = \langle s, t : s^4 = 1, s^2 = t^2, st = ts^3 \rangle,$$

where $s = i, /; t = j$. (It is best to forget at this point that one of the elements of Q_8 is called -1, and another i, since otherwise we shall fall into endless confusion.)

We know that Q_8 has four 1-dimensional representations, given by

$$s \mapsto \pm 1,\ t \mapsto \pm 1.$$

Suppose α is a simple representation of Q_8 in V, of dimension > 1. Let e be an eigenvector of s:

$$se = \lambda e,$$

where

$$s^4 = 1 \Longrightarrow \lambda = \pm 1, \pm i.$$

Let

$$te = f.$$

Then

$$sf = ste = ts^3 e = \lambda^3 te = \lambda^3 f.$$

So as in the previous example, f is also an eigenvector of s, but with eigenvalue λ^3.

Again, as in that example, the subspace

$$U = \langle e, f \rangle$$

is stable under Q_8, since

$$se = \lambda e,\ sf = \lambda^3 f,\ te = f,\ tf = t^2 e = s^2 e = \lambda^2 e.$$

So $V = U$, and $\{e, f\}$ is a basis for V. With respect to this basis our representation takes the form

$$s \mapsto \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^3 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & \lambda^2 \\ 1 & 0 \end{pmatrix},$$

where $\lambda = \pm 1, \pm i$.

If $\lambda = 1$ this representation is not simple, since the 1-dimensional subspace

$$\langle (1, 1) \rangle$$

is stable under Q_8. (This is the same argument as before. Every vector is an eigenvector of s, so we can find a simultaneous eigenvector by taking any eigenvector of t.)

The same argument holds if $\lambda = -1$, since s is represented by $-I$ with respect to one basis, and so also with respect to any basis. Again, the subspace

$$\langle (1, 1) \rangle$$

is stable under Q_8, contradicting our assumption that the representation is simple, and of dimension > 1.

We are left with the cases $\lambda = \pm i$. In fact these are equivalent. For if $\lambda = -i$, then f is an s-eigenvector with eigenvalue $\lambda^3 = i$. So taking f in place of e we may assume that $\lambda = i$.

We conclude that Q_8 has just 5 simple representations, of dimensions 1,1,1,1,2, given by

$$\begin{array}{rl} 1: & s \mapsto 1, /; t \mapsto 1 \\ \mu: & s \mapsto 1, /; t \mapsto -1 \\ \nu: & s \mapsto -1, /; t \mapsto 1 \\ \rho: & s \mapsto -1, /; t \mapsto -1 \\ \alpha: & s \mapsto \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \end{array}$$

We end by considering a very important case: *abelian* (or commutative) groups.

Proposition 3.2 *A simple representation of a finite abelian group over $\mathbb{C}$ is necessarily 1-dimensional.*

Proof ► Suppose $a \in A$. Let λ be an eigenvalue of a, and let

$$E(\lambda) = \{v \in V : av = \lambda v\}.$$

be the corresponding eigenspace.

Then $E(\lambda)$ is *stable under* A. For

$$\begin{aligned} b \in A, v \in E(\lambda) &\implies a(bv) = (ab)v = (ba)v = b(av) = b(\lambda v) = \lambda(bv) \\ &\implies bv \in E(\lambda). \end{aligned}$$

Thus $E(\lambda)$ is stable under b, and so under A. But since V is simple, by hypothesis, it follows that

$$E(\lambda) = V.$$

In other words a acts as a scalar multiple of the identity:

$$a = \lambda I.$$

It follows that *every* subspace of V is stable under a. Since that is true for each $a \in A$, we conclude that every subspace of V is stable under A. Therefore, since α is simple, V has no proper subspaces. But that is only true if $\dim V = 1$. ◄

Example: Consider the group

$$D_2 = \{1, a, b, c : a^2 = b^2 = c^2 = 1,\ bc = cb = a,\ ca = ac = b,\ ab = ca = c\}.$$

This has just four 1-dimensional representations, as shown in the following table.

	1	a	b	c
1	1	1	1	1
μ	1	1	-1	-1
ν	1	-1	1	-1
ρ	1	-1	-1	1

Chapter 4

The Arithmetic of Representations

4.1 Addition

Representations can be added and multiplied, like numbers; and the usual laws of arithmetic hold. There is even a conjugacy operation, analogous to complex conjugation.

Definition 4.1 *Suppose α, β are representations of G in the vector spaces U, V over k. Then $\alpha + \beta$ is the representation of G in $U \bigoplus V$ defined by the action*

$$g(u \oplus v) = gu \oplus gv.$$

Remarks:

1. Recall that $U \bigoplus V$ is the cartesian product of U and V, where however we write $u \oplus v$ rather than (u, v). The structure of a vector space is defined on this set in the natural way.

2. Note that $\alpha + \beta$ is only defined when α, β are representations of the *same* group G over the *same* scalar field k.

3. Suppose α, β are given in matrix form

$$\alpha : g \mapsto A(g), \quad \beta : g \mapsto B(g).$$

Then $\alpha + \beta$ is the representation

$$g \mapsto \begin{pmatrix} A(g) & 0 \\ 0 & B(g) \end{pmatrix}$$

Example: Let us look again at the 2-dimensional representations $\gamma_1, \gamma_2, \gamma_3$ of S_3 over $\mathbb{C}$ defined in Chapter 2

$$\gamma_1 : s \mapsto \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix}, t \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \gamma_2 : s \mapsto \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix}, t \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\gamma_3 : s \mapsto \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix}, t \mapsto \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We see now that

$$\gamma_1 = 1 + 1,\ \gamma_2 = 1 + \epsilon,\ \gamma_3 = \epsilon + \epsilon,$$

where 1 is the trivial 1-dimensional representation of S_3, and ϵ is the 1-dimensional parity representation

$$s \mapsto 1, \quad t \mapsto -1.$$

(We can safely write $1 + 1 = 2, \epsilon + \epsilon = 2\epsilon$.)

Proposition 4.1 *1.* $\dim(\alpha + \beta) = \dim \alpha + \dim \beta$*;*

2. $\beta + \alpha = \alpha + \beta$*;*

3. $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$*.*

Proof ► These are all immediate. For example, the second part follows from the natural isomorphism

$$V \bigoplus U \to U \bigoplus V : v \oplus u \mapsto u \oplus v.$$

◄

4.2 Multiplication

Definition 4.2 *Suppose α, β are representations of G in the vector spaces U, V over k. Then $\alpha\beta$ is the representation of G in $U \bigotimes V$ defined by the action*

$$g(u_1 \otimes v_1 + \cdots + u_r \otimes v_r) = gu_1 \otimes gv_1 + \cdots gu_r \otimes gv_r.$$

Remarks:

1. The *tensor product* $U \bigotimes V$ of 2 vector spaces U and V may be unfamiliar. Each element of $U \bigotimes V$ is expressible as a finite sum

$$u_1 \otimes v_1 + \cdots + u_r \otimes v_r.$$

If U has basis $\{e_1, \ldots, e_m\}$ and V has basis $\{f_1, \ldots, f_n\}$ then the mn elements

$$u_i \otimes v_j \quad (i = 1, \ldots, m;\ j = 1, \ldots, n)$$

form a basis for $U \bigotimes V$. In particular

$$\dim(U \otimes V) = \dim U \dim V.$$

(It is a common mistake to suppose that every element of $U \bigotimes V$ is expressible in the form $u \otimes v$. That is not so; the general element requires a finite sum.)

Formally, the tensor product is defined as the set of formal sums

$$u_1 \otimes v_1 + \cdots + u_r \otimes v_r,$$

where 2 sums define the same element if one can be derived from the other by applying the rules

$$(u_1+u_2)\otimes v \otimes u_1 \otimes v + u_2 \otimes v,\ u \otimes (v_1+v_2) \otimes u \otimes v_1 + u \otimes v_2,\ (\rho u) \otimes v = u \otimes (\rho v).$$

The structure of a vector space is defined on this set in the natural way.

2. As with $\alpha + \beta$, $\alpha\beta$ is only defined when α, β are representations of the same group G over the same scalar field k.

3. It is important *not* to write $\alpha \times \beta$ for $\alpha\beta$, as we shall attach a different meaning to $\alpha \times \beta$ later.

4. Suppose α, β are given in matrix form

$$\alpha : g \mapsto A(g), \quad \beta : g \mapsto B(g).$$

Then $\alpha\beta$ is the representation

$$\alpha\beta : g \mapsto A(g) \otimes B(g).$$

But what do we mean by the tensor product $S \otimes T$ of 2 square matrices S, T? If $S = s_{ij}$ is an $m \times m$-matrix, and $T = t_{kl}$ is an $n \times n$-matrix, then

$S \otimes T$ is the $mn \times mn$-matrix whose rows and columns are indexed by the pairs (i, k) where $1 \le i \le m, 1 \le k \le n$, with matrix entries

$$(S \otimes T)_{(i,k)(j,l)} = S_{ij} T_{kl}.$$

To write out this matrix $S \otimes T$ we must order the index-pairs. Let us settle for the 'lexicographic order'

$$(1,1), (1,2), \ldots, (1,n), (2,1), \ldots, (2,n), \ldots, (m,1), \ldots, (m,n).$$

(In fact the ordering *does not matter* for our purposes. For if we choose a different ordering of the rows, then we shall have to make the same change in the ordering of the columns; and this double change simply corresponds to a change of basis in the underlying vector space, leading to a similar matrix to $S \otimes T$.)

Example: Consider the 2-dimensional representation α of S_3 over $\mathbb{C}$

$$\alpha : s \mapsto \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

We shall determine the 4-dimensional representation $\alpha^2 = \alpha\alpha$. (The notation α^2 causes no problems.) We have

$$\alpha^2 : s \mapsto \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix} \bigotimes \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bigotimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

It is simply (!) a matter of working out these 2 tensor products. In fact

$$\begin{aligned} \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix} \bigotimes \begin{pmatrix} \omega & 0 \\ 0 & \omega^2 \end{pmatrix} &= \begin{pmatrix} \omega \cdot \omega & \omega \cdot 0 & 0 \cdot \omega & 0 \cdot 0 \\ \omega \cdot 0 & \omega \cdot \omega^2 & 0 \cdot 0 & 0 \cdot \omega^2 \\ 0 \cdot \omega & 0 \cdot 0 & \omega^2 \cdot \omega & \omega^2 \cdot 0 \\ 0 \cdot 0 & 0 \cdot \omega^2 & \omega^2 \cdot 0 & \omega^2 \cdot \omega^2 \end{pmatrix} \\ &= \begin{pmatrix} \omega^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \omega \end{pmatrix}, \end{aligned}$$

while

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bigotimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

We can simplify this by the change of coordinates $(x, y, z, t) \mapsto (y, z, t, x)$. This will give the equivalent representation (which we may still denote by $\alpha\beta$):

$$\alpha\beta : s \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \omega & 0 \\ 0 & 0 & 0 & \omega^2 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

But now we see that this splits into 2 2-dimensional representations, the second of which is α itself:

$$\alpha^2 = \beta + \alpha,$$

where β is the representation

$$\beta : s \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The representation β can be split further. That is evident if we note that since s is represented by I, we can diagonalise t without affecting s. Since t has eigenvalues ± 1, this must yield the representation

$$\beta : s \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Concretely, the change of coordinates $(x, y) \mapsto (x + y, x - y)$ brings this about.) Thus

$$\beta = 1 + \epsilon,$$

and so

$$\alpha^2 = 1 + \epsilon + \alpha.$$

(We hasten to add that this kind of matrix manipulation is not an essential part of representation theory! We shall rapidly develop techniques which will enable us to dispense with matrices altogether.)

Proposition 4.2

1. $\dim(\alpha\beta) = \dim \alpha \dim \beta$;
2. $\beta\alpha = \alpha\beta$;
3. $\alpha(\beta\gamma) = (\alpha\beta)\gamma$;
4. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$;
5. $1\alpha = \alpha$.

All these results, again, are immediate consequences of 'canonical isomorphisms' which it would be tedious to explicate.

We have seen that the representations of G over k can be added and multiplied. They almost form a ring—only subtraction is missing. In fact if we introduce 'virtual representations' $\alpha - \beta$ (where α, β are representations) then we will indeed obtain a ring

$$R(G) = R(G, k),$$

the *representation-ring* if G over k. (By convention if k is omitted then we assume that $k = \mathbb{C}$.)

We shall see later that

$$\alpha + \beta = \alpha + \gamma \Longrightarrow \beta = \gamma.$$

It follows that nothing is lost in passing from representations to $R(G)$; if $\alpha = \beta$ in $R(G)$ then $\alpha = \beta$ in 'real life'.

4.3 Conjugacy

Definition 4.3 *Suppose $\alpha =$ is a representation of G in the vector space V over k. Then α^* is the representation of G in the dual vector space U^* defined by the action*

$$(g\pi)(v) = \pi(g^{-1}v) \quad (g \in G,\ \pi \in V^*,\ v \in V)$$

Remarks:

1. Recall that the dual vector space V^* is the space of linear functionals

$$\pi : V \to k.$$

   To any basis $\{e_1, \ldots, e_n\}$ of V there corresponds a dual basis $\{\pi_1, \ldots, \pi_n\}$ of V^*, where

$$\pi_j(e_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

2. Suppose α is given in matrix form

$$\alpha : g \mapsto A(g).$$

   Then α^* is the representation

$$g \mapsto \left(A(g)^{-1}\right)',$$

where T' denotes the transpose of T. Notice the mysterious way in which the inverse and transpose, each of which is 'contravariant', ie

$$(RS)^{-1} = S^{-1}T^{-1}, \quad (RS)' = S'R',$$

combine to give the required property

$$\left((RS)^{-1}\right)' = (R^{-1})'(S^{-1})'.$$

Example: Consider α^*, where α is the 2-dimensional representation of S_3 over $\mathbb{C}$ considered above. By the rule above, α^* is given by

$$\alpha^* : s \mapsto \begin{pmatrix} \omega^2 & 0 \\ 0 & \omega \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

It is easy to see that swapping the coordinates, $(x, y) \mapsto (y, x)$, gives

$$\alpha^* = \alpha$$

Many of the representations we shall meet will share this property of self-conjugacy.

Proposition 4.3 *1.* $\dim(\alpha^*) = \dim \alpha$*;*

2. $(\alpha^*)^* = \alpha$*;*

3. $(\alpha + \beta)^* = \alpha^* + \beta^*$*.*

4. $(\alpha\beta)^* = \alpha^*\beta^*$*.*

5. $1^* = 1$*.*

Summary: We have defined the representation ring $R(G)$ of a group G, and shown that it carries a conjugacy operation $\alpha \mapsto \alpha^*$.

Chapter 5

Semisimple Representations

Definition 5.1 *The represenation α of G is said to be* semisimple *if it is expressible as a sum of simple representations:*

$$\alpha = \sigma_1 + \cdots + \sigma_r.$$

Example: Consider the permutation representation ρ of S_3 in k^3. (It doesn't matter for the following argument if $k = \mathbb{R}$ or $\mathbb{C}$.)

Recall that

$$g(x_1, x_2, x_3) = (x_{g^{-1}1}, x_{g^{-1}2}, x_{g^{-1}3}).$$

We have seen that k^3 has 2 proper stable subspaces:

$$U = \{(x, x, x) : x \in k\}, \quad W = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 0\}.$$

U has dimension 1, with basis $\{(1, 1, 1)\}$; W has dimension 2, with basis $\{(1, -1, 0), (-1, 0, 1)\}$. Evidently

$$U \cap V = 0.$$

Recall that a sum $U + V$ of vector subspaces is direct,

$$U + V = U \oplus V,$$

if (and only if) $U \cap V = 0$. So it follows here, by considering dimensions, that

$$k^3 = U \bigoplus W.$$

The representation on U is the trivial representation 1. Thus

$$\rho = 1 + \alpha,$$

where α is the representation of S_3 in W.

We can see that α is simple as follows. Suppose $V \subset W$ is stable under S_3, where $V \neq 0$. Take any element $v \neq 0$ in V: say

$$v = (x, y, z) \quad (x + y + z = 0).$$

The coefficients x, y, z cannot all be equal. Suppose $x \neq y$. Then

$$(12)v = (y, x, z) \in V;$$

and so

$$v - (12)v = (x - y, y - x, 0) = (x - y)(1, -1, 0) \in V.$$

Hence

$$(1, -1, 0) \in V.$$

It follows that

$$(-1, 0, 1) = (132)(1, -1, 0) \in V$$

also. But these 2 elements generate W; hence

$$V = W.$$

So we have shown that W is a simple S_3-space, whence the corresponding representation α is simple.

We conclude that the representation

$$\rho = 1 + \alpha$$

is a sum of simple representations, and so is semisimple.

It is easy to see that U and W are the only subspaces of k^3 stable under S_3, apart from 0 and the whole space. So it is evident that the splitting $U \oplus V$ is unique. In general this is not so; in fact we shall show later that there is a unique split into simple subspaces if and only if the representations corresponding to these subspaces are distinct. (So in this case the split is unique because $1 \neq \alpha$.) However the simple representations that appear *are* unique. This fact, which we shall prove in the next chapter, is the foundation stone of representation theory.

Most of the time we do not need to look behind a representation at the underlying representation-space. But sometimes we do; and the following results should help to clarify the structure of semisimple representation-spaces.

Proposition 5.1 *Suppose V is a sum (not necessarily direct) of simple subspaces:*

$$V = S_1 + \cdots + S_r.$$

Then V is semisimple.

Proof ► Since S_2 is simple,

$$S_1 \cap S_2 = 0 \text{ or } S_2.$$

In the former case

$$S_1 + S_2 = S_1 \bigoplus S_2;$$

in the latter case $S_2 \subset S_1$ and so

$$S_1 + S_2 = S_1.$$

Repeating the argument with $S_1 + S_2$ in place of S_1, and S_3 in place of S_2,

$$(S_1 + S_2) \cap S_3 = 0 \text{ or } S_3,$$

since S_3 is simple. In the former case

$$S_1 + S_2 + S_3 = (S_1 + S_2) \bigoplus S_3;$$

in the latter case $S_3 \subset S_1 + S_2$ and so

$$S_1 + S_2 + S_3 = S_1 + S_2.$$

Combining this with the previous step

$$S_1 + S_2 + S_3 = S_1 \bigoplus S_2 \bigoplus S_3 \text{ or } S_1 \bigoplus S_3 \text{ or } S_1 \bigoplus S_2 \text{ or } S_1.$$

Continuing in this style, at the ith step, since S_i is simple,

$$S_1 + \cdots + S_i = (S_1 + \cdots + S_{i-1}) \bigoplus S_i \text{ or } S_1 + \cdots + S_{i-1}.$$

We conclude, finally, that

$$V = S_1 + \cdots + S_r = S_{i_1} \bigoplus \cdots S_{i_s},$$

where $\{S_{i_1}, \ldots, S_{i_s}\}$ is a subset of $\{S_1, \ldots, S_r\}$. ◄

Remark: The subset $\{S_{i_1}, \ldots, S_{i_s}\}$ depends in general on the *order* in which we take $S_1, \ldots, S_r$. In particular, since $S_{i_1} = S_1$, we can always specify that any *one* of $S_1, \ldots, S_n$ appears in the direct sum.

Proposition 5.2 *The following 2 properties of the G-space V are equivalent:*

1. *V is semisimple;*

2. *each stable subspace* $U \subset V$ *has at least one complementary stable subspace* W*, ie*
$$V = U \bigoplus W.$$

Proof ► Suppose first that V is semisimple, say
$$V = S_1 \bigoplus \cdots \bigoplus S_r.$$

Let us follow the proof of the preceding proposition, but starting with U rather than S_1. Thus our first step is to note that since S_1 is simple,
$$U + S_1 = U \bigoplus S_1 \text{ or } U.$$

Continuing as before, we conclude that
$$V = U \bigoplus S_{i_1} \bigoplus \cdots \bigoplus S_{i_s},$$

from which the result follows, with
$$W = S_{i_1} \bigoplus \cdots \bigoplus S_{i_s}.$$

Now suppose that condition (2) holds. Since V is finite-dimensional, we can find a stable subspace S_1 of minimal dimension. Evidently S_1 is simple; and by our hypothesis
$$V = S_1 \bigoplus W_1.$$

Now let us find a stable subspace S_2 of W_1 of minimal dimension. As before, this subspace is simple; and
$$S_1 \cap S_2 \subset S_1 \cap W = 0,$$

so that
$$S_1 + S_2 = S_1 \bigoplus S_2.$$

Applying the hypothesis again to this space, we can find a stable complement W_2:
$$V = S_1 \bigoplus S_2 \bigoplus W_2.$$

Continuing in this way, since V is finite-dimensional we must conclude with an expression for V as a direct sum of simple subspaces:
$$V = S_1 \bigoplus \cdots \bigoplus S_r.$$

Hence V is semisimple. ◄

Remark: This Proposition gives an alternative definition of semisimplicity: *V is semisimple if every stable subspace $U \subset V$ posseses a complementary stable subspace W.* This alternative definition allows us to extend the concept of semisimplicity to infinite-dimensional representations.

Exercises

In Exercises 01–15 calculate e^X for the given matrix X:

1. Show that any *commuting* set of diagonalisable matrices can be simultaneously diagonalised. Hence show that any representation of a finite abelian group

2. Show that for all n the natural representation ρ of S_n in k^n is semisimple.

3. If $T \in \mathbf{GL}(n, k)$ then the map

$$\mathbb{Z} \to \mathbf{GL}(n, k) : m \mapsto T^m$$

   defines a representation τ of the infinite abelian group Z.

   Show that if $k = \mathbb{C}$ then τ is semisimple if and only if T is semisimple.

4. Prove the same result when $k = \mathbb{R}$.

5. Suppose $k = \mathbf{GF}(2) = \{0, 1\}$, the finite field with 2 elements. Show that the representation of $C_2 = \{e, g\}$ given by

$$g \mapsto \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

   is not semisimple.

Chapter 6

Every Representation of a Finite Group is Semisimple

Theorem 6.1 *(Maschke's Theorem) Suppose α is a representation of the finite group G over k, where $k = \mathbb{R}$ or $\mathbb{C}$. Then α is semisimple.*

Proof ► Suppose α is a representation on V. We take the alternative definition of semisimplicity: every stable subspace $U \subset V$ must have a stable complement W.

Our idea is to construct an invariant positive-definite form P on V. (By 'form' we mean here *quadratic form* if $k = \mathbb{R}$, or *hermitian form* if $k = \mathbb{C}$.) Then we can take W to be the *orthogonal complement* of U with respect to this form:

$$W = U^{\perp} = \{v \in V : P(u, v) = 0 \text{ for all } u \in U\}.$$

We can construct such a form by taking *any* positive-definite form Q, and *averaging* it over the group:

$$P(u, v) = \frac{1}{\|G\|} \sum_{g \in G} Q(gu, gv).$$

(It's not really necessary to divide by the order of the group; we do it because the idea of 'averaging over the group' occurs in other contexts.)

It is easy to see that the resulting form is invariant:

$$\begin{aligned} P(gu, gv) &= \frac{1}{\|G\|} \sum_{h \in G} Q(hgu, hgv) \\ &= \frac{1}{\|G\|} \sum_{h \in G} Q(hu, hv) \\ &= P(u, v) \end{aligned}$$

since hg runs over the group as h does so.

It is a straightforward matter to verify that if P is invariant and U is stable then so is $U^\perp$. Writing $\langle u, v\rangle$ for $P(u, v)$,

$$\begin{aligned} g \in G, w \in U^\perp &\implies \langle u, w\rangle = 0\ \forall u \in U \\ &\implies \langle gu, gw\rangle = \langle u, w\rangle = 0\ \forall u \in U \\ &\implies \langle u, gw\rangle = \langle g(g^{-1}u), w\rangle = 0\ \forall u \in U \\ &\implies gw \in U^\perp. \end{aligned}$$

◀

Examples:

1. Consider the representation of S_3 in $\mathbb{R}^3$. There is an obvious invariant quadratic form—as is often the case—namely

$$x_1^2 + x_2^2 + x_3^2.$$

But as an exercise in averaging, let us take the positive-definite form

$$Q(x_1, x_2, x_3) = 2x_1^2 - 2x_1x_2 + 3x_2^2 + x_3^2.$$

Then

$$\begin{aligned} Q\left(e(x_1, x_2, x_3)\right) &= Q(x_1, x_2, x_3) = 2x_1^2 - 2x_1x_2 + 3x_2^2 + x_3^2 \\ Q\left((23)(x_1, x_2, x_3)\right) &= Q(x_1, x_3, x_2) = 2x_1^2 - 2x_1x_3 + 3x_3^2 + x_2^2 \\ Q\left((13)(x_1, x_2, x_3)\right) &= Q(x_3, x_2, x_1) = 2x_3^2 - 2x_3x_2 + 3x_2^2 + x_1^2 \\ Q\left((12)(x_1, x_2, x_3)\right) &= Q(x_2, x_1, x_3) = 2x_2^2 - 2x_2x_1 + 3x_1^2 + x_3^2 \\ Q\left((123)(x_1, x_2, x_3)\right) &= Q(x_3, x_1, x_2) = 2x_3^2 - 2x_3x_1 + 3x_1^2 + x_2^2 \\ Q\left((132)(x_1, x_2, x_3)\right) &= Q(x_2, x_3, x_1) = 2x_2^2 - 2x_2x_3 + 3x_3^2 + x_1^2 \end{aligned}$$

Adding, and dividing by 6,

$$\begin{aligned} P(x_1, x_2, x_3) &= 2\left(x_1^2 + x_2^2 + x_3^2\right) - \frac{2}{3}\left(x_2x_3 + x_1x_3 + x_1x_2\right) \\ &= \frac{7}{3}\left(x_1^2 + x_2^2 + x_3^2\right) - \frac{1}{3}\left(x_1 + x_2 + x_3\right)^2. \end{aligned}$$

The corresponding inner product is given by

$$\langle (x_1, x_2, x_3), (y_1, y_2, y_3)\rangle = 2(x_1y_1 + x_2y_2 + x_3y_3) - \frac{1}{3}(x_2y_3 + x_3y_2 + x_3y_1 + x_1y_3 + x_1y_2 + x_2y_1)$$

To see how this is used, let

$$U = \{(x, x, x) : x \in \mathbb{R}\}.$$

Evidently U is stable. Its orthogonal complement with respect to the form above is

$$\begin{aligned} U^\perp &= \{(x_1, x_2, x_3) : \langle(1, 1, 1), (x_1, x_2, x_3)\rangle = 0\} \\ &= \{(x_1, x_2, x_3) : \frac{4}{3}(x_1 + x_2 + x_3) = 0\} \\ &= \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 0\}, \end{aligned}$$

which is just the complement we found before. This is not surprising since—as we observed earlier—U and $U^\perp$ are the only proper stable subspaces of $\mathbb{R}^3$.

2. For an example using hermitian forms, consider the simple representation of D_4 over $\mathbb{C}$ defined by

$$s \mapsto \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Again, there is an obvious invariant hermitian form, namely

$$|x_1^2 + x_2^2| = \overline{x_1}x_1 + \overline{x_2}x_2.$$

But this will not give us much exercise.

The general hermitian form on $\mathbb{C}^2$ is

$$a\bar{x}x + b\bar{y}y + c\bar{x}y + \bar{c}\bar{y}x \quad (a, b \in \mathbb{R},\ c \in \mathbb{C})$$

Let us take

$$Q(x, y) = 2\bar{x}x + \bar{y}y - i\bar{x}y + i\bar{y}x.$$

Note that

$$D_4 = \{e, s, s^2, s^3, t, ts, ts^2, ts^3\}.$$

For these 8 elements are certainly distinct, eg

$$s^2 = ts^3 \Longrightarrow ts = 1 \Longrightarrow s = t.$$

Now

$$\begin{aligned}
Q\left(e(x,y)\right) &= Q(x,y) = 2\bar{x}x + \bar{y}y - i\bar{x}y + i\bar{y}x,\\
Q\left(s(x,y)\right) &= Q(ix,-iy) = 2\bar{x}x + \bar{y}y + i\bar{x}y - i\bar{y}x,\\
Q\left(s^2(x,y)\right) &= Q(-x,-y) = 2\bar{x}x + \bar{y}y - i\bar{x}y + i\bar{y}x,\\
Q\left(s^3(x,y)\right) &= Q(-ix,iy) = 2\bar{x}x + \bar{y}y + i\bar{x}y - i\bar{y}x,\\
Q\left(t(x,y)\right) &= Q(y,x) = \bar{x}x + 2\bar{y}y + i\bar{x}y - i\bar{y}x,\\
Q\left(ts(x,y)\right) &= Q(iy,-ix) = \bar{x}x + 2\bar{y}y - i\bar{x}y + i\bar{y}x,\\
Q\left(ts^2(x,y)\right) &= Q(-y,-x) = \bar{x}x + 2\bar{y}y + i\bar{x}y - i\bar{y}x,\\
Q\left(ts^3(x,y)\right) &= Q(-iy,ix) = \bar{x}x + 2\bar{y}y - i\bar{x}y + i\bar{y}x.
\end{aligned}$$

Averaging,

$$\begin{aligned}
P(x,y) &= \frac{1}{8}\sum g \in D_4 Q(g(x,y))\\
&= \frac{3}{2}(\bar{x}x + \bar{y}y)
\end{aligned}$$

It is no coincidence that we have ended up with a scalar multiple of $|x|^2 + |y|^2$. For it is easy to see that a *simple* G-space carries a *unique* invariant hermitian form, up to a scalar multiple. Suppose P, Q were 2 such forms. Let λ be an eigenvalue of Q with respect to P, ie a solution of

$$\det(A - \lambda B) = 0,$$

where A, B are the matrices of P, Q. Then the corresponding eigenspace

$$E = \{v : Av = \lambda Bv\}$$

would be stable under G.

The alternative proof of Maschke's Theory below may be preferred by the algebraically-minded. It has the advantage of extending to scalar fields other than $\mathbb{R}$ and $\mathbb{C}$. Against that, it lacks the intuitive appeal of the earlier proof.

Alternative proof ► Recall that a projection $p : V \to V$ is a linear map satisfying the relation

$$p^2 = p$$

(ie p is *idempotent*).

If p is a projection then so is $1 - p$:

$$(1-p)^2 = 1 - 2p + p^2 = 1 - 2p + p = 1 - p.$$

The projections $(p, 1-p)$ define a splitting of V into a direct sum

$$V = \operatorname{im} p \bigoplus \operatorname{im}(1-p).$$

Note that

$$v \in \operatorname{im} p \iff pv = v.$$

Note also that

$$\operatorname{im}(1-p) = \ker p,$$

since

$$v = (1-p)w \implies pv = (p - p^2)w = 0,$$

while

$$pv = 0 \implies v = (1-p)v.$$

Thus the splitting can equally well be written

$$V = \operatorname{im} p \bigoplus \ker p.$$

Conversely, every splitting

$$V = U \bigoplus W$$

arises from a projection p in this way: if

$$v = u + w \quad (u \in U, w \in W)$$

then we set

$$pv = u.$$

(Although the projection p is often referred to as 'the projection onto U' it depends on W as well as U. In general there are an infinity of projections onto U, corresponding to the infinity of complements W. When there is a positive-definite form on V—quadratic or hermitian, according as $k = \mathbb{R}$ or $\mathbb{C}$—then one of these projections is distinguished: namely the 'orthogonal projection' corresponding to the splitting

$$V = U \bigoplus U^{\perp}.$$

But we are not assuming the existence of such a form at the moment.)

Now suppose U is a stable subspace of V. Choose any complementary subspace W:

$$V = U \bigoplus W.$$

In general W will not be stable under G. Our task is to find a stable complementary subspace W_0:

$$V = U \bigoplus W = U \bigoplus W_0.$$

Let p be the projection onto U with complement W. We know that U is stable under G, ie

$$g \in G,\ u \in U \Longrightarrow gu \in U.$$

Thus

$$g \in G, u = pv \Longrightarrow pgu = gu \Longrightarrow pgpv = gpv.$$

Since this holds for all $v \in V$,

$$pgp = gp$$

for all $g \in G$. Conversely, if this is so then $U = \operatorname{im} p$ is stable.

By the same argument, $W = \operatorname{im}(1-p)$ will be stable if and only if

$$(1-p)g(1-p) = g(1-p)$$

for all $g \in G$. This reduces to

$$pgp = pg.$$

Both U and W are stable if and only if

$$gp = pg.$$

For in that case

$$pgp = p(gp) = p(pg) = p^2 g = pg = gp.$$

Now

$$gp = pg \Longleftrightarrow g^{-1}pg = p.$$

In other words, p defines a splitting into stable subspaces if and only if it is invariant under G.

In general, we can construct an invariant element by averaging over G. Let us therefore set

$$P = \frac{1}{\|G\|} \sum_{g \in G} g^{-1}pg.$$

This will certainly be invariant under G:

$$\begin{aligned} g^{-1}Pg &= \frac{1}{\|G\|} \sum_{h \in G} g^{-1}h^{-1}phg \\ &= \frac{1}{\|G\|} \sum_{h \in G} (hg)^{-1}p(hg)^{-1} \\ &= \frac{1}{\|G\|} \sum_{h \in G} h^{-1}ph^{-1} \\ &= P, \end{aligned}$$

since hg runs over G as h does so.

What is less obvious is that P *is a projection*, and in fact a projection onto U. To see that, note that

$$u \in U \Longrightarrow gu \in U \Longrightarrow p(gu) = gu.$$

Hence by addition

$$u \in U \Longrightarrow Pu = u.$$

Conversely,

$$v \in V \Longrightarrow pgv \in U \Longrightarrow gpgv \in U.$$

So by addition

$$v \in V \Longrightarrow Pv \in U.$$

These 2 results imply that $P^2 = P$, and that P projects onto U. ◂

Remarks:

1. We can show directly that P is a projection, as follows:

$$\begin{aligned} P^2 &= \frac{1}{\|G\|^2} \sum_{g,h} g^{-1}pgh^{-1}ph \\ &= \frac{1}{\|G\|^2} \sum_{g,h} g^{-1}gh^{-1}ph \\ &= \frac{1}{\|G\|^2} \sum_{g,h} h^{-1}ph \\ &= \frac{1}{\|G\|} \sum_{h} h^{-1}ph \\ &= P. \end{aligned}$$

Two projections p, q project onto the same (first) subspace if

$$qp = p, \; pq = q.$$

So to prove that P projects onto the same subspace U as p, we must show that $Pp = p$ and $pP = P$. These follow in much the same way:

$$\begin{aligned} Pp &= \frac{1}{\|G\|} \sum_{g} g^{-1}pgp \\ &= \frac{1}{\|G\|} \sum_{g} g^{-1}gp \\ &= p, \end{aligned}$$

$$
\begin{aligned}
pP &= \frac{1}{\|G\|}\sum_g pg^{-1}pg \\
&= \frac{1}{\|G\|}\sum_g g^{-1}pg \\
&= P.
\end{aligned}
$$

2. Both proofs of Maschke's Theorem rely on the same idea: obtaining an invariant element (in the first proof, an invariant form; in the second, and invariant projection) by averaging over transforms of a non-invariant element.

   In general, if V is a G-space (in other words, we have a representation of G in V) then the invariant elements form a subspace

   $$V^G = \{v \in V : gv = v \forall g \in G\}.$$

   The averaging operation defines a projection of V onto V^G:

   $$v \mapsto \frac{1}{\|G\|}\sum_g gv.$$

   Clearly V^G is a stable subspace of V. Thus if V is simple, either $V^G = 0$ or $V^G = V$. In the first case, all averages vanish. In the second case, the representation in V is trivial, and so V must be 1-dimensional.

3. It is worth noting that our alternative proof works in any scalar field k, *provided* $\|G\| \neq 0$ in k. Thus it even works over the finite field $\mathbf{GF}(p^n)$, unless $p \mid \|G\|$.

   Of course we are not considering such *modular representations* (as representations over finite fields are known); but our argument shows that semisimplicity still holds unless the characteristic p if the scalar field divides the order of the group.

Chapter 7

Uniqueness and the Intertwining Number

Definition 7.1 *Suppose α, β are representations of G over k in the vector spaces U, V respectively. The* intertwining number $I(\alpha, \beta)$ *is defined to be the dimension of the space of G-maps $t : U \to V$,*

$$I(\alpha, \beta) = \dim \hom^G(U, V).$$

Remarks:

1. A G-map $t : U \to V$ is a linear map which preserves the action of G:

   $$t(gu) = g(tu) \quad (g \in G, u \in G).$$

   These G-maps evidently form a vector space over k.

2. The intertwining number will remain somewhat abstract until we give a formula for it (in terms of characters) in Chapter . But intuitively $I(\alpha, \beta)$ measures how much the representations α, β have in common.

3. The intertwining number of finite-dimensional representations is certainly finite, as the following result shows.

Proposition 7.1 *We have*

$$I(\alpha, \beta) \leq \dim \alpha \dim \beta.$$

Proof ► The space $\hom(U, V)$ of *all* linear maps $t : U \to V$ has dimension $\dim U \dim V$, since we can represent each such map by an $m \times n$-matrix, where $m = \dim U, n = \dim V$.

The result follows, since

$$\hom^G(U, V) \subset \hom(U, V).$$

◄

Proposition 7.2 *Suppose α, β are* simple *representations over k. Then*

$$I(\alpha, \beta) = \begin{cases} 0 & \text{if } \alpha \neq \beta, \\ \geq 1 & \text{if } \alpha = \beta. \end{cases}$$

Proof ► Suppose α, β are representations in U, V, respectively; and suppose

$$t : U \to V$$

is a G-map. Then the subspaces

$$\ker t = \{u \in U : tu = 0\} \text{ and } \operatorname{im} t = \{v \in V : \exists u \in U,\ tu = v\}$$

are both stable under G. Thus

$$\begin{aligned} u \in \ker t &\implies tu = 0 \\ &\implies t(gu) = g(tu) = 0 \\ &\implies gu \in \ker t, \end{aligned}$$

while

$$\begin{aligned} v \in \operatorname{im} t &\implies v = tu \\ &\implies t(gu) = g(tu) = gv \\ &\implies gv \in \operatorname{im} t. \end{aligned}$$

But since U and V are both simple, by hypothesis, it follows that

$$\ker t = 0 \text{ or } U, \quad \operatorname{im} t = 0 \text{ or } V.$$

Now $\ker t = U \implies t = 0$, and $\operatorname{im} t = 0 \implies t = 0$. So if $t \neq 0$,

$$\ker t = 0, \quad \operatorname{im} t = V.$$

But in this case t is an *isomorphism* of G-spaces, and so $\alpha = \beta$.

On the other hand, if $\alpha = \beta$ then (by the definition of equivalent representations) there exists a G-isomorphis $t : U \to V$, and so $I(\alpha, \beta) \geq 1$. ◄

When $k = \mathbb{C}$ we can be more precise.

Proposition 7.3 *If α is a simple representation over $\mathbb{C}$ then*

$$I(\alpha, \alpha) = 1.$$

Proof ► Suppose V carries the representation α. We have to show that

$$\dim \hom^G(V, V) = 1.$$

Since the identity map $1 : V \to V$ is certainly a G-map, we have to show that every G-map $t : V \to V$is a scalar multiple $\rho 1$ of the identity.

Let λ be an eigenvector of t. Then the corresponding eigenspace

$$E = E(\lambda) = \{v \in V : tv = \lambda v\}$$

is stable under G. For

$$g \in G,\ v \in E \Longrightarrow t(gv) = g(tv) = \lambda gv \Longrightarrow gv \in E.$$

Since α is simple, this implies that $E = V$, ie

$$t = \lambda 1.$$

◄

Proposition 7.4 *Suppose α, β, γ are representations over k. Then*

1. $I(\alpha + \beta, \gamma) = I(\alpha, \gamma) + I(\beta, \gamma)$;
2. $I(\alpha, \beta + \gamma) = I(\alpha, \beta) + I(\alpha, \gamma)$;
3. $I(\alpha\beta, \gamma) = I(\alpha, \beta^*\gamma)$.

Proof ► Suppose α, β, γ are representations in U, V, W respectively. The first 2 results are immediate, arising from the more-or-less self-evident isomorphisms

$$\begin{aligned} \hom(U \bigoplus V, W) &\cong \hom(U, W) \bigoplus \hom(V, W) \\ \hom(U, V \bigoplus W) &\cong \hom(U, V) \bigoplus \hom(U, W). \end{aligned}$$

Take the first. This expresses the fact that a linear map

$$t : U \bigoplus V \to W$$

can be defined by giving 2 linear maps

$$t_1 : U \to W,\ t_2 : V \to W.$$

In fact t_1 is the restriction of t to $U \subset U \bigoplus V$, and t_2 the restriction of t to $V \subset U \bigoplus V$; and
$$t(u \oplus v) = t_1 u \oplus t_2 v.$$
In much the same way, the second result expresses the fact that a linear map
$$t : U \to V \bigoplus W$$
can be defined by giving 2 linear maps
$$t_1 : U \to V,\ t_2 : U \to W.$$
In fact
$$t_1 = \pi_1 t, \quad t_2 = \pi_2 t,$$
where π_1, π_2 are the projections of $U \bigoplus V$ onto V, W respectively; and
$$tu = t_1 u \oplus t_2 u.$$

The third result, although following from a similar 'natural equivalence'
$$\hom(U \bigotimes V, W) \cong \hom(U, V^* \bigotimes W),$$
where
$$V^* = \hom(V, k),$$
is rather more difficult to establish.

We can divide the task in two. First, there is a natural equivalence
$$\hom(U, \hom(V, W)) \cong \hom(U \bigotimes V, W).$$
For this, note that there is a 1–1 correspondence between *linear* maps $b : \mathbf{U} \otimes V \to W$ and *bilinear* maps
$$B : U \times V \to W.$$
(This is sometimes taken as the definition of $U \otimes V$.) So we have to show how such a bilinear map $B(u, v)$ gives rise to a linear map
$$t : U \to \hom(V, W).$$
But that is evident:
$$t(u)(v) = B(u, v).$$
It is a straightforward matter to verify that every such linear map t arises in this way from a unique bilinear map B.

It remains to show that

$$\hom(V, W) \cong V^* \bigotimes W.$$

For this, note first that both sides are 'additive functors' in W, ie

$$\begin{aligned} \hom(V, W_1 \bigoplus W_2) &= \hom(V, W_1) \bigoplus \hom(V, W_2), \\ V^* \bigotimes (W_1 \bigoplus W_2) &= (V^* \bigotimes W_1) \bigoplus (V^* \bigotimes W_2). \end{aligned}$$

This allows us to reduce the problem, by expressing W as a sum of 1-dimensional subspaces, to the case where W is 1-dimensional. In that case, we may take $W = k$; so the result to be proved is

$$\hom(V, k) \cong V^* \bigotimes k.$$

But there is a natural isomorphism

$$U \bigotimes k \cong U$$

for every vector space U. So our result reduces to the tautology $V^* \cong V^*$.

It's a straightforward (if tedious) matter to verify that these isomorphisms are all compatible with the actions of the group G. In particular the G-invariant elements on each side correspond:

$$\begin{aligned} \hom^G(U \bigoplus V, W) &\cong \hom^G(U, W) \bigoplus \hom^G(V, W), \\ \hom^G(U, V \bigoplus W) &\cong \hom^G(U, V) \bigoplus \hom^G(U, W), \\ \hom^G(U \bigotimes V, W) &\cong \hom^G(U, V^* \bigotimes W). \end{aligned}$$

The 3 results follow on taking the dimensions of each side. ◄

Theorem 7.1 *The expression for a semisimple representation α as a sum of simple parts*

$$\alpha = \sigma_1 + \cdots + \sigma_r$$

is unique up to order.

Proof ► Suppose σ is a simple representation of G over k. We can use the intertwining number to compute the number of times, m say, that σ occurs amongst the σ_i. For

$$\begin{aligned} I(\sigma, \alpha) &= I(\sigma, \sigma_1) + \cdots + I(\sigma, \sigma_r) \\ &= mI(\sigma, \sigma), \end{aligned}$$

since only those summands for which $\sigma_i = \sigma$ will contribute to the sum. Thus

$$m = \frac{I(\sigma,\alpha)}{I(\sigma,\sigma)}.$$

It follows that σ will occur the same number m times in every expression for α as a sum of simple parts. Hence two such expressions can only differ in the order of their summands. ◂

Although the expression

$$\alpha = \sigma_1 + \cdots + \sigma_r$$

for the representation α is unique, the corresponding splitting

$$V = U_1 \bigoplus \cdots \bigoplus \mathbf{U}_r$$

of the representation-space is not in general unique. It's perfectly possible for 2 different expressions for V as a direct sum of simple G-subspaces to give rise to the *same* expression for α: say

$$V = U_1 \bigoplus \cdots \bigoplus U_r, \quad V = W_1 \bigoplus \cdots \bigoplus W_r$$

where U_i and W_i both carry the representation σ_i.

For example, consider the trivial representation $\alpha = 1 + 1$ of a group G in the 2-dimensional space $V = k^2$. Every subspace of V is stable under G; so if we choose *any* 2 different 1-dimensional subspaces $U, W \subset V$, we will have

$$V = U \bigoplus W.$$

However, the splitting of V into *isotypic components* is unique, as we shall see.

Definition 7.2 *The representation α, and the underlying representation-space V, are said to be* isotypic *of type σ, where σ is a simple representation, if*

$$\alpha = e\sigma = \sigma + \cdots + \sigma.$$

In other words, σ is the only simple representation appearing in α.

Proposition 7.5 *Suppose V is a G-space.*

1. *If V is isotypic of type σ then so is every G-subspace $U \subset V$.*

2. *If $U, W \subset V$ are isotypic of type σ then so is $U + W$.*

Proof ► These results follow easily from the Uniqueness Theorem. But it is useful to give an independent proof, since we can use this to construct an alternative proof of the Uniqueness Theorem.

Lemma 7.1 *Suppose*

$$V = U_1 + \cdots + U_r$$

is an expression for the G-space V as a sum of simple spaces; and suppose the subspace $U \subset V$ is also simple. Then U is isomorphic (as a G-space) to one of the summands:

$$U \cong U_i$$

for some i.

Proof of Lemma ▷ We know that

$$V = U_{i_1} \bigoplus \cdots \bigoplus U_{i_t}$$

for some subset $\{U_{i_1}, \ldots, U_{i_t}\} \subset \{U_1, \ldots, U_r\}$. Thus we may assume that the sum is direct:

$$V = U_1 \bigoplus \cdots \bigoplus U_r.$$

For each i, consider the composition

$$U \to V \to U_i,$$

where the second map is the projection of V onto its component U_i. Since U and U_i are both simple, this map is either an isomorphism, or else 0.

But it cannot be 0 for all i. For suppose $u \in U, u \neq 0$. We can express u as a sum

$$u = u_1 \oplus \cdots \oplus u_r \quad (u_i \in U_i).$$

Not all the u_i vanish. Now $u \mapsto u_i$ under the composition $U \to V \to U_i$. Thus one (at least) of these compositions is $\neq 0$. Hence $U \cong U_i$ for some i. ◁

Turning to the first part of the Proposition, if $U \subset V$, where V is σ-isotypic, then each simple summand of U must be of type σ, by the Lemma. It follows that U is also σ-isotypic.

For the second part, if U and W are both σ-isotypic, then $U + W$ is a sum (not necessarily direct) of simple subspaces X_i of type σ:

$$U + W = X_1 + \cdots + X_r.$$

But then

$$U + W = X_{i_1} \bigoplus \cdots \bigoplus X_{i_t},$$

where $\{X_{i_1}, \ldots, X_{i_t}\}$ are some of the $X_1, \ldots, X_r$. In particular $U + W$ is σ-isotypic. ◄

Corollary 7.1 *Suppose σ is a simple representation of G over k, Then each G-space V over k possesses a maximal σ-isotypic subspace V_σ, which contains every other σ-isotypic subspace.*

Definition 7.3 *This subspace V_σ is called the σ-component of V.*

Proposition 7.6 *Every semimsimple G-space V is the direct sum of its isotypic components:*

$$V = V_{\sigma_1} \bigoplus \cdots \bigoplus V_{\sigma_r}.$$

Proof ► If we take an expression for V as a direct sum of simple subspaces, and combine those that are isomorphic, we will obtain an expression for V as a direct sum of isotypic spaces of different types, each of which will be contained in the corresponding isotypic component. It follows that

$$V = V_{\sigma_1} + \cdots + V_{\sigma_r}.$$

We have to show that this sum is direct.

It is sufficient to show that

$$(V_{\sigma_1} + \cdots + V_{\sigma_{i-1}}) \bigcap V_{\sigma_i} = 0$$

for $i = 2, \ldots, r$.

Suppose not. Then we can find a simple subspace

$$U \subset V_{\sigma_i},\ U \subset V_{\sigma_1} + \cdots + V_{\sigma_{i-1}}.$$

By the Lemma to the last Proposition, U must be of type σ_i, as a subspace of V_{σ_i}. On the other hand, as a subspace of $V_{\sigma_1} + \cdots + V_{\sigma_{i-1}}$ it must be of one of the types $\sigma_1, \ldots, \sigma_{i-1}$, by the same Lemma.

This is a contradiction. Hence the sum is direct:

$$V = V_{\sigma_1} \bigoplus \cdots \bigoplus V_{\sigma_r}.$$

◄

Corollary 7.2 *If the G-space V carries a multiple-free representation*

$$\alpha = \sigma_1 + \cdots + \sigma_r$$

(where the σ_i are distinct) then V has a unique expression as a direct sum of simple subspaces.

Remark: It is easy to see that multiplicity *does* give rise to non-uniqueness. For suppose
$$V = U \bigoplus U,$$
where U is simple. For each $\lambda \in k$ consider the map

$$u \mapsto u \oplus \lambda u : U \to U \bigoplus U = V.$$

The image of this map is a subspace

$$U(\lambda) = \{u \oplus \lambda u : u \in U\}.$$

This subspace is isomorphic to U, since U is simple.

It is readily verified that

$$U(\lambda) \neq U(\mu) \Longleftrightarrow \lambda = \mu.$$

It follows that
$$V = U(\lambda) \bigoplus U(\mu)$$
for any λ, μ with $\lambda \neq \mu$.

Chapter 8

The Character of a Representation

Amazingly, all the information about a representation of a group G can be encoded in a single function on G, the *character* of the representation.

Definition 8.1 *Suppose α is a representation of G over k. The* character $\chi = \chi_\alpha$ *of α is the function $\chi : G \to k$ defined by*

$$\chi(g) = \operatorname{tr}(\alpha(g)).$$

Remarks:

1. Recall that the *trace* of an $n \times n$-matrix A is the sum of the diagonal elements:
$$\operatorname{tr} A = \sum_{1 \le i \le n} A_{ii}.$$
The trace has the following properties:
   (a) $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$;
   (b) $\operatorname{tr}(\lambda A) = \lambda \operatorname{tr} A$.
   (c) $\operatorname{tr} AB = \operatorname{tr} BA$;
   (d) $\operatorname{tr} A' = \operatorname{tr} A$;
   (e) $\operatorname{tr} A^* = \overline{\operatorname{tr} A}$.

   Here A' denotes the transpose of A, and A^* the conjugate transpose:
$$A'_{ij} = A_{ji}, \quad A^*_{ij} = \overline{A_{ji}}.$$
The third property is the only one that is not immediate:
$$\operatorname{tr} AB = \sum_i (AB)_{ii} = \sum_i \sum_j A_{ij} B_{ji} = \sum_j \sum_i B_{ji} A_{ij} = \operatorname{tr} BA.$$

Note that

$$\operatorname{tr} ABC \neq \operatorname{tr} BAC$$

in general. However the trace *is* invariant under *cyclic* permutations, eg

$$\operatorname{tr} ABC = \operatorname{tr} BCA = \operatorname{tr} CAB.$$

In particular, if P is invertible (non-singular) then

$$\operatorname{tr} PAP^{-1} = \operatorname{tr} P^{-1}PA = \operatorname{tr} A :$$

similar matrices have the same trace.

It follows from this that we can speak without ambiguity of the trace $\operatorname{tr} t$ of a linear transformation $t : V \to V$; for the matrix T representing t with respect to one basis will be changed to PTP^{-1} with respect to another basis, where P is the matrix of the change of basis.

Example: Consider the 2-dimensional representation α of D_4 over $\mathbb{C}$ given by

$$s \mapsto \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad t \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Writing χ for χ_α

$$\begin{aligned}
\chi(e) &= \dim \alpha = 2 \\
\chi(s) &= i - i = 0 \\
\chi(s^2) &= \operatorname{tr} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -1 - 1 = -2 \\
\chi(s^3) &= \operatorname{tr} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = -i + i = 0 \\
\chi(t) &= i - i = 0 \\
\chi(st) &= \operatorname{tr} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = 0 \\
\chi(s^2t) &= \operatorname{tr} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = 0 \\
\chi(s^3t) &= \operatorname{tr} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = 0
\end{aligned}$$

In summary

$$\chi(e) = 2, \chi(s^2) = -2, \chi(g) = 0 \text{ if } g \neq e, s^2.$$

Proposition 8.1 *1.* $\chi_{\alpha+\beta}(g) = \chi_\alpha(g) + \chi_\beta(g)$

2. $\chi_{\alpha\beta}(g) = \chi_\alpha(g)\chi_\beta(g)$

3. $\chi_{\alpha^*}(g) = \chi_\alpha(g^{-1})$

4. $\chi_1(g) = 1$

5. $\chi_\alpha(e) = \dim \alpha$

Proof ► (1) follows from the matrix form

$$g \mapsto \begin{pmatrix} A(g) & 0 \\ 0 & B(g) \end{pmatrix}$$

for $\alpha + \beta$.

(2) follows from the fact that if A is an $m \times m$-matrix and B is an $n \times n$-matrix then the diagonal elements of the tensor product $A \otimes B$ are just the products

$$A_{ii}B_{jj} \quad (1 \le i \le m,\ 1 \le j \le n)$$

Thus

$$\mathrm{tr}(A \otimes B) = \mathrm{tr}\, A \,\mathrm{tr}\, B.$$

(3) If α takes the matrix form

$$g \mapsto A(g)$$

then its dual is given (with respect to the dual basis) by

$$g \mapsto A(g)'^{-1} = A(g^{-1})'.$$

Hence

$$\chi_{\alpha^*}(g) = \mathrm{tr}\, A(g^{-1})' = \mathrm{tr}\, A(g^{-1}) = \chi_\alpha(g^{-1}).$$

(4) and (5) are immediate. ◄

Remark: In effect the character defines a *ring-homomorphism*

$$\chi : R(G, k) \to C(G, k)$$

from the representation-ring $R(G) = R(G, k)$ to the ring $C(G, k)$ of functions on G (with values in k).

Theorem 8.1 *Suppose α, β are representations of G over k. Then*

$$I(\alpha, \beta) = \frac{1}{\|G\|} \sum_{g \in G} \chi_\alpha(g^{-1})\chi_\beta(g).$$

Proof ► It is sufficient to prove the result when $\alpha = 1$. For on the left-hand side

$$I(\alpha, \beta) = I(1, \alpha^*\beta);$$

while on the right-hand side

$$\begin{aligned}
\sum_{g\in G} \chi_\alpha(g^{-1})\chi_\beta(g) &= \sum_g \chi_{\alpha^*}(g)\chi_\beta(g) \\
&= \sum_g \chi\alpha^*\beta(g) \\
&= \sum_g \chi\chi_1(g)\alpha^*\beta(g).
\end{aligned}$$

Thus the result for α, β follows from that for $1, \alpha^*\beta$.

We have to show therefore that

$$I(1, \alpha) = \frac{1}{\|G\|} \sum_{g\in G} \chi_\alpha(g).$$

By definition, if α is a representation in V,

$$I(1, \alpha) = \dim \hom^G(k, V).$$

Now

$$\hom(k, V) = V,$$

with the vector $v \in V$ corresponding to the map

$$\lambda \mapsto \lambda v : \ k \to V.$$

Moreover, the action of G is preserved under this identification; so we may write

$$\hom^G(k, V) = V^G,$$

where V^G denotes the space of G-invariant elements of V:

$$V^G = \{v \in V : gv = v \ \forall g \in G\}$$

Thus we have to prove that

$$\dim V^G = \frac{1}{\|G\|} \sum_{g\in G} \chi_\alpha(g).$$

Consider the 'averaging map' $\pi : V \to V$ defined by

$$v \mapsto \frac{1}{\|G\|} \sum_{g\in G} gv,$$

that is,

$$\pi = \frac{1}{\|G\|} \sum_{g \in G} \alpha(g).$$

It is evident that $\pi v \in V^G$ for all $v \in V$, ie πv is invariant under G. For

$$\begin{aligned} g\pi v &= \frac{1}{\|G\|} \sum_{h \in G} ghv \\ &= \frac{1}{\|G\|} \sum_{h \in G} hv \\ &= \pi v, \end{aligned}$$

since gh runs over G as h does.

On the other hand, if $v \in V^G$ then $gv = v$ for all g and so

$$\pi v = \frac{1}{\|G\|} \sum_{g \in G} gv = v.$$

It follws that π *is a projection onto* V^G.

Lemma 8.1 *Suppose $p : V \to V$ is a projection onto the subspace $U \subset V$. Then*

$$\operatorname{tr} p = \dim U.$$

Proof of Lemma $\triangleright$ We know that

$$V = \operatorname{im} p \oplus \ker p.$$

Let $e_1, \ldots, e_m$ be a basis for $imp = U$, and let $e_{m+1}, \ldots, e_n$ be a basis for $\ker p$. Then

$$pe_i = \begin{cases} e_i & 1 \leq i \leq m, \\ 0 & m+1 \leq ilen. \end{cases}$$

It follows that the matrix of p with respect to the basis $e_1, \ldots, e_n$ is

$$P = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

with m 1's down the diagonal and 0's elsewhere. Hence

$$\operatorname{tr} p = \operatorname{tr} P = m = \dim U.$$

$\triangleleft$

Applying this to the averaging map π,

$$\operatorname{tr} \pi = \dim V^G.$$

On the other hand, by the linearity of the trace,

$$\begin{aligned} \operatorname{tr} \pi &= \frac{1}{\|G\|} \sum_g \operatorname{tr} \alpha(g) \\ &= \frac{1}{\|G\|} \sum_g \chi_\alpha(g) \end{aligned}$$

Thus

$$\dim V^G = \frac{1}{\|G\|} \sum_g \chi_\alpha(g),$$

as we had to show. ◄

Proposition 8.2 *If* $k = \mathbb{R}$,

$$\chi_{\alpha^*}(g) = \chi_\alpha(g^{-1}) = \chi_\alpha(g).$$

If $k = \mathbb{C}$,

$$\chi_{\alpha^*}(g) = \chi_\alpha(g^{-1}) = \overline{\chi_\alpha(g)}.$$

Proof ► First suppose $k = \mathbb{C}$. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $\alpha(g)$. Then

$$\chi_\alpha(g) = \operatorname{tr} \alpha(g) = \lambda_1 + \cdots + \lambda_n.$$

In fact, we can diagonalise $\alpha(g)$, ie we can find a basis with respect to which

$$g \mapsto A(g) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Now

$$A(g^{-1}) = A(g)^{-1} = \begin{pmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{pmatrix}$$

and so

$$\chi_\alpha(g^{-1}) = \operatorname{tr} A(g^{-1}) = \lambda_1^{-1} + \cdots + \lambda_n^{-1}.$$

But since G is a finite group, $g^n = e$ for some n (eg for $n = \|G\|$), and so

$$\lambda_i^n = 1 \Longrightarrow |\lambda_i| = 1 \Longrightarrow \lambda_i^{-1} = \overline{\lambda_i}$$

for each eigenvalue λ_i. Hence

$$\chi_\alpha(g^{-1}) = \overline{\lambda_1} + \cdots + \overline{\lambda_n} = \overline{\chi_\alpha(g)}.$$

The result for $k = \mathbb{R}$ follows from this. For if A is a real matrix satisfying $A^n = I$ then we may regard A as a complex matrix, and so deduce by the argument above that

$$\operatorname{tr}(A^{-1}) = \overline{\operatorname{tr} A}.$$

But since A is real, so is $\operatorname{tr} A$, and thereforeHence

$$\operatorname{tr}(A^{-1}) = \operatorname{tr} A.$$

◄

Corollary 8.1 *Suppose α, β are representations of G over k. Then*

$$I(\alpha, \beta) = \begin{cases} \frac{1}{\|G\|} \sum_{g \in G} \chi_\alpha(g)\chi_\beta(g) & \text{if } k = \mathbb{R} \\ \frac{1}{\|G\|} \sum_{g \in G} \overline{\chi_\alpha(g)}\chi_\beta(g) & \text{if } k = \mathbb{C} \end{cases}$$

Definition 8.2 *We define the inner product*

$$\langle u, v \rangle \quad (u(g), v(g) \in C(G, k))$$

by

$$\langle u, v \rangle = \begin{cases} \frac{1}{\|G\|} \sum_{g \in G} \overline{u(g)}v(g) & \text{if } k = \mathbb{C} \\ \frac{1}{\|G\|} \sum_{g \in G} u(g)v(g) & \text{if } k = \mathbb{R} \end{cases}$$

Proposition 8.3 *1. The inner product $\langle u, v \rangle$ is positive-definite.*

2. $I(\alpha, \beta) = \langle \chi_\alpha, \chi_\beta \rangle$.

Proposition 8.4 *Two representations are equivalent if and only if their characters are equal:*

$$\alpha = \beta \Longleftrightarrow \chi_\alpha(g) = \chi_\beta(g) \text{ for all } g \in G.$$

Proof ► If $\alpha = \beta$ then
$$B(g) = PA(g)P^{-1}$$
for some P. Hence
$$\chi_\beta(g) = \operatorname{tr} B(g) = \operatorname{tr} A(g) = \chi_\alpha(g).$$

On the other hand, suppose $\chi_\alpha(g) = \chi_\beta(g)$ for all $g \in G$. Then for each simple representation σ of G over k,

$$\begin{aligned} I(\sigma, \alpha) &= \frac{1}{\|G\|} \sum_{g \in G} \chi_\sigma(g^{-1}) \chi_\alpha(g) \\ &= \frac{1}{\|G\|} \sum_{g \in G} \chi_\sigma(g^{-1}) \chi_\beta(g) \\ &= I(\sigma, \beta). \end{aligned}$$

It follows that σ occurs the same number of times in α and β. Since this is true for all simple representations σ,
$$\alpha = \beta.$$

◄

Proposition 8.5 *Characters are* class functions, *ie*
$$g' \sim g \Longrightarrow \chi_\alpha(g') = \chi_\alpha(g).$$

Remark: Recall that we write $g' \sim g$ to mean that g', g are *conjugate,* ie there exists an $x \in G$ such that
$$g' = xgx^{-1}.$$

Proof ► If
$$g' = xgx^{-1}$$
then (since a representation $g \mapsto A(g)$ is a homomorphism)

$$\begin{aligned} A(g') &= A(x)A(g)A(x^{-1}) \\ &= A(x)A(g)A(x)^{-1}. \end{aligned}$$

It follows from the basic property of the trace that
$$\chi_\alpha(g') = \operatorname{tr} A(g') = \operatorname{tr} A(g) = \chi_\alpha(g).$$

◄

Proposition 8.6 *Simple characters are orthogonal, ie if α, β are distinct simple representations of G over k then*

$$\langle \chi_\alpha, \chi_\beta \rangle = 0.$$

Proof ► This is just a restatement of the fact that

$$I(\alpha, \beta) = 0.$$

◄

When $k = \mathbb{C}$ we can be a little more precise.

Proposition 8.7 *The simple characters of G over $\mathbb{C}$ form an orthonormal set, ie*

$$\langle \chi_\alpha, \chi_\beta \rangle = \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{otherwise.} \end{cases}$$

Proof ► Again, this is simply a restatement of the result for the intertwining number. ◄

Theorem 8.2 *The group G has at most s simple represenations over k, where s is the number of classes in G.*

Proof ► The class functions on G form a vector space

$$X \subset C(G, k).$$

Lemma 8.2 $\dim X = s$.

Proof of Lemma ▷ Suppose the conjugacy classes are $C_1, \ldots, C_n$. Let $c_i(g)$ denote the *characteristic function* of C_i, ie

$$c_i(g) = \begin{cases} 1 & \text{if } g \in C_i, \\ 0 & \text{otherwise} \end{cases}$$

Then the functions

$$c_i(g) \quad (1 \leq i \leq s)$$

form a basis for the class functions on G. ◁

Lemma 8.3 *Mutually orthogonal vectors (with respect to a positive-definite form) are necessarily linearly independent.*

Proof of Lemma ▷ Suppose $v_1, \ldots, v_r$ are mutually orthogonal:

$$\langle v_i, v_j \rangle = 0 \text{ if } i \neq j.$$

Suppose

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0.$$

Taking the inner product of v_i with this relation,

$$\lambda_1 \langle v_i, v_1 \rangle + \cdots + \lambda_r \langle v_i, v_r \rangle = 0 \Longrightarrow \lambda_i = 0.$$

Since this is true for all i, the vectors $v_1, \ldots, v_r$ must be linearly independent. ◁

Now consider the simple characters of G over k. They are mutually orthogonal, by the last Proposition; and so they are linearly independent, by the Lemma. But they belong to the space X of class functions. Hence their number cannot exceed the dimension of this space, which by Lemma 1 is s. ◄

Remark: We shall see that when $k = \mathbb{C}$, the number of simple representations is actually *equal to* the number of classes. This is equivalent, by the reasoning above, to the statement that *the characters span the space of class functions.*

Our major aim now is to establish this result. We shall give 2 proofs, one based on induced representations, and one of the representation theory of product groups.

Example: Since characters are class functions, it is only necessary to compute their values for 1 representative from each class. The *character table* of a group G over k tabulates the values of the simple representations on the various classes. By convention, if the scalar field k is not specified it is understood that we are speaking of representations over $\mathbb{C}$.

As an illustration, let us take the group S_3. The 6 elements divide into 3 classes, corresponding to the 3 cylic types:

1^3 e

21 $(bc), (ac), (ab)$

3 $(abc), (acb)$

It follows that S_3 has at most 3 simple characters over $\mathbb{C}$. Since we already know 3, namely the 2 1-dimensional representations $1, \epsilon$ and the 2-dimensional representation α, we have the full panoply.

We draw up the character table as follows:

class size	$[1^3]$ 1	[21] 3	[3] 2
1	1	1	1
ϵ	1	-1	1
α	2	0	-1

Let us verify that the simple characters form an orthonormal set:

$$\begin{aligned}
I(1,1) &= \frac{1}{6}(1\cdot 1\cdot 1 + 3\cdot 1\cdot 1 + 2\cdot 1\cdot 1) = 1\\
I(1,\epsilon) &= \frac{1}{6}(1\cdot 1\cdot 1 + 3\cdot 1\cdot -1 + 2\cdot 1\cdot 1) = 0\\
I(1,\alpha) &= \frac{1}{6}(1\cdot 1\cdot 2 + 3\cdot 1\cdot 0 + 2\cdot 1\cdot -1) = 0\\
I(\epsilon,\epsilon) &= \frac{1}{6}(1\cdot 1\cdot 1 + 3\cdot -1\cdot -1 + 2\cdot 1\cdot 1) = 1\\
I(\epsilon,\alpha) &= \frac{1}{6}(1\cdot 1\cdot 2 + 3\cdot -1\cdot 0 + 2\cdot 1\cdot -1) = 0\\
I(\alpha,\alpha) &= \frac{1}{6}(1\cdot 2\cdot 2 + 3\cdot 0\cdot 0 + 2\cdot -1\cdot -1) = 1
\end{aligned}$$

It is very easy to compute the character of a permutational representation, that is, a representation arising from the action of the group G on the finite set X. Recall that this is the representation in the function-space $C(X,k)$ given by

$$(gf)(x) = f(g^{-1}x).$$

Proposition 8.8 *Suppose α is the permutational representation of G arising from the action of G on the finite set X. Then*

$$\chi_a lpha(g) = \|\{x : gx = x\}\|,$$

ie $\chi(g)$ is equal to the number of elements of X left fixed by g.

Proof ► Let $c_x(t)$ denote the characteristic function of the 1-point subset $\{x\}$, ie

$$c_x(t) = \begin{cases} 1 & \text{if } t = x,\\ 0 & \text{otherwise.}\end{cases}$$

The $\|X\|$ functions $c_x(t)$ form a basis for the vector space $C(X,k)$; and the action of $g \in G$ on this basis is given by

$$gc_x = c_{gx},$$

since

$$gc_x(t) = c_x(g^{-1}t) = 1 \Longleftrightarrow g^{-1}t = x \Longleftrightarrow t = gx.$$

It follows that with respect to this basis

$$g \mapsto A(g),$$

where $A = A(g)$ is the matrix with entries

$$A_{xy} = \begin{cases} 1 & \text{if } x = gy, \\ 0 & \text{otherwise.} \end{cases}$$

In particular

$$A_{xx} = \begin{cases} 1 & \text{if } x = gx, \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$\chi_a lpha(g) = \operatorname{tr} A = \sum_x A_{xx} = \|\{x : gx = x\}\|.$$

◄

Example: Consider the action of the group S_3 on $X = \{a, b, c\}$, Let us denote the resulting representation by ρ. We only need to compute $\chi_\rho(g)$ for 3 values of g, namely 1 representative of each class.

We know that

$$\chi_\rho(e) = \dim \rho = \|X\| = 3.$$

The transposition (bc) (for example) has just 1 fixed point, namely a. Hence

$$\chi_\rho(bc) = 1.$$

On the other hand, the 3-cycle (abc) has no fixed points, so

$$\chi_\rho(abc) = 0.$$

Let us add this character to our table:

class size	$[1^3]$ 1	$[21]$ 3	$[3]$ 2
1	1	1	1
ϵ	1	-1	1
α	2	0	-1
ρ	3	1	0

We know that ρ is some integral combination of the simple characters, say

$$\rho = r \cdot 1 + s \cdot \epsilon + t \cdot \alpha,$$

where $r, s, t \in \mathbb{N}$. These 'coefficients' r, s, t are unique, since the simple characters are linearly independent.

It would be easy to determine them by observation. But let us compute them from the character of ρ. Thus

$$\begin{aligned} r = I(1, \rho) &= \frac{1}{6}(1 \cdot 1 \cdot 3 + 3 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot 0) = 1 \\ s = I(\epsilon, \rho) &= \frac{1}{6}(1 \cdot 1 \cdot 3 + 3 \cdot -1 \cdot 1 + 2 \cdot 1 \cdot 0) = 0 \\ t = I(\alpha, \rho) &= \frac{1}{6}(1 \cdot 2 \cdot 3 + 3 \cdot 0 \cdot 1 + 2 \cdot -1 \cdot 0) = 0 \end{aligned}$$

Thus

$$\rho = 1 + \alpha.$$

Chapter 9

The Regular Representation

The group G acts on itself in 3 ways:

- *By left multiplication:* $(g, x) \mapsto gx$
- *By right multiplication:* $(g, x) \mapsto xg^{-1}$
- *By inner automorphism:* $(g, x) \mapsto gxg^{-1}$

The first action leads to the *regular* representation defined below. The second action also leads to the regular representation, as we shall see. The third action leads to the *adjoint* representation, which we shall consider later.

Definition 9.1 *The* regular representation $\mathbf{reg}$ *of the group G over k is the permutational representation defined by the action*

$$(g, x) \mapsto gx$$

of G on itself.

Proposition 9.1 *The character of the regular representation is given by*

$$\chi_{\mathbf{reg}}(g) = \begin{cases} 1 & \text{if } g = e, \\ 0 & \text{otherwise}. \end{cases}$$

Proof ► We have to determine, for each $g \in G$, the number of elements $x \in G$ left fixed by g, ie satisfying

$$gx = x.$$

But

$$gx = x \Longrightarrow g = e.$$

Thus no element $g \neq e$ leaves any element fixed; while $g = e$ leaves every element fixed. ◄

Proposition 9.2 *The permutational representation defined by right multiplication*

$$(g,x) \mapsto xg^{-1}$$

is equivalent to the regular representation.

Proof ▶ No element $g \neq e$ leaves any element fixed; while $g = e$ leaves every element fixed:

$$xg^{-1} = x \Longleftrightarrow g = e.$$

Thus this representation has the same character as the regular representation; and so it is equal (that is, equivalent) to it. ◀

Alternative proof ▶ In fact it is readily verified that the representation defined by right multiplication is the *dual* $\mathbf{reg}^*$ of the regular representation. But the regular representation is self-dual, since its character is real. ◀

Proposition 9.3 *Suppose α is a representation of G over k. Then*

$$I(\alpha, \mathbf{reg}) = \dim \alpha.$$

Proof ▶ Plugging the result for the character of $\mathbf{reg}$ above into the formula for the intertwining number,

$$\begin{aligned} I(\alpha, \mathbf{reg}) &= \frac{1}{\|G\|} \sum_{g \in G} \chi_\alpha(g^{-1}) \chi_{\mathbf{reg}}(g) \\ &= \frac{1}{\|G\|} \|G\| \chi_\alpha(e) \\ &= \dim \alpha. \end{aligned}$$

◀

This result shows that *every simple representation occurs in the regular representation*, since $I(\sigma, \mathbf{reg}) > 0$. When $k = \mathbb{C}$ we can be more precise.

Proposition 9.4 *Each simple representation σ of G over $\mathbb{C}$ occurs just $\dim \sigma$ times in the regular representation $\mathbf{reg}$ of G over $\mathbb{C}$:*

$$\mathbf{reg} = \sum_\sigma (\dim \sigma) sigma,$$

where the sum extends over all simple representations σ of G over $\mathbb{C}$.

Proof ► We know that $\mathbf{reg}$, as a semisimple representation, is expressible in the form

$$\mathbf{reg} = \sum_\sigma e_\sigma \sigma \quad (e_\sigma \in \mathbb{N}).$$

Taking the intertwining number of a particular simple representation σ with each side,

$$\begin{aligned} I(\sigma, \mathbf{reg}) = e_\sigma I(\sigma, \sigma) = e_\sigma \\ &= \dim \sigma, \end{aligned}$$

by the Proposition. ◄

Theorem 9.1 *The dimensions of the simple representations* $\sigma_1, \ldots, \sigma_r$ *of* G *over* $\mathbb{C}$ *satisfy the relation*

$$\dim^2 \sigma_1 + \cdots + \dim^2 \sigma_r = \|G\|.$$

Proof ► This follows at once on taking the dimensions on each side of the identity

$$\mathbf{reg} = \sum_\sigma (\dim \sigma) sigma.$$

◄

Example: Consider S_5. We have

$$\|S_5\| = 120;$$

while S_5 has 7 classes:

$$[1^5],\ [21^3],\ [2^2 1],\ [31^2],\ [32],\ [41],\ [5].$$

Thus S_5 has at most 7 simple representations over $\mathbb{C}$.

Let us review the information on these representations that we already have:

1. S_5 has just 2 1-dimensional representations, 1 and ϵ;

2. The natural 5-dimensional representation ρ of S_5 splits into 2 parts:

   $$\rho = 1 + \alpha,$$

   where α is a simple 4-dimensional representation of S_5;

3. If σ is a simple representation of S_5 of odd dimension then $\epsilon\sigma \neq \sigma$;

4. More generally, if σ is a simple representation of S_5 with $\sigma([21^3]) \neq 0$ then $\epsilon\sigma \neq \sigma$;

We can apply this last result to α. For

$$\begin{aligned} \chi_\alpha([21^3]) &= \chi_\rho([21^3]) - 1 \\ &= 3 - 1 \\ &= 2. \end{aligned}$$

Hence

$$\epsilon\alpha \neq \alpha.$$

Thus we have found 4 of the 7 (or fewer) simple representations of S_5: $1, \epsilon, \alpha, \epsilon\alpha$. Our dimensional equation reads

$$120 = 1^2 + 1^2 + 4^2 + 4^2 + a^2 + b^2 + c^2,$$

where $a, b, c \in \mathbb{N}$, with $a, b, c \neq 1$. (We are allowing for the fact that S_5 might have < 7 simple representations.) In other words,

$$a^2 + b^2 + c^2 = 86.$$

It follows that

$$a^2 + b^2 + c^2 \equiv 6 \pmod 8.$$

Now

$$n^2 \equiv 0, 1, \text{ or } 4 \pmod 8$$

according as $n \equiv 0 \pmod 4$, or n is odd, or $n \equiv 2 \pmod 4$. The only way to get 6 is as $4 + 1 + 1$. In other words, 2 of a, b, c must be odd, and the other must be $\equiv 2 \pmod 4$. (In particular $a, b, c \neq 0$. So S_5 must in fact have 7 simple representations.)

By (3) above, the 2 odd dimensions must be equal: say $a = b$. Thus

$$2a^2 + c^2 = 86.$$

Evidently $a = 3$ or 5. Checking, the only solution is

$$a = b = 5, c = 6.$$

We conclude that S_5 has 7 simple representations, of dimensions

$$1, 1, 4, 4, 5, 5, 6.$$

Chapter 10

Induced Representations

> Each representation of a group defines a representation of a subgroup, by restriction; that much is obvious. More subtly, each representation of the subgroup defines a representation of the full group, by a process called *induction*. This provides the most powerful tool we have for constructing group representations.

Definition 10.1 *Suppose H is a subgroup of G; and suppose α is a representation of G in V. Then we denote by α_H the representation of H in the same space V defined by restricting the group action from G to H. We call α_H the* restriction *of α to H.*

Proposition 10.1

1. $(\alpha + \beta)_H = \alpha_H + \beta_H$
2. $(\alpha\beta)_H = \alpha_H\beta_H$
3. $(\alpha^*)_H = (\alpha_H)^*$
4. $1_H = 1$
5. $\dim \alpha_H = \dim \alpha$
6. $\chi_{\alpha_H}(h) = \chi_\alpha(h)$

Example: We can learn much about the representations of G by considering their restrictions to subgroups $H \subset G$. But induced representations give us the same information—and more—much more easily, as we shall see; so the following example is of more intellectual interest than practical value.

Let us see what we can discover about the simple characters of S_4 (over $\mathbb{C}$) from the character table for S_3. Let's assume we know—as we shall prove later in this chapter—that the number of simple characters of S_4 is equal to the number of classes, 5. Let's suppose too that we know S_4 has just 2 1-dimensional representations, 1 and ϵ. Let γ be one of the 3 other simple representations of S_4.

Let

$$\gamma_{S_3} = a1 + b\epsilon + c\alpha \quad (a, b, c \in \mathbb{N}).$$

By the Proposition above, if $\bar{h} \subset \bar{g}$ (where $\bar{h}$ is a class in H and $\bar{g}$ a class in G) then

$$\chi_\gamma(\bar{g}) = \chi_{\gamma_H}(\bar{h}).$$

So we know some of the values of χ_γ:

Class size	$[1^4]$ 1	$[21^2]$ 6	$[2^2]$ 3	$[31]$ 8	$[4]$ 6
1	1	1	1	1	1
ϵ	1	-1	1	1	-1
γ	$a+b+2c$	$a-b$	x	$a+b-c$	y

We have found nothing about $\chi([2^2])$ and $\chi([4])$, since these 2 classes don't intersect S_3. However, if we call the values x and y as shown, then the 2 equations

$$I(1, \gamma) = 0, \qquad I(\epsilon, \gamma) = 1$$

give

$$\begin{aligned} 15a + 3b - 6c + 3x + 6y &= 0 \\ 3a + 15b - 6c + 3x + 6y &= 0 \end{aligned}$$

Setting

$$s = a + b, \quad t = a - b,$$

for simplicity, these yield

$$x = -3s + 2t, \quad y = -t.$$

The table now reads

Class size	$[1^4]$ 1	$[21^2]$ 6	$[2^2]$ 3	$[31]$ 8	$[4]$ 6
1	1	1	1	1	1
ϵ	1	-1	1	1	-1
γ	$s+2c$	t	$-3s+2c$	$s-c$	$-t$

Since γ is—by hypothesis—simple,

$$I(\gamma, \gamma) = 1.$$

Thus

$$24 = (s+2c)^2 + 6t^2 + 3(-3s+2c)^2 + 8(s-c)^2 + 6t^2.$$

On simplification this becomes

$$\begin{aligned} 2 &= 3s^2 - 4sc + 2c^2 + t^2 \\ &= s^2 + 2(s-c)^2 + t^2. \end{aligned}$$

Noting that s, t, c are all integral, and that $s, c \geq 0$, we see that there are just 3 solutions to this diophantine equation:

$$(a, b, c) = (1, 0, 1),\ (0, 1, 1),\ (0, 0, 1).$$

These must yield the 3 missing characters.

We have determined the character table of S_4 without constructing—even implicitly—the corresponding representations. This has an interesting parallel in recent mathematical history. One of the great achievements of the last 25 years has been *the determination of all finite simple groups*, ie groups possessing no proper normal (or self-conjugate) subgroups. The last link in the chain was the determination of the *exceptional* simple groups, ie those not belonging to the known infinite families (such as the family of alternating groups A_n for $n \geq 5$). Finally, all was known except for the largest exceptional group—the so-called *mammoth group*. The character table of this group had been determined several years before it was established that a group did indeed exist with this table.

As we remarked earlier, the technique above is *not* recommended for serious character hunting. The method of choice must be induced representations, our next topic.

Suppose V is a vector space. Then we denote by $C(G, V)$ the G-space of maps $f : G \to V$, with the action of G defined by

$$(gf)(x) = f(g^{-1}x)$$

(This extends our earlier definition of $C(G, k)$.)

Definition 10.2 *Suppose H is a subgroup of G; and suppose α is a representation of H in U. Then we define the induced representation α^G of G as follows. Let*

$$V = \{F \in C(G, U) : F(gh) = h^{-1}F(g) \quad \text{for all } g \in G, h \in H\}.$$

Then V is a G-subspace of $C(G, U)$; and α^G is the representation of G in this subspace.

Remark: That V is a G-subspace follows from the fact that we are acting on G with G and H from opposite sides (G on the left, H on the right); and their actions therefore commute:

$$(gx)h = g(xh).$$

Thus if $F \in V$ then

$$\begin{aligned}(gF)(xh) &= F(g^{-1}xh)\\ &= h^{-1}F(g^{-1}x)\\ &= h^{-1}((gF)(x)),\end{aligned}$$

ie $gF \in V$.

This definition is too cumbersome to be of much practical use. The following result offers an alternative, and usually more convenient, starting point.

Lemma 10.1 *Suppose $e = g_1, g_2, ..., g_r$ are representatives of the cosets of H in G, ie*

$$G = g_1H \cup g_2H \cup ... \cup g_rH.$$

Then there exists an H-subspace $U' \subset V$ such that

(a) U' is isomorphic to U as an H-space,

(b) $V = g_1U' \bigoplus g_2U' \bigoplus ... \bigoplus g_rU'$.

Moreover the induced representation α^G is uniquely characterised by the existence of such a subspace.

Remarks:

1. After the lemma, we may write

$$V = g_1U \bigoplus g_2U \bigoplus ... \bigoplus g_rU.$$

2. The action of G on V is implicit in this description of V. For suppose v is in the ith summand, say

$$v = g_iu;$$

and suppose gg_i is in the jth coset, say

$$gg_i = g_jh.$$

Then gv is in the jth summand:

$$gv = gg_iu = g_j(hu).$$

3. The difficulty of taking this result as the definition of α^G lies in the awkwardness of showing that the resulting representation does not depend on the choice of coset-representatives.

Proof ► To each $u \in U$ let us associate the function $u' = u'(g) \in C(G, U)$ by

$$u'(g) = \begin{cases} gu & \text{if } g \in H \\ 0 & \text{otherwise} \end{cases}$$

Then it is readily verified that

(a) $u' \in V$, ie $u'(gh) = h^{-1}u'(g) \quad$ for all $h \in H$.

(b) If $u \mapsto u'$ then $hu \mapsto hu'$.

Thus the map $u \mapsto u'$ sets up an H-isomorphism between U and an H-subspace $U' \subset V$.

Suppose $F \in V$. From the definition of V,

$$F(gh) = h^{-1}F(g).$$

It follows that the values of F on any coset g_iH are completely determined by its value at one point g_i. Thus F is completely determined by its r values

$$u_1 = F(e), u_2 = F(g_2), ..., u_r = F(g_r).$$

Let us write

$$F \longleftrightarrow (u_1, u_2, ..., u_r).$$

Then it is readily verified that

$$u' \longleftrightarrow (u, 0, ..., 0);$$

and more generally

$$g_iu' \longleftrightarrow (0, .., u, .., 0),$$

ie the function g_iu' vanishes on all except the ith coset g_iH, and takes the value u at g_i.

It follows that

$$F = g_1u_1' + g_2u_2' + \ldots + g_ru_r'$$

since the 2 functions take the same values at the r points g_i. Moreover the argument shows that this expression for $F \in V$ as a sum of functions in the subspaces $U' = g_1U', g_2U', ..., g_rU'$, respectively, is unique: so that

$$V = g_1U' \bigoplus g_2U' \bigoplus \cdots \bigoplus g_rU'.$$

Finally this uniquely characterises the representation α^G, since the action of G on V is completely determined by the action of H on U, as we saw in Remark 1 above. ◀

Example: Suppose α is the representation of S_3 in $U = \mathbb{C}^2$ given by

$$(abc) \mapsto \begin{pmatrix} \omega & 0 \\ 0 & \omega^{-1} \end{pmatrix} \qquad (ab) \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Let us consider the representation of S_4 induced by α (where we identify S_3 with the subgroup of S_4 leaving d fixed).

First we must choose representatives of the S_3-cosets in S_4. The nicest way to choose coset representatives of H in G is to find—if we can—a subgroup $T \subset G$ *transverse* to H, ie such that

1. $T \cap H = \{e\}$
2. $\|T\| \, \|H\| = \|G\|$.

It is readily verified that these 2 conditions imply that each element $g \in G$ is *uniquely* expressible in the form

$$g = th \quad (t \in T, h \in H)$$

It follows that the elements of T represent the cosets gH of H in G.

In the present case we could take T to be the subgroup generated by a 4-cycle: say

$$\{e, (abcd), (ac)(bd), (adcb)\}.$$

Or we could take

$$T = V_4 = \{e, (ab)(cd), (ac)(bd), (ad)(bc)\}$$

(the Viergruppe). Let's make the latter choice; the fact that T is normal (self-conjugate) in G should simplify the calculations. We have

$$S_4 = S_3 \cup (ab)(cd)S_3 \cup (ac)(bd)S_3 \cup (ad)(bc)S_3;$$

and so α^G is the represention in the 8-dimensional vector space

$$V = U \bigoplus (ab)(cd)U \bigoplus (ac)(bd)U \bigoplus (ad)(bc)U.$$

As basis for this space we may take

$$\begin{array}{llll} e_1 = e, & e_2 = f, & e_3 = (ab)(cd)e, & e_4 = (ab)(cd)f, \\ e_5 = (ac)(bd)e, & e_6 = (ac)(bd)f, & e_7 = (ad)(cb)e, & e_8 = (ad)(bc)f, \end{array}$$

where $e = (1, 0), f = (0, 1)$.

To simplify our calculations, recall that if $g, x \in S_n$, and

$$x = (a_1 a_2 \dots a_r)(b_1 b_2 \dots b_s) \dots$$

in cyclic notation, then

$$gxg^{-1} = (ga_1, ga_2, \dots, ga_r)(gb_1, gb_2, \dots, gb_s) \dots,$$

since, for example,

$$(gxg^{-1})(ga_1) = gxa_1 = ga_2.$$

(This is how we show that 2 elements of S_n are conjugate if and only if they are of the same type.) In our case, suppose $h \in S_3, t \in V_4$. Then

$$hth^{-1} \in V_4$$

since V_4 is normal. In other words,

$$ht = sh,$$

where $s \in V_4$.

Now let's determine the matrix representing (ab). By the result above, we have

$$\begin{aligned}
(ab) \cdot (ab)(cd) &= (ab)(cd) \cdot (ab) \\
(ab) \cdot (ac)(bd) &= (bc)(ad) \cdot (ab) \\
(ab) \cdot (ad)(bc) &= (bd)(ac) \cdot (ab).
\end{aligned}$$

Thus

$$\begin{aligned}
(ab)e_6 &= (ab) \cdot (ac)(bd)f \\
&= (ad)(bc) \cdot (ab)f \\
&= (ad)(bc)e \\
&= e_7.
\end{aligned}$$

In fact

$$\begin{aligned}
(ab)e_1 &= e_2, & (ab)e_2 &= e_1, \\
(ab)e_3 &= e_4, & (ab)e_4 &= e_3, \\
(ab)e_5 &= e_8, & (ab)e_6 &= e_7, \\
(ab)e_7 &= e_6, & (ab)e_8 &= e_5.
\end{aligned}$$

Hence

$$(ab) \mapsto \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

It is not hard to see that (ab) and $(abcd)$ generate S_4. So the representation α^{S_4} will be completely determined—in principle, at least—if we establish the matrix representing $(abcd)$. We see now that it was easier to detemine the matrix representing (ab), because $(ab) \in S_3$. But the general case is not difficult. Notice that

$$(abcd) = (ad)(bc) \cdot (ac)$$

It follows that (for example)

$$\begin{aligned} (abcd) \cdot (ac)(bd) &= (ad)(bc) \cdot (ac) \cdot (ac)(bd) \\ &= (ad)(bc) \cdot (ac)(bd) \cdot (ac) \\ &= (ab)(cd) \cdot (ac). \end{aligned}$$

Now $(ac) = (abc)(ab)$; so under α,

$$(ac) \mapsto \begin{pmatrix} 0 & \omega \\ \omega^{-1} & 0 \end{pmatrix}$$

We see that, for example,

$$\begin{aligned} (abcd)e_5 &= (abcd) \cdot (ac)(bd)e \\ &= (ab)(cd) \cdot (ac)e \\ &= (ab)(cd)\omega^{-1} f \\ &= \omega^{-1} e_2. \end{aligned}$$

We leave it to the reader to complete this calculation of the matrix representing $(abcd)$.

Clearly this is too time-consuming a hobby to pursue.

It is evident that

$$h \sim h' \text{ in } H \Longrightarrow h \sim h' \text{ in } G$$

In other words, each class $\bar{h}$ in H lies in a unique class $\bar{g}$ in G:

$$\bar{h} \subset \bar{g}.$$

Or, to put it the other way round, each class $\bar{g}$ in G *splits* into classes $\overline{h_1}, \ldots, \overline{h_r}$ in H:

$$\bar{g} \cap H = \overline{h_1} \cup \cdots \cup \overline{h_r}.$$

Theorem 10.1 *Suppose H is a subgroup of G; and suppose β is a representation of H. Then*

$$\chi_{\beta^G}(\bar{g}) = \frac{\|G\|}{\|H\|\ \|\bar{g}\|} \sum_{\bar{h} \subset \bar{g}} \|\bar{h}\| \chi_\beta(\bar{h}),$$

where the sum runs over those H-classes $\bar{h}$ contained in $\bar{g}$.

Proof ► Let $g_1, \ldots, g_r$ be representatives of the cosets of H in G, so that β^G is the representation in

$$V = g_1 U \bigoplus \cdots \bigoplus g_r U.$$

Lemma 10.2 *With the notation above*

$$\chi_{\beta^G}(g) = \sum_{i : g_i^{-1} g g_i = h \in H} \chi_\beta(h),$$

where the sum extends over those coset-representatives g_i for which $g_i^{-1} g g_i \in H$.

Proof ► Let us extend the function χ_β (which is of course defined on H) to G by setting

$$\chi_\beta(g) = 0 \text{ if } g \notin H,$$

then our formula can be written:

$$\chi_{\beta^G}(g) = \sum_i \chi_\beta(g_i^{-1} g g_i),$$

with the sum now extending over all coset-representatives.

Suppose $e_1, ..., e_m$ is a basis for U. Then $g_i e_j$ $(1 \le i \le r, 1 \le j \le m)$ is a basis for V.

Suppose v belongs to the ith summand of V, say

$$v = g_i u;$$

and suppose $g g_i$ belongs to the jth coset, say

$$g g_i = g_j h.$$

Then

$$gv = g g_i u = g_j (hu).$$

So

$$g(g_iU) \subset g_jU.$$

Thus the basis elements in g_iU cannot contribute to $\chi_{\beta^G}(g)$ unless $i = j$, that is, unless $gg_i = g_ih$, ie

$$g_i^{-1}gg_i = h \in H.$$

Moreover if this is so then

$$g(g_ie_j) = g_i(he_j),$$

ie the $m \times m$ matrix defining the action of g on g_iU with respect to the basis $g_ie_1, ..., g_ie_m$ is just $B(h)$; so that its contribution to $\chi_{\beta^G}(g)$ is

$$\chi_\beta(h).$$

The result follows on adding the contributions from all those summands sent into themselves by g. ◄

Lemma 10.3 *For each $g \in G$,*

$$\chi_{\beta^G}(g) = \frac{1}{\|H\|} \sum_{g' \in G: g'^{-1}gg' = h \in H} \chi_\beta(h)$$

Proof ► Suppose we take a different representative of the ith coset, say

$$g_i' = g_ih.$$

This will make the same contribution to the sum, since

$$g_i'^{-1}gg_i' = h^{-1}(g_i^{-1}gg_i)h;$$

and

$$\chi_\beta(h^{-1}h'h) = \chi_\beta(h').$$

Thus if we sum over all the elements of G, we shall get each coset-contribution just $\|H\|$ times. ◄

To return to the proof of the Proposition, we compute how many times each element $h \in H$ occurs in the sum above.

Two elements g', g'' define the same conjugate of g in G, ie

$$g'^{-1}gg' = g''^{-1}gg'',$$

if and only if $g''g'^{-1}$ and g commute, ie if and only

$$g''N(g) = g'N(g),$$

where

$$N(g) = \{x \in G : gx = xg\}.$$

It follows that each G-conjugate h of g in H will occur just $\|N(g)\|$ times in the sum of Corollary 1. Thus if we sum over these elements h we must multiply by $\|N(g)\|$.

The result follows, since

$$|N(g)\| = \frac{\|G\|}{\|\bar{g}\|}$$

by the same argument, each conjugate $x^{-1}gx$ of g arising from $\|N(g)\|$ elements x. ◄

Examples:

1. Let us look again at $\alpha^{S_3 \to S_4}$. The classes of S_4 and S_3 are related as follows:

$$\begin{aligned}
[1^4] \cap S_3 &= [1^3] \\
[21^2] \cap S_3 &= [21] \\
[2^2] \cap S_3 &= \emptyset \\
[31] \cap S_3 &= [3] \\
[4] \cap S_3 &= \emptyset
\end{aligned}$$

Hence

$$\begin{aligned}
\chi_{\alpha^{S_4}}([1^4]) &= \frac{24}{6 \cdot 1}\chi_\alpha(1^3) = 8 \\
\chi_{\alpha^{S_4}}([21^2]) &= \frac{24}{6 \cdot 6}3\chi_\alpha(21) = 0 \\
\chi_{\alpha^{S_4}}([2^2]) &= 0 \\
\chi_{\alpha^{S_4}}([31]) &= \frac{24}{6 \cdot 8}2\chi_\alpha(3) = -1 \\
\chi_{\alpha^{S_4}}(4) &= 0
\end{aligned}$$

Class	$[1^4]$	$[21^2]$	$[2^2]$	$[31]$	$[4]$
size	1	6	3	8	6
α^{S_4}	8	0	0	-1	0

Since

$$I(\alpha^{S_4}, \alpha^{S_4}) = \frac{1}{24}\left(8^2 + 8 \cdot 1^2\right) = 3,$$

α^{S_4} has just 3 distinct simple parts. whose determination is left to the reader.

The relation between S_4 and S_3 is unusual, in that *classes never split*. If $\bar{g}$ is a class in S_4 then $\bar{g} \cap S_3$ is either a whole class $\bar{h}$ in H, or else is empty. This is true more generally for S_n and S_m ($m < n$), where S_m is identified with the subgroup of S_n leaving the last $n - m$ elements fixed. If $\bar{g}$ is a class in S_n, then

$$\bar{g} \cap S_m = \bar{h} \text{ or } \emptyset.$$

2. Now let's look at the cyclic subgroup

$$C_4 = \langle (abcd) \rangle = \{e, (abcd), (ac)(bd), (adcb)\}$$

of S_4. Since C_4 is abelian, each element is in a class by itself. Let θ be the 1-dimensional representation of C_4 defined by

$$(abcd) \mapsto i$$

We have

$$\begin{aligned}
[1^4] \cap C_4 &= \{e\} \\
[21^2] \cap C_4 &= \emptyset \\
[2^2] \cap C_4 &= \{(ac)(bd)\} \\
[31] \cap C_4 &= \emptyset \\
[4] \cap C_4 &= \{(abcd), (adcb)\}
\end{aligned}$$

Hence

$$\begin{aligned}
\chi_{\theta^{S_4}}([1^4]) &= \frac{24}{4 \cdot 1} \chi_\theta(e) = 6 \\
\chi_{\theta^{S_4}}([21^2]) &= 0 \\
\chi_{\theta^{S_4}}([2^2]) &= \frac{24}{4 \cdot 3} \chi_\theta((ac)(bd)) = -2 \\
\chi_{\theta^{S_4}}([31]) &= 0 \\
\chi_{\theta^{S_4}}([4]) &= \frac{24}{4 \cdot 6} \left(\chi_\theta((abcd)) + \chi_\theta((adcb)) \right) \\
&= i + (-i) = 0
\end{aligned}$$

Class	$[1^4]$	$[21^2]$	$[2^2]$	$[31]$	$[4]$
size	1	6	3	8	6
θ^{S_4}	6	-2	0	0	0

Since

$$I(\theta^{S_4}, \theta^{S_4}) = \frac{1}{24}\left(6^2 + 6 \cdot 2^2\right) = 3,$$

θ^{S_4} has just 3 distinct simple parts. whose elucidation we again leave to the reader.

Proposition 10.2 *1.* $(\alpha + \alpha')^G = \alpha^G + \alpha'^G$*;*

2. $(\alpha^*)^G = (\alpha^G)^*$*;*

3. $\dim \alpha^G = [G : H] \dim \alpha$.

It is worth noting that permutation representations are a particular case of induced representations.

Lemma 10.4 *Suppose G acts transitively on the finite set X. Let α be the corresponding representation of G. Take $x \in X$; and let*

$$S_x = \{g \in G : gx = x\}$$

be the corresponding stabilizer subgroup. Then

$$\alpha = 1^{S_x \to G},$$

ie α is the representation of G obtained by induction from the trivial representation of S_x.

Remark: The result is easily extended to non-transitive actions. For in that case the set splits into a number of orbits, on each of which G acts transitively. On applying the Proposition to each orbit, we conclude that any permutation representation can be expressed as a sum of representations, each of which arises by induction from the trivial representation of some subgroup of G.

Proof ► By Definition 1,

$$\alpha' = 1^{S_x \to G}$$

is the representation in the subspace

$$V \subset C(G)$$

consisting of those functions $F : G \to k$ satisfying

$$F(gh) = h^{-1}F(g) \quad \forall h \in S_x.$$

But since S_x acts trivially on k this condition reduces to

$$F(gh) = F(g),$$

ie F is constant on each coset gS_x. Thus V can be identified with the space $C(G/S_x, k)$ of functions on the set G/S_x of S_x-cosets in G.

On the other hand, the G-sets X and G/S_x can be identified, with the element $gx \in X$ corresponding to the coset gS_x. Thus

$$C(G/S_x, k) = C(X, k).$$

Since α is by definition the representation of G in C_X the result follows. ◄

Proof ► (Alternative) By Proposition 2,

$$\begin{aligned} \chi_{\alpha'}(g) &= \|\{i : g_i^{-1} g g_i \in S_x\}\| \\ &= \|\{i : g g_i x = g_i x\}\| \\ &= \|\{y \in X : gy = y\}\|, \end{aligned}$$

since each $y \in X$ is uniquely expressible in the form $y = g_i x$. But by Proposition ???,

$$\chi_\alpha(g) = \|\{y \in X : gy = y\}\|.$$

Thus

$$\chi_\alpha = \chi_{\alpha'},$$

and so

$$\alpha = \alpha' = 1^{S_x \to G}.$$

◄

Induced representations are of great practical value. But we end with an extremely important theoretical application.

Proposition 10.3 *The number of simple representations of a finite group G is equal to the number of conjugacy classes in G.*

Proof ► Let s denote the number of classes in G. We already know that

- The characters of G are class functions.
- The simple characters are linearly independent.

Thus G has at most s simple characters; and the result will follow if we can show that every class function is a linear combination of characters.

It suffices for the latter to show that we can find a linear combination of characters taking the value 1 on a given class $\bar{g}$, and vanishing on all other classes.

We can extend the formula in the Theorem above to define a map

$$f(\bar{h}) \mapsto f^G(\bar{g}) : \; X(H,k) \to X(G,k)$$

from the space $X(H,k)$ of class functions on the subgroup $H \subset G$ to the space $X(G,k)$ of class functions on G, by

$$f^G(\bar{g}) = \frac{\|G\|}{\|H\|\|\bar{g}\|} \sum_{\bar{h} \subset \bar{g}} \|\bar{h}\| f(\bar{h}).$$

Evidently this map is linear:

$$F(h) = af(h) + bg(h) \Longrightarrow F^G(g) = af^G(g) + bf^G(g).$$

Choose any $g \in \bar{g}$. Let H be the subgroup generated by g. Thus if g is of order d,

$$H = C_d = \langle g \rangle = \{e, g, g^2, \ldots, g^{d-1}\}.$$

Let θ denote the 1-dimensional character on H defined by

$$\theta(g) = \omega = e^{2\pi i/d}.$$

Since H is abelian, each element is in a class by itself, so all functions on H are class functions. The d characters on H are

$$1, \theta, \theta^2, \ldots, \theta^{d-1}.$$

Let $f(h)$ denote the linear combination

$$f = 1 + \omega^{-1}\theta + \omega^{-2}\theta^2 + \cdots + \omega^{-(d-1)}\theta^{d-1}.$$

Then

$$f(h^i) = \begin{cases} d & \text{if } i = 1, \\ 0 & \text{if } i = 0. \end{cases},$$

ie f vanishes off the H-class $\{g\}$, but is not identically 0.

It follows that the induced function $f^G(g)$ has the required property; it vanishes off $\bar{g}$, while

$$f^G(\bar{g}) = \frac{\|G\|d}{\|H\|\|\bar{g}\|} \neq 0.$$

This proves the result, since f^G is a linear combination of characters:

$$f^G = 1^G + \omega\theta^G + \omega^2(\theta^2)^G + \cdots + \omega^{-(d-1)}(\theta^{d-1})^G.$$

◄

Examples:

1. S_3 has 3 classes: 1^3, 21 and 3. So it has 3 simple representations over $\mathbb{C}$, as of course we already knew: namely 1, ϵ and α.

2. D(4) has 5 classes: $\{e\}$, $\{s^2\}$, $\{s, s^3\}$, $\{c, d\}$ and $\{h, v\}$. So it has 5 simple representations over $\mathbb{C}$. We already know of 4 1-dimensional representations. In addition the matrices defining the natural 2-dimensional representation in $\mathbb{R}^2$ also define a 2-dimensional complex representation. (We shall consider this process of complexification more carefully in Chapter ???.) This representation must be simple, since the matrices do not commute, as they would if it were the sum of 2 1-dimensional representations. Thus all the representations of D_4 are accounted for.

Proposition 10.4 *(Frobenius' Reciprocity Theorem) Suppose α is a representation of G, and β a representation of $H \subset G$. Then*

$$I_G(\alpha, \beta^G) = I_H(\alpha_H, \beta).$$

Proof ► We have

$$\begin{aligned} I_G(\alpha, \beta^G) &= \frac{1}{\|G\|} \sum_{\bar{g}} \|\bar{g}\| \overline{\chi_\alpha(\bar{g})} \frac{\|G\|}{\|H\|\|\bar{g}\|} \sum_{\bar{h} \subset \bar{g}} \|\bar{h}\| \chi_\beta(\bar{h}) \\ &= \frac{1}{\|H\|} \sum_{\bar{h}} \|\bar{h}\| \overline{\chi_\alpha(\bar{h})} \chi_\beta(\bar{h}) \\ &= I_H(\alpha_H, \beta). \end{aligned}$$

◄

This short proof does not explain *why* Frobenius' Reciprocity Theorem holds. For that we must take a brief excursion into category theory.

Let $\mathcal{C}_G$ denote the category of G-spaces and G-maps. Then restriction and induction define *functors*

$$S : \mathcal{C}_G \to \mathcal{C}_H, \quad , I : \mathcal{C}_H \to \mathcal{C}_G.$$

Now 2 functors

$$E : \mathcal{C}_1 \to \mathcal{C}_2, \quad F : \mathcal{C}_2 \to \mathcal{C}_1$$

are said to be *adjoint* if for any 2 objects $X \in \mathcal{C}_1, Y \in \mathcal{C}_2$ there are bijections

$$\mathcal{M}_{\mathcal{C}_1}(X, FY) = \mathcal{M}_{\mathcal{C}_2}(EX, Y)$$

which are natural in the sense that given any morphism

$$f : X \to X'$$

in $\mathcal{C}_1$ the diagram

$$\begin{array}{ccc} \mathcal{M}(X, FY) & \longleftarrow & \mathcal{M}(X', FY) \\ \| & & \| \\ \mathcal{M}(EX, Y) & \longleftarrow & \mathcal{M}(EX', Y) \end{array}$$

is commutative, and similarly given any morphism

$$e : Y \to Y'$$

in $\mathcal{C}_2$ the diagram

$$\begin{array}{ccc} \mathcal{M}(X, FY) & \longrightarrow & \mathcal{M}(X, FY') \\ \| & & \| \\ \mathcal{M}(EX, Y) & \longrightarrow & \mathcal{M}(EX, Y') \end{array}$$

is commutative.

It's not difficult to establish—but would take us too far out of our way—that the *induction and restriction functors are adjoint* in this sense: if V is a G-space, and U a H-space, then

$$\mathrm{hom}^H(V_H, U) = \mathrm{hom}^G(V, U^G).$$

On taking dimensions, this gives Frobenius' Theorem:

$$I_H(\alpha_H, \beta) = I_G(\alpha, \beta^G).$$

Chapter 11

Representations of Product Groups

The representations of a product group $G \times H$ can be expressed—in as neat a way as one could wish—in terms of the representations of G and H.

Definition 11.1 *Suppose α is a representation of G in the vector space U over k, and β a representation of H in the vector space V over k. Then we denote by $\alpha \times \beta$ the representation of the product group $G \times H$ in the tensor product $U \otimes V$ defined by*

$$(g, h) \sum u \otimes v = \sum gu \otimes hv.$$

Lemma 11.1 *1.* $\chi_{\alpha\times\beta}(g, h) = \chi_\alpha(g)\chi_\beta(h)$

2. $\dim(\alpha \times \beta) = \dim \alpha \dim \beta$

3. if α and β are both representations of G then

$$(\alpha \times \beta)_G = \alpha\beta,$$

where the restriction is to the diagonal subgroup

$$G = \{(g, g) : g \in G\} \subset G \times G.$$

Proposition 11.1 *The representation $\alpha \times \beta$ of $G \times H$ over $\mathbb{C}$ is simple if and only if α and β are both simple. Moreover, every simple representation of $G \times H$ is of this form.*

Proof ►

Lemma 11.2 *If α_1, α_2 are representations of G, and β_1, β_2 are representations of H, all over k, then*

$$I(\alpha_1 \times \beta_1, \alpha_2 \times \beta_2) = I(\alpha_1, \beta_1)I(\alpha_2, \beta_2)$$

Proof ► We have

$$\begin{aligned} I(\alpha_1 \times \beta_1, \alpha_2 \times \beta_2) &= \frac{1}{|G||H|} \sum_{(g,h)\in G\times H} \overline{\chi_{\alpha_1\times\beta_1}(g,h)} \chi_{\alpha_2\times\beta_2}(g,h) \\ &= \frac{1}{|G||H|} \sum_{(g,h)\in G\times H} \overline{\chi_{\alpha_1}(g)\chi_{\beta_1}(h)} \chi_{\alpha_2}(g)\chi_{\beta_2}(h) \\ &= \frac{1}{|G|} \sum_{g\in G} \overline{\chi_{\alpha_1}(g)} \chi_{\alpha_2}(g) \frac{1}{|G|} \sum_{h\in H} \overline{\chi_{\beta_1}(h)} \chi_{\beta_2}(h) \\ &= I(\alpha_1, \beta_1) I(\alpha_2, \beta_2) \end{aligned}$$

◄

Recall that a representation α over $\mathbb{C}$ is simple if and only if

$$I(\alpha, \alpha) = 1.$$

Thus if α is a simple representation of G and β is a simple representation of H (both over $\mathbb{C}$) then

$$I(\alpha \times \beta, \alpha \times \beta) = I(\alpha, \alpha) I(\beta, \beta) = 1;$$

and therefore $\alpha \times \beta$ is simple.

Now suppose G has r classes and H has s classes. Then $G \times H$ has rs classes, since

$$(g, h) \sim (g', h') \iff g \sim g' \text{ and } h \sim h'.$$

But we have just produced rs simple representations $\alpha \times \beta$ of $G \times G$; so these are in fact the full complement.

(The lemma shows that these representations are distinct; for

$$I(\alpha_1 \times \beta_1, \alpha_2 \times \beta_1) = I(\alpha_1, \alpha_2) I(\beta_1, \beta_2) = 0$$

unless $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$.) ◄

It is useful to give a proof of the last part of the Proposition not using the fundamental result that the number of simple representations is equal to the number of classes; for we can give an alternative proof of this result using product groups.

Proof ► (of last part of Proposition). Suppose γ is a representation of $G \times H$ in W over $\mathbb{C}$.

Consider the restriction γ_H of γ to the subgroup $H = e \times H \subset G \times H$. Let V be a simple part of W_H:

$$W_H = V \oplus \cdots$$

Let

$$X = \hom^H(V, W)$$

be the vector space formed by the H-maps $t : V \to W$. This is non-trivial since V is a subspace of W.

Now X is a G-space, under the action

$$(gt)(v) = g(tv)$$

Let U be a simple G-subspace of X. Then

$$\begin{aligned} \hom^G(U, X) &= \hom^G(U, \hom^H(V, W)) \\ &= \hom^{G\times H}(U \otimes V, W). \end{aligned}$$

Since this space is non-trivial, there exists a $G \times H$ map

$$\theta : U \otimes V \to W.$$

But since both $U \otimes V$ and W are simple, we must have

$$\ker \theta = 0, \quad \operatorname{im} \theta = W.$$

Hence θ is an isomorphism, ie

$$W = U \otimes V.$$

Thus

$$\gamma = \alpha \times \beta,$$

where α is the representation of G in U, and β is the representation of H in V. ◄

Theorem 11.1 *Suppose G has n elements and s classes. Then*

1. *G has s simple representations over $\mathbb{C}$;*
2. *if these are $\sigma_1, \dots, \sigma_s$ then*

$$\dim^2 \sigma_1 + \cdots + \dim^2 \sigma_s = n.$$

Proof ► Let τ be the permutation representation of $G \times G$ in $C(G, k)$ induced by the action

$$(g, h)x = gxh^{-1}$$

of $G \times G$ on G.

Lemma 11.3 *The character of τ is given by*

$$\chi_\tau(g, h) = \begin{cases} |G|/|\bar{g}| & \textit{if } g \sim h \\ 0 & \textit{otherwise} \end{cases}$$

Proof ▶ Since τ is a permutational representation,

$$\begin{aligned}\chi_\tau(g,h) &= |\{x : (g,h)x = x\}| \\ &= |\{x : gxh^{-1} = x\}| \\ &= |\{x : x^{-1}gx = h\}|.\end{aligned}$$

If $g \not\sim h$ then clearly no such x exists.
Suppose $g \sim h$. Then there exists at least one x, say x_0, such that

$$h = x_0^{-1}gx_0.$$

Now

$$\begin{aligned}x^{-1}gx = h &\iff x^{-1}gx = x_0^{-1}gx_0 \\ &\iff (xx_0^{-1})g = g(xx_0^{-1}) \\ &\iff xx_0^{-1} \in Z(g) \\ &\iff x \in Z(g)x_0.\end{aligned}$$

Thus

$$\begin{aligned}\chi_\tau(g,h) &= |\{x : gxh^{-1} = x| \\ &= |Z(g)| \\ &= |G|/|\bar{g}|.\end{aligned}$$

◀

Lemma 11.4 *Suppose G has simple representations $\sigma_1, \ldots \sigma_s$. Then*

$$\tau = \sigma_1^* \times \sigma_1 + \cdots + \sigma_s^* \times \sigma_s.$$

Proof ▶ We know that the simple representations of $G \times G$ are $\sigma_i \times \sigma_j$. Thus

$$\tau = \sum_{i,j} e(i,j)\sigma_i \times \sigma_j,$$

where $e(i.j) \in \mathbb{N}$.

To determine $e(i,j)$ we must compute the intertwining number

$$\begin{aligned}
I(\tau, \sigma_i \times \sigma_j) &= \frac{1}{|G|^2} \sum_{g,h} \overline{\chi_\tau(g,h)} \chi_{\sigma_i \times \sigma_j}(g,h) \\
&= \frac{1}{|G|^2} \sum_{g,h} \overline{\chi_\tau(g,h)} \chi_{\sigma_i}(g) \chi_{\sigma_j}(h) \\
&= \frac{1}{|G|^2} \sum_{h \sim g} \frac{|G|}{|\bar{g}|} \chi_{\sigma_i}(g) \chi_{\sigma_j}(h) \\
&= \frac{1}{|G|} \chi_{\sigma_i}(g) \chi_{\sigma_j}(g) \\
&= \frac{1}{|G|} \overline{\chi_{\sigma_i^*}(g)} \chi_{\sigma_j}(g) \\
&= I(\sigma_i^*, \sigma_j).
\end{aligned}$$

Thus

$$I(\tau, \sigma_i \times \sigma_j) = \begin{cases} 1 & \text{if } \sigma_i^* = \sigma_j \\ 0 & \text{otherwise} \end{cases}$$

In other words, $\sigma_i \times \sigma_j$ occurs in τ if and only if $\sigma_i^* = \sigma_j$, and then occurs just once. ◀

It follows from this result in particular that the number of simple representations is equal to $I(\tau, \tau)$.

Lemma 11.5 *$I(\tau, \tau)$ is equal to the number of classes in G.*

Proof ▶ We have

$$\begin{aligned}
I(\tau, \tau) &= \frac{1}{|G|^2} \sum_{g,h} |\chi_\tau(g,h)|^2 \\
&= \frac{1}{|G|^2} \sum_g \sum_{h \sim g} \left(\frac{|G|}{|\bar{g}|} \right)^2 \\
&= \frac{1}{|G|^2} \sum_g |\bar{g}| \frac{|G|^2}{|\bar{g}|^2} \\
&= \sum_g \frac{1}{|\bar{g}|}.
\end{aligned}$$

Since each class contributes $|\bar{g}|$ terms to this sum, each equal to $1/|\bar{g}|$, the sum is equal to the number of classes. ◀

That proves the first part of the Theorem; the number of simple representations is equal to $I(\tau, \tau)$, which in turn is equal to the number of classes.

The second part follows at once on taking dimensions in

$$\tau = \sigma_1^* \times \sigma_1 + \cdots + \sigma_s^* \times \sigma_s.$$

◀

Example: We can think of product groups in 2 ways—as a method of constructing new groups, or as a way of splitting up a given group into factors.

We say that $G = H \times K$, where H, K are subgroups of G, if the map

$$H \times K \to G : (h, k) \mapsto hk$$

is an isomorphism.

A necessary and sufficient condition for this— supposing G finite—is that

1. elements of H and K commute, ie

   $$hk = kh$$

   for all $h \in H, k \in K$; and

2. $|G| = |H||K|$.

Now consider the symmetry group G of a cube. This has 48 elements; for there are 8 vertices, and 6 symmetries leaving a given vertex fixed.

Of these 48 symmetries, half are proper and half improper. The proper symmetries form a subgroup $P \subset G$.

Let $Z = \{I, J\}$, where J denotes reflection in the centre of the cube. In fact Z is the centre of G:

$$Z = ZG = \{z \in G : zg = gz \text{ for all } g \in G\}.$$

By the criterion above,

$$G = Z \times P.$$

Moreover,

$$P = S_4,$$

as we can see by considering the action of symmetries on the 4 diagonals of the cube. This defines a homomorphism

$$\Theta : P \to S_4.$$

Since no symmetry send every diagonal into itself,

$$\ker \Theta = \{I\}.$$

Thus Θ is injective; and so it is bijective, since

$$|P| = 24 = |S_4|.$$

Hence Θ is an isomorphism.

Thus

$$G = C_2 \times S_4.$$

In theory this allows us to dtermine the character table of G from that of S_4. However, to make use of this table we must know how the classes of $C_2 \times S_4$ are to be interpreted geometrically. This is described in the following table.

class in $C_2 \times P$	size	order	$geometrical description$
$\{I\} \times 1^4$	1	1	identity
$\{J\} \times 1^4$	1	1	reflection in centre
$\{I\} \times 21^2$	6	2	half-turn about axis joining centres of opposite edges
$\{J\} \times 21^2$	6	2	relflection in plane through opposite edges
$\{I\} \times 2^2$	3	2	rotation about axis parallel to edge through π
$\{J\} \times 2^2$	3	2	relflection in central plane parallel to face
$\{I\} \times 31$	8	3	rotation about diagonal through $\pm\frac{\pi}{3}$
$\{J\} \times 31$	8	6	screw reflection about diagonal
$\{I\} \times 4$	6	4	rotation about axis parallel to edge through $\pm\frac{\pi}{2}$
$\{J\} \times 4$	6	4	screw reflection about axis parallel to edge

The character table of $C_2 \times S_4$ id readily derived from that of S_4. We denote the non-trivial character of C_2 ($J \mapsto -1$) by η.

Class	$I \times 1^4$	$J \times 1^4$	$I \times 21^2$	$J \times 21^2$	$I \times 2^2$	$J \times 2^2$	$I \times 31$	$J \times 31$	$I \times 4$	$J \times 4$
1×1	1	1	1	1	1	1	1	1	1	1
$\eta \times 1$	1	-1	1	-1	1	-1	1	-1	1	-1
$1 \times \epsilon$	1	1	-1	-1	1	1	1	1	-1	-1
$\eta \times \epsilon$	1	-1	-1	1	1	-1	1	-1	-1	1
$1 \times \alpha$	2	2	0	0	2	2	-1	-1	0	0
$\eta \times \alpha$	2	-2	0	0	2	-2	-1	1	0	0
$1 \times \beta$	3	3	1	1	-1	-1	0	0	-1	-1
$\eta \times \beta$	3	-3	1	-1	-1	1	0	0	-1	1
$1 \times \epsilon\beta$	3	3	-1	-1	-1	-1	0	0	1	1
$\eta \times \epsilon\beta$	3	-3	-1	1	-1	1	0	0	1	-1

Now suppose π is the 6-dimensional permutational representation of G induced by its action on the 6 faces of the cube. Its character is readily determined:

Class	$I \times 1^4$	$J \times 1^4$	$I \times 21^2$	$J \times 21^2$	$I \times 2^2$	$J \times 2^2$	$I \times 31$	$J \times 31$	$I \times 4$	$J \times 4$
π	6	0	0	2	2	4	0	0	2	0

For example, to determine $\chi_\pi(\{J\} \times 4)$ we note that an element of this class is a rotation about an axis parallel to an edge followed by reflection in the centre. This will send each of the 4 faces parallel to the edge into an adjacent face, and will swap the other 2 faces. Thus it will leave no face fixed; and so

$$\chi_\pi(\{J\} \times 4) = 0.$$

We have

$$I(\pi, 1 \times 1) = \frac{1}{48}(1 \cdot 1 \cdot 6 + 6 \cdot 2 \cdot 1 + 3 \cdot 2 \cdot 1 + 3 \cdot 4 \cdot 1 + 6 \cdot 2 \cdot 1) = 1$$

(as we knew it would be, since the action is transitive). Similarly,

$$\begin{aligned}
I(\pi, \eta \times 1) &= \frac{1}{48}(1 \cdot 1 \cdot 6 - 6 \cdot 2 \cdot 1 + 3 \cdot 2 \cdot 1 - 3 \cdot 4 \cdot 1 + 6 \cdot 2 \cdot 1) = 0,\\
I(\pi, 1 \times \epsilon) &= \frac{1}{48}(1 \cdot 1 \cdot 6 - 6 \cdot 2 \cdot 1 + 3 \cdot 2 \cdot 1 + 3 \cdot 4 \cdot 1 - 6 \cdot 2 \cdot 1) = 0,\\
I(\pi, \eta \times \epsilon) &= \frac{1}{48}(1 \cdot 1 \cdot 6 + 6 \cdot 2 \cdot 1 + 3 \cdot 2 \cdot 1 - 3 \cdot 4 \cdot 1 - 6 \cdot 2 \cdot 1) = 0.
\end{aligned}$$

It is clear at this point that the remaining simple parts of π must be of dimensions 2 and 3. Thus π contains either $1 \times \alpha$ or $\eta \times \alpha$. In fact

$$I(\pi, 1 \times \alpha) = \frac{1}{48}(1 \cdot 6 \cdot 2 + 3 \cdot 2 \cdot 2 + 3 \cdot 4 \cdot 2) = 1.$$

The remaining part drops out by subtraction; and we find that

$$\pi = 1 \times 1 + 1 \times \alpha + \eta \times \epsilon\beta.$$

Chapter 12

Exterior Products

12.1 The exterior products of a vector space

Suppose V is a vector space. Recall that its rth exterior product $\wedge^r V$ is a vector space, spanned by elements of the form

$$v_1 \wedge \cdots \wedge v_r \qquad (v_1, \ldots, v_r \in V),$$

where

$$v_{\pi 1} \wedge \cdots \wedge v_{\pi_r} = \epsilon(\pi) v_1 \wedge \cdots \wedge v_r$$

for any permutation $\pi \in S_r$.

This implies in particular that any product containing a repeated element vanishes:

$$\cdots \wedge v \wedge \cdots \wedge v \wedge \cdots = 0.$$

(We are assuming here that the characteristic of the scalar field k is not 2. In fact we shall only be concerned with the cases $k = \mathbb{R}$ or $\mathbb{C}$.)

The exterior product $\wedge^r V$ could be defined rigorously as the quotient-space

$$\wedge^r V = V^{\otimes r}/X,$$

where X is the subspace of $V^{\otimes r}$ spanned by all elements of the form

$$v_{\pi 1} \wedge \ldots v_{\pi_r} - \epsilon(\pi) v_1 \wedge \cdots \wedge v_r,$$

where $v_1, \ldots, v_r \in V, \pi \in S_r$, and ϵ denotes the parity representation of S_r.

Suppose $e_1, \ldots, e_n$ is a basis for V. Then

$$e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_r} \qquad (i_1 < i_2 < \cdots < i_r)$$

is a basis for $\wedge^r V$. (Note that there is one basis element corresponding to each subset of $\{e_1, \dots, e_n\}$ containing r elements.) It follows that if $\dim V = n$ then

$$\wedge^r V = 0 \text{ if } r > n;$$

while if $r \leq n$ then

$$\dim \wedge^r V = \binom{n}{r}.$$

Now suppose $T : V \to V$ is a linear map. Then we can define a linear map

$$\wedge^r T : \wedge^r V \to \wedge^r V$$

by

$$(\wedge^r T)(v_1 \wedge \cdots \wedge v_r) = (Tv_1) \wedge \cdots \wedge (Tv_r).$$

(To see that this action is properly defined, it is sufficient to see that it sends the subspace $X \subset V^{\otimes n}$ described above into itself; and that follows at once since

$$(\wedge^r T)\left(v_{\pi 1} \wedge \cdots \wedge v_{\pi_r}\right) - \epsilon(\pi) v_1 \wedge \ldots v_r = (Tv_{\pi 1}) \wedge \cdots \wedge (Tv_{\pi_r}) - \epsilon(\pi)(Tv_1) \wedge \cdots \wedge (Tv_r)$$

is again one of the spanning elements of X.)

In the case $r = n$, $\wedge^n V$ is 1-dimensional, with the basis element

$$e_1 \wedge \cdots \wedge e_n;$$

and

$$\wedge^n T = (\det T) I.$$

This is in fact the "true" definition of the determinant.

Although we shall not make use of this, the spaces $\wedge^r V$ can be combined to form the *exterior algebra $\wedge V$ of V*

$$\wedge V = \bigoplus \wedge^r V,$$

with the "wedge multiplication"

$$\wedge : \wedge^r V \times \wedge^s V \to \wedge^{r+s} V$$

defined by

$$(u_1 \wedge \cdots \wedge u_r) \wedge (v_1 \wedge \cdots \wedge v_s) = u_1 \wedge \cdots \wedge u_r \wedge v_1 \wedge \cdots \wedge v_s,$$

extended to $\wedge V$ by linearity.

Observe that if $a \in \wedge^r V, b \in \wedge^s V$ then

$$b \wedge a = (-1)^{rs} a \wedge b.$$

In particular the elements of even order form a *commutative* subalgebra of $\wedge V$.

12.2 The exterior products of a group representation

Definition 12.1 *Suppose α is a representation of G in V. Then we denote by $\wedge^r \alpha$ the representation of G in $\wedge^r V$ defined by*

$$g(v_1 \wedge \cdots \wedge v_r) = (gv_1) \wedge \cdots \wedge (gv_r).$$

In other words, g acts through the linear map $\wedge^r (\alpha(g))$.

Proposition 12.1 *Suppose $g \in G$ has eigenvalues $\lambda_1, \ldots, \lambda_n$ in the representation α. Then the character of $\wedge^r \alpha$ is the rth symmetric sum of the λ's, ie*

$$\chi_{\wedge^r \alpha}(g) = \sum_{i_1 < i_2 < \cdots < i_r} \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_r}.$$

Proof ► Let us suppose that $k = \mathbb{C}$. We know that $\alpha(g)$ can be diagonalised, ie we can find a basis $e_1, \ldots, e_n$ of the representation-space V such that

$$ge_i = \lambda_i e_i \qquad (i = 1, \ldots, n).$$

But now

$$ge_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_r} = \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_1} e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_r},$$

from which the result follows, since these products form a basis for $\wedge^r V$. ◄

12.3 Symmetric polynomials

We usually denote the symmetric product in the Proposition above by

$$\sum \lambda_1 \ldots \lambda_r.$$

It is an example of a *symmetric polynomial* in $\lambda_1, \ldots, \lambda_n$.

More generally, suppose A is a commutative ring, with 1. (In fact we shall only be interested in the rings $\mathbb{Z}$ and $\mathbb{Q}$.) As usual, $A[x_1, \ldots, x_n]$ denotes the ring of polynomials in $x_1, \ldots, x_n$ with coefficients in A.

The symmetric group S_n acts on this ring, by permutation of the variables:

$$(\pi P)(x_1, \ldots, x_n) = P\left(x_{\pi^{-1}(1)}, \ldots, x_{\pi^{-1}(n)}\right) \quad (\pi \in S_n).$$

The polynomial $P(x_1, \ldots, x_n)$ is said to be *symmetric* if it is left invariant by this action of S_n. The symmetric polynomials evidently form a *sub-ring* of $A[x_1, \ldots, x_n]$, which we shall denote by $\Sigma_n(A)$.

The n polynomials

$$a_1 = \sum x_i,\ a_2 = \sum_{i_1<i_2} x_{i_1}x_{i_2},\ \ldots, a_n = x_1\cdots x_n$$

are symmetric; as are

$$s_1 = \sum x_i,\ s_2 = \sum x_i^2, ; s_3 = \sum x_i^3, \ldots.$$

Proposition 12.2 *The ring $\Sigma_{\mathbb{Z}}(n)$ is freely generated over $\mathbb{Z}$ by $a_1, \ldots, a_n$, ie the map*

$$p(x_1, \ldots, x_n) \mapsto p(a_1, \ldots, a_n) : \mathbb{Z}[x_1, \ldots, x_n] \to \Sigma_n(\mathbb{Z})$$

is a ring-isomorphism.

Proof ► We have to show that

1. Every symmetric polynomial $P(x_1, \ldots, x_n)$ over $\mathbb{Z}$ (ie with integer coefficients) is expressible as a polynomial in $a_1, \ldots, a_n$ over $\mathbb{Z}$:

   $$P(x_1, \ldots, x_n) = p(a_1, \ldots, a_n).$$

   This will show that the map is surjective.

2. The map is injective, ie

   $$p(a_1, \ldots, a_n) \equiv 0 \Longrightarrow p \equiv 0.$$

1. Any polynomial is a linear combination of *monomials* $x_1^{e_1} \cdots x_n^{e_n}$. We order the monomials first by degree, with higher degree first, and then within each degree *lexicographically*, eg if $n = 2$ then

   $$1 < x_2 < x_1 < x_2^2 < x_1x_2 < x_1^2 < x_2^3 < \cdots.$$

   The *leading term* in $p(x_1, \ldots, x_n)$ is the non-zero term $cx_1^{e_1} \cdots x_n^{e_n}$ containing the greatest monomial in this ordering.

   Now suppose the polynomial $P(x_1, \ldots, x_n)$ is symmetric. Evidently $e_1 \geq e_2 \geq \cdots \geq e_n$ in the leading term. For if say $e_1 < e_2$ then the term $cx_1^{e_2}x_2^{e_1} \cdots x_n^{e_n}$ — which must also appear in $P(x_1, \ldots, x_n)$.

   ◄

Corollary 12.1 *The ring $\Sigma_{\mathbb{Q}}(n)$ is freely generated over $\mathbb{Q}$ by $a_1, \ldots, a_n$,*

Proposition 12.3 *The ring $\Sigma_{\mathbb{Q}}(n)$ is freely generated over $\mathbb{Q}$ by $s_1, \ldots, s_n$,*

Proof ► ◄

12.4 Newton's formula

It follows from the Propositions above that the power-sums s_n are expressible in terms of the symmetric products a_n, and vice versa. More precisely, there exist polynomials $S_n(x_1, \ldots, x_n)$ and $A_n(x_1, \ldots, x_n)$ such that

$$s_n = S_n(a_1, \ldots, \_n), \quad a_n = A_n(a_1, \ldots, \_n),$$

with the coefficients of S_n integral and those of A_n rational. Newton's formula allows these polynomials to be determined recursively.

Let

$$\begin{aligned} f(t) &= (1 - x_1 t) \cdots (1 - x_n t) \\ &= 1 - a_1 t + a_2 t^2 - \cdots + (-1)^n a_n t^n. \end{aligned}$$

Then

$$\begin{aligned} \frac{f'(t)}{f(t)} &= frac{-x_1}1 - x_1 t + \cdots + frac{-x_n}1 - x_n t \\ &= -s_1 - s_2 t - s_3 t^2 - \cdots . \end{aligned}$$

Thus

$$-a_1 + 2a_2 t - 3a_3 t^2 + \cdots + (-1)^n n a_n t^{n-1} = (1 - a_1 t + a_2 t^2 - \cdots + (-1)^n a_n t^n)(-s_1 - s_2 t - s_3 t^2 - \cdots).$$

Equating coefficients,

$$\begin{array}{rcl} a_1 & = & s_1 \\ 2a_2 & = & s_1 a_1 - s_2 \\ 3a_3 & = & s_1 a_2 - s_2 a_1 + s_3 \\ \ldots & & \\ r a_r & = & s_1 a_r - s_2 a_{r-1} + \cdots + (-1)^{r-1} s_r \\ \ldots & & \end{array}$$

Evidently these equations allow us to express $s_1, s_2, s_3, \ldots$ successively in terms of $a_1, a_2, a_3, \ldots$, or vice versa:

$$\begin{aligned}
s_1 &= a_1 \\
s_2 &= a_1^2 - 2a_2 \\
s_3 &= a_1^3 - 3a_1a_2 + 3a_3 \\
&\ldots \\
a_1 &= s_1 \\
2a_2 &= s_1^2 - s_2 \\
6a_3 &= s_1^3 - 3s_1s_2 + 2s_3 \\
&\ldots
\end{aligned}$$

12.5 Plethysm

There is another way of looking at the exterior product — as a particular case of the *plethysm* operator on the representation-ring $R(G)$.

Suppose V is a vector space over a field k of characteristic 0. (We shall only be interested in the cases $k = R$ or $\mathbb{C}$.) Then the symmetric group S_n acts on the tensor product $V^{\otimes n}$ by permutation of the factors:

$$\pi(v_1 \otimes \cdots \otimes v_n) = v_{\pi^{-1}1} \otimes \cdots \otimes v_{\pi^{-1}n}.$$

Thus $V^{\otimes n}$ carries a representation of S_n. As we know this splits into components V^Σ corresponding to the simple representations Σ of S_n:

$$V^{\otimes n} = V^{\Sigma_1} \oplus \cdots \oplus V^{\Sigma_s},$$

where $\Sigma_1, \ldots, \Sigma_s$ are the simple representations of S_n. (We shall find it convenient to use superfixes rather than suffixes for objects corresponding to representations of S_n.)

We are particularly interested in the components corresponding to the 2 1-dimensional representations of S_n: the trivial representation 1_n and the parity representation ϵ_n, and we shall write

$$V^P = V^{1_n}, \quad V^N = V^{\epsilon_n}.$$

We also use P and N to denote the operations of *symmetrisation* and *skew-symmetrisation* on $V^{\otimes n}$; that is, the linear maps

$$P, N : V^{\otimes n} \to V^{\otimes n}$$

defined by

$$P(v_1 \otimes \cdots \otimes v_n) = \frac{1}{n!} \sum_{\pi \in S_n} \pi(v_1 \otimes \cdots \otimes v_n),$$
$$N(v_1 \otimes \cdots \otimes v_n) = \frac{1}{n!} \sum_{\pi \in S_n} \epsilon(\pi)\pi(v_1 \otimes \cdots \otimes v_n).$$

Suppose $\pi \in S_n$. Regarding π as a map $V^{\otimes n} \to V^{\otimes n}$, we have

$$\pi P = P = P\pi, \quad \pi N = \epsilon(\pi)N = N\pi.$$

It follows that

$$\{P^2 = P, \quad N^2 = N,$$

ie P and N are both *projections* onto subspaces of $V^{\otimes n}$.

We say that $x \in V^{\otimes n}$ is *symmetric* if

$$\pi x = x$$

for all $\pi \in S_n$; and we say that x is *skew-symmetric* if

$$\pi x = \epsilon(\pi)x$$

for all π. It follows at once from the relations $\pi P = P, \pi N = \epsilon N$ that x is symmetric if and only if

$$Px = x;$$

while x is skew-symmetric if and only if

$$Nx = x.$$

Thus P is a projection onto the symmetric elements in $V^{\otimes n}$, and N is a projection onto the skew-symmetric elements.

To see the connection with the exterior product $\wedge^n V$, recall that we could define the latter by

$$\wedge^n V = V^{\otimes n}/X,$$

where $X \subset V^{\otimes n}$ is the subspace spanned by elements of the form

$$\pi x - \epsilon(\pi)x.$$

It is easy to see that $Nx = 0$ for such an element x; while conversely, for any $x \in V^{\otimes n}$

$$\begin{aligned} x - Nx &= \frac{1}{n!} \sum_{\pi \in S_n} \epsilon(\pi) \left(\epsilon(\pi)x - \pi x s\right) \\ &\in X \end{aligned}$$

It follows that

$$X = \ker N;$$

and so (since N is a projection)

$$\wedge^n V = V/X \cong \operatorname{im} N = V^N;$$

that is, *the nth exterior product of V can be identified with the ϵ-component of $V^{\otimes n}$.*

Now suppose that V carries a representation α of some group G. Then G acts on $V^{\otimes n}$ through the representation α^n.

Proposition 12.4 *Suppose α is a represenation of G in V. Then the actions of G and S_n on $V^{\otimes n}$ commute.*

For each simple representation Σ of S_n, the component V^Σ of $V^{\otimes n}$ is stable under G, and so carries a representation α^Σ of G. Thus

$$\alpha^n = \alpha^{\Sigma_1} + \cdots + \alpha^{\Sigma_s},$$

where $\Sigma_1, \ldots, \Sigma_s$ are the simple representations of S_n.

Proof ► We have

$$\begin{aligned}
\pi g(v_1 \otimes \cdots \otimes v_n) &= \pi(gv_1 \otimes \cdots \otimes gv_n) \\
&= (gv_{\pi^{-1}1} \otimes \cdots \otimes gv_{\pi^{-1}n}) \\
&= g\pi(v_1 \otimes \cdots \otimes v_n).
\end{aligned}$$

◄

Since the actions of G and S_n on $V^{\otimes n}$ commute, they combine to define a representation of the product group $G \times S_n$ on this space.

Corollary 12.2 *The representation of $G \times S_n$ on $V^{\otimes n}$ is given by*

$$\alpha^{\Sigma_1} \times \Sigma_1 + \cdots + \alpha^{\Sigma_s} \times \Sigma_s.$$

Suppose $g \in G$ (or more accurately, $\alpha(g)$) has eigenvalues $\lambda_1, \ldots, \lambda_d$. We know that the character of

$$\wedge^n \alpha = \alpha^{\epsilon_n}$$

is the nth symmetric product of the λ_i:

$$\chi_{\wedge^n \alpha}(g) = a_n(\lambda_1, \ldots, \lambda_d).$$

Proposition 12.5 *To each simple representation Σ of S_n there corresponds a unique symmetric function S_Σ of degree n such that for any representation α of G, and for any $g \in G$ with eigenvalues $\lambda_1, \ldots, \lambda_d$,*

$$\chi_{\wedge^n \alpha}(g) = S_\Sigma(\lambda_1, \ldots, \lambda_d).$$

Proof ► We begin by establishing an important result which should perhaps have been proved when we discussed the splitting of a G-space V into components

$$V = V_{\sigma_1} \oplus \cdots \oplus V_{\sigma_s}$$

corresponding to the simple representations $\sigma_1, \ldots, \sigma_s$ of G.

Lemma 12.1 *The projection P_σ onto the σ-component of V is given by*

$$P_\sigma = \frac{\dim \sigma}{\|G\|} \sum_{g \in G} \chi_\sigma(g^{-1}) g.$$

Proof of Lemma ▷ Suppose α is a representation of G in V. Then the formula above defines a linear map

$$P : V \to V.$$

Suppose $h \in G$. Then (writing d for $\dim \sigma$)

$$\begin{aligned} hPh^{-1} &= \frac{d}{\|G\|} \sum_g \chi_\sigma(g^{-1}) hgh^{-1} \\ &= \frac{d}{\|G\|} \sum_{g'} \chi_\sigma(h^{-1}g'^{-1}h) g' \\ &= \frac{d}{\|G\|} \sum_{g'} \chi_\sigma(g'^{-1}) g' \\ &= P. \end{aligned}$$

Now suppose α is simple. By Schur's Lemma, the only linear transformations commuting with all $\alpha(g)$ are multiples of the identity. Thus

$$P = \rho I$$

for some $\rho \in \mathbb{C}$. Taking traces,

$$\frac{d}{\|G\|} \sum_g \chi_\sigma(g^{-1}) \chi_\alpha(g) = \rho d.$$

It follows that

$$\rho = \begin{cases} 1 & \alpha = \sigma \\ 0 & \alpha \neq \sigma \end{cases}$$

◁ In other words,

$$P = \begin{cases} I & \alpha = \sigma \\ 0 & \alpha \neq \sigma \end{cases}$$

It follows that P acts as the identity on all simple G-subspaces carrying the representation σ, and as 0 on all simple subspaces carrying a representation $\sigma' \neq \sigma$. In particular, $P = I$ on V_σ and $P = 0$ on $V_{\sigma'}$ for all $\sigma' \neq \sigma$. In other words, P is the projection onto the component V_σ.

◀

Chapter 13

Real Representations

> Representation theory over $\mathbb{C}$ is much simpler than representation theory over $\mathbb{R}$. For that reason, we usually *complexify* real representations—extend the scalars from $\mathbb{R}$ to $\mathbb{C}$—just as we do with polynomial equations. But at the end of the day we must determine if the representations—or solutions—that we have obtained are in fact real.

Suppose U is a vector space over $\mathbb{R}$. Then we can define a vector space $V = \mathbb{C}U$ over $\mathbb{C}$ by "extension of scalars". More precisely,

$$V = \mathbb{C} \otimes_{\mathbb{R}} U.$$

In practical terms,

$$V = U \oplus iU,$$

ie each element $v \in V$ is uniquely expressible in the form

$$v = u_1 + iu_2 \quad (u_1, u_2 \in U).$$

If $e_1, \ldots, e_n$ is a basis for U over $\mathbb{R}$, then it is also a basis for V over $\mathbb{C}$. In particular,

$$\dim_{\mathbb{C}} V = \dim_{\mathbb{R}} U.$$

On the other hand, suppose V is a vector space over $\mathbb{C}$. Then we can define a vector space $U = \mathbb{R}V$ over $\mathbb{R}$ by "forgetting" scalar multiplication by non-reals. Thus the elements of U are precisely the same as those of V. If $e_1, \ldots, e_n$ is a basis for V over V, then $e_1, ie_1, e_2, ie_2, \ldots, e_n, ie_n$ is a basis for U over $\mathbb{R}$. In particular,

$$\dim_{\mathbb{R}} U = 2 \dim_{\mathbb{C}} V.$$

Now suppose G acts on the vector space U over $\mathbb{R}$. Then G also acts on $\mathbb{C}U$, by

$$g(u_1 + iu_2) = (gu_1) + i(gu_2).$$

On the other hand, suppose G acts on the vector space V over $\mathbb{C}$. Then G also acts on $\mathbb{R}V$ by the same rule

$$(g, v) \mapsto gv.$$

Definition 13.1 *1. Suppose β is a real representation of G in U. Then we denote by $\mathbb{C}\beta$ the complex representation of G in the vector space*

$$\mathbb{C}U = U \oplus iU$$

derived from U by extending the scalars from $\mathbb{R}$ to $\mathbb{C}$.

2. Suppose α is a complex representation of G in V. Then we denote by $\mathbb{R}\alpha$ the real representation of G in the vector space $\mathbb{R}V$ derived from V by "forgetting" scalar multiplication by non-reals.

Remarks:

1. Suppose β is described in matrix terms, by choosing a basis for U and giving the matrices $B(g)$ representing $\beta(g)$ with respect to this basis. Then we can take the same basis for $\mathbb{C}U$, and the same matrices to represent $\mathbb{C}\beta(g)$. Thus from the matrix point of view, β and $\mathbb{C}\beta$ appear the same. The essential difference is that $\mathbb{C}\beta$ may split even if β is simple, ie we may be able to find a complex matrix P such that

$$PB(g)P^{-1} = \begin{pmatrix} C(g) & 0 \\ 0 & D(g) \end{pmatrix}$$

for all $g \in G$, although no real matrix P has this property.

2. Suppose α is described in matrix form, by choosing a basis $e_1, e_2, \ldots, e_n$ for V, and giving the matrices $A(g)$ representing $\alpha(g)$ with respect to this basis. Then we can take the $2n$ elements $e_1, ie_1, e_2, ie_2, \ldots, e_n, ie_n$ as a basis for $\mathbb{R}V$; and the matrix representing $\mathbb{R}\alpha(g)$ with respect to this basis can be derived from the matrix $A = A(g)$ representing $\alpha(g)$ as follows. By definition,

$$ge_r = \sum_s A_{sr} e_s.$$

Let

$$A_{r,s} = X_{r,s} + iY_{r,s},$$

where $X_{r,s}, Y_{r,s} \in \mathbb{R}$. Then

$$\begin{aligned} ge_r &= X_{sr}e_s + Y_{sr}ie_s \\ g(ie_r) &= -Y_{sr}e_s + X_{sr}ie_s \end{aligned}$$

Thus the entry A_{rs} is replaced in $\mathbb{R}\alpha(g)$ by the 2×2-matrix

$$\begin{pmatrix} X_{r,s} & -Y_{r,s} \\ Y_{r,s} & X_{r,s} \end{pmatrix}$$

Proposition 13.1 *1.* $\mathbb{C}(\beta + \beta') = \mathbb{C}\beta + \mathbb{C}\beta'$

2. $\mathbb{C}(\beta\beta') = (\mathbb{C}\beta)(\mathbb{C}\beta')$

3. $\mathbb{C}(\beta^*) = (\mathbb{C}\beta)^*$

4. $\mathbb{C}1 = 1$

5. $\dim \mathbb{C}\beta = \dim \beta$

6. $\chi_{\mathbb{C}\beta}(g) = \chi_\beta(g)$

7. $I(\mathbb{C}\beta, \mathbb{C}\beta') = I(\beta, \beta')$

8. $\mathbb{R}(\alpha + \alpha') = \mathbb{R}\alpha + \mathbb{R}\alpha'$

9. $\mathbb{R}(\alpha^*) = (\mathbb{R}\alpha)^*$

10. $\dim \mathbb{R}\alpha = 2 \dim \alpha$

11. $\chi_{\mathbb{R}\alpha}(g) = 2\Re\chi_\alpha(g) = \chi_\alpha(g) + \chi_\alpha(g^{-1})$

12. $\mathbb{R}\mathbb{C}\beta = 2\beta$

13. $\mathbb{C}\mathbb{R}\alpha = \alpha + \alpha^*$

Proof ► All is immediate except (perhaps) parts (11) and (13).

11. Suppose $\alpha(g)$ is represented by the $n \times n$ matrix

$$A = X + iY,$$

where X, Y are real. Then—as we saw above—the entry A_{rs} is replaced in $\mathbb{R}\alpha(g)$ by the 2×2 matrix

$$\begin{pmatrix} X_{r,s} & -Y_{r,s} \\ Y_{r,s} & X_{r,s} \end{pmatrix}$$

Thus

$$\begin{aligned}
\operatorname{tr} \mathbb{R}\alpha(g) &= 2\sum_r X_{rr} \\
&= 2\Re(\sum_r A_{rr}) \\
&= 2\Re \operatorname{tr} \alpha(g) \\
&= 2\Re\chi_\alpha(g) \\
&= \chi_\alpha(g) + \chi_\alpha(g^{-1})
\end{aligned}$$

since

$$\overline{\chi(g)} = \chi(g^{-1}).$$

13. This now follows on taking characters, since

$$\begin{aligned}
\chi_{\mathbb{C}\mathbb{R}\alpha}(g) &= \chi_{\mathbb{R}\alpha}(g) \\
&= \chi_\alpha(g) + \chi_\alpha(g^{-1}) \\
&= \chi_\alpha(g) + \chi_{\alpha^*}(g)
\end{aligned}$$

Since this holds for all g,

$$\mathbb{C}\mathbb{R}\alpha = \alpha + \alpha^*.$$

◄

Lemma 13.1 *Given a representation α of G over $\mathbb{C}$ there exists at most one real representation β of G over $\mathbb{R}$ such that*

$$\alpha = \mathbb{C}\beta.$$

Proof ► By Proposition 1,

$$\begin{aligned}
\mathbb{C}\beta = \mathbb{C}\beta' &\implies \chi_{\mathbb{C}\beta}(g) = \chi_{\mathbb{C}\beta'}(g) \\
&\implies \chi_\beta(g) = \chi_{\beta'}(g) \\
&\implies \beta = \beta'.
\end{aligned}$$

◄

Definition 13.2 *A representation α of G over $\mathbb{C}$ is said to be real if $\alpha = \mathbb{C}\beta$ for some representation β over $\mathbb{R}$.*

Remarks:

1. In matrix terms α is real if we can find a complex matrix P such that the matrices
$$PB(g)P^{-1}$$
are real for all $g \in G$.

2. Since β is uniquely determined by α in this case, one can to some extent confuse the two (as indeed in speaking of α as real), although eg if discussing simplicity it must be made clear whether the reference is to α or to β.

Lemma 13.2 *Consider the following 3 properties of the representation α over $\mathbb{C}$:*

1. *α is real*
2. *χ_α is real, ie $\chi_\alpha(g) \in \mathbb{R}$ for all $g \in G$*
3. *$\alpha = \alpha^*$*

We have
$$(1) \implies (2) \iff (3).$$

Proof ► $(1) \implies (2)$: If $\alpha = \mathbb{C}\beta$ then
$$\chi_\alpha(g) = \chi_\beta(g).$$
But the trace of a real matrix is necessarily real.
$(2) \iff (3)$: If χ_α is real then
$$\chi_\alpha(g) = \overline{\chi_\alpha(g)} = \chi_{\alpha^*}(g)$$
for all $g \in G$. Hence
$$\alpha = \alpha^*.$$

◄

Problems involving representations over $\mathbb{R}$ often arise in classical physics, since the spaces there are normally real, eg those given by the electric and magnetic fields, or the vibrations of a system. The best way of tackling such a problem is usually to complexify, ie to extend the scalars from $\mathbb{R}$ to $\mathbb{C}$. This allows the powerful techniques developed in the earlier chapters to be applied. But at the end of the day it may be necessary to determine whether or not the representations that arise are real. The Lemma above gives a necessary condition: if α is real then its character must be real. But this condition is not sufficient; and our aim in the rest of the Chapter is to find necessary and sufficient conditions for reality, of as practical a nature as possible.

Definition 13.3 *Suppose α is a simple representation over $\mathbb{C}$. Then we say that α is* strictly complex *if χ_α is not real; and we say that α is* quaternionic *if χ_α is real, but α itself is not real;*

Thus the simple representations of G over $\mathbb{C}$ fall into 3 mutually exclusive classes:

$\mathbb{R}$ *real:* $\alpha = \mathbb{C}\beta$

$\mathbb{C}$ *strictly complex:* χ_α not real

$\mathbb{H}$ *quaternionic:* χ_α real but α not real

Lemma 13.3 *Suppose α is a simple representation over $\mathbb{C}$. Then*

1. *If α is real, $\mathbb{R}\alpha = 2\beta$, where β is a simple representation over $\mathbb{R}$;*
2. *if α is strictly complex or quaternionic, $\mathbb{R}\alpha = \beta$ is a simple representation over $\mathbb{R}$.*

In particular, if χ_α is not real then $\mathbb{R}\alpha$ must be simple.

Proof ► If α is real, say $\alpha = \mathbb{C}\beta$, then by Proposition 1

$$\mathbb{R}\alpha = \mathbb{R}\mathbb{C}\beta = 2\beta.$$

Conversely, suppose $\mathbb{R}\alpha$ splits, say

$$\mathbb{R}\alpha = \beta + \beta'.$$

Then by Proposition 1,

$$\alpha + \alpha^* = \mathbb{C}\mathbb{R}\alpha = \mathbb{C}\beta + \mathbb{C}\beta'.$$

But since α and α^* are simple, this implies (by the unique factorisation theorem) that

$$\alpha = \mathbb{C}\beta \text{ or } \alpha = \mathbb{C}\beta'.$$

In either case α is real. ◄

This gives a (not very practical) way of distinguishing between the 3 classes:

$\mathbb{R}$: α real $\iff \chi_\alpha$ real and $\mathbb{R}\alpha$ splits

$\mathbb{C}$: α quaternionic $\iff \chi_\alpha$ real and $\mathbb{R}\alpha$ simple

$\mathbb{H}$: α strictly complex $\iff \chi_\alpha$ not real ($\implies \mathbb{R}\alpha$ simple)

The next Proposition shows that the classification of simple representations over $\mathbb{C}$ into 3 classes leads to a similar classification of simple representations over $\mathbb{R}$.

Proposition 13.2 *Suppose β is a simple representation over $\mathbb{R}$. Then there are 3 (mutually exclusive) possibilities:*

$\mathbb{R}$: $\mathbb{C}\beta = \alpha$ *is simple*

$\mathbb{C}$: $\mathbb{C}\beta = \alpha + \alpha^*$,

$\mathbb{H}$: $\mathbb{C}\beta = 2\alpha$*, with α simple with α (and α^*) simple, and $\alpha \neq \alpha^*$*

In case ($\mathbb{R}$), α is real and

$$I(\beta, \beta) = 1.$$

In case ($\mathbb{C}$), α is strictly complex and

$$I(\beta, \beta) = 2.$$

In case ($\mathbb{H}$), α is quaternionic and

$$I(\beta, \beta) = 4.$$

Proof ► Since

$$\mathbb{R}\mathbb{C}\beta = 2\beta,$$

$\mathbb{C}\beta$ cannot split into more than 2 parts. Thus there are 3 possibilities:

1. $\mathbb{C}\beta = \alpha$ is simple
2. $\mathbb{C}\beta = 2\alpha$, with α simple
3. $\mathbb{C}\beta = \alpha + \alpha'$, with α, α' simple and $\alpha \neq \alpha'$

Since

$$I(\beta, \beta) = I(\mathbb{C}\beta, \mathbb{C}\beta)$$

by Proposition 1, the values of $I(\beta, \beta)$ in the 3 cases follow at once. Thus it only remains to show that α is in the class specified in each case, and that $\alpha' = \alpha^*$ in case (3).

In case (1), α is real by definition.

In case (2),

$$2\chi_\alpha(g) = \chi_{2\alpha}(g) = \chi_\beta(g)$$

is real for all $g \in G$. Hence $\chi_\alpha(g)$ is real, and so α is either real or quaternionic. If α were real, say $\alpha = \mathbb{C}\beta'$, we should have

$$\mathbb{C}\beta = 2\mathbb{C}\beta'$$

which would imply that

$$\beta = 2\beta'$$

by Proposition 2. Hence α is quaternionic.

In case (3)

$$2\beta = \mathbb{R}\mathbb{C}\beta = \mathbb{R}\alpha + \mathbb{R}\alpha'.$$

Hence

$$\mathbb{R}\alpha = \beta = \mathbb{R}\alpha'.$$

But then

$$\alpha + \alpha' = \mathbb{C}\beta = \alpha + \alpha^*.$$

Hence

$$\alpha' = \alpha^*.$$

Finally, since $\alpha^* = \alpha' \neq \alpha$, α is strictly complex. ◄

Proposition 5 gives a practical criterion for determining which of the 3 classes a simple representation β over $\mathbb{R}$ belongs to, namely by computing $I(\beta, \beta)$ from χ_β. Unfortunately, the question that more often arises is: which class does a given simple representation α over $\mathbb{C}$ belong to? and this is more difficult to determine.

Lemma 13.4 *Suppose α is a simple representation of G over $\mathbb{C}$ in V. Then*

$\mathbb{R}$: *if α is real,there exists an invariant symmetric (quadratic) form on V, unique up to a scalar multiple—but there is no invariant skew-symmetric form on V;*

$\mathbb{C}$: *if α is complex, there is no invariant bilinear form on V.*

$\mathbb{H}$: *if α is quaternionic, there exists an invariant skew-symmetric form on V, unique up to a scalar multiple—but there is no invariant symmetric form on V;*

Proof ► A bilinear form on V is a linear map

$$V \otimes V \to \mathbb{C},$$

ie an element of

$$(V \otimes V)^* = V^* \otimes V^*.$$

Thus the space of bilinear maps carries the representation $(\alpha^*)^2$ of G. Hence the invariant bilinear maps form a space of dimension

$$I(1, (\alpha^*)^2) = I(1, \alpha^*\alpha^*) = I(\alpha, \alpha^*)$$

Since α and α^* are simple, this is 0 or 1 according as α = α^* or not, ie according as α is either real or quaternionic, or strictly complex. In other words, if α is complex there is no invariant bilinear form; while if α is real or quaternionic there is an invariant bilinear form on V, say $F(u, v)$, unique up to a scalar multiple.

Now any bilinear form can be split into a symmetric (or quadratic) part and a skew-symmetric part; say

$$F(u, v) = Q(u, v) + S(u, v),$$

where

$$Q(u, v) = \frac{1}{2}\left(F(u, v) + F(v, u)\right), \; S(u, v) = \frac{1}{2}\left(F(u, v) - F(v, u)\right)$$

But it is easy to see that if F is invariant then so are Q and S. Since F is the only invariant bilinear form on V, it follows that either

$$F = Q \text{ or } F = S,$$

ie F is either symmetric or skew-symmetric. It remains to show that the former occurs in the real case, the latter in the quaternionic case.

Suppose α is real, say $\alpha = \mathbb{C}\beta$, where β is a representation in the real vector space U. We know that U carries an invariant symmetric form (in fact a positive-definite one), say $Q(u, u')$. But this defines an invariant symmetric form $\mathbb{C}Q$ on $V = \mathbb{C}U$ by extension of scalars. So if α is real, V carries an invariant symmetric form.

Finally, suppose α is quaternionic. Then V carries either a symmetric or a skew-symmetric invariant form (but not both). Suppose the former; say $Q(v, v')$ is invariant. By Proposition 3, $\beta = \mathbb{R}\alpha$ is simple. Hence there exists a real invariant positive-definite symmetric form on $\mathbb{R}V$; and this is the only invariant symmetric form on $\mathbb{R}V$, up to a scalar multiple. But the real part of $Q(v, v')$ is also an invariant form on $\mathbb{R}V$; and it is certainly not positive-definite, since

$$\Re Q(iv, iv) = -\Re Q(v, v).$$

This contradiction shows that V cannot carry an invariant symmetric form. We conclude that it must carry an invariant skew-symmetric form. ◄

We deduce from this Proposition the following more practical criterion for reality.

Proposition 13.3 *Suppose α is a simple representation over $\mathbb{C}$. Then*

$$\frac{1}{|G|}\sum_{g\in G}\chi_\alpha(g^2) = \begin{cases} 1 & \text{if } \alpha \text{ is real} \\ 0 & \text{if } \alpha \text{ is strictly complex} \\ -1 & \text{if } \alpha \text{ is quaternionic} \end{cases}$$

Proof ► Every bilinear form has a unique expression as the sum of its symmetric and skew-symmetric parts. In other words, the space of bilinear forms is the direct sum of the spaces of symmetric and of skew-symmetric forms; say

$$V^* \otimes V^* = V^Q \oplus V^S.$$

Moreover, each of these subspaces is stable under G; so the representation $(\alpha^*)^2$ in the space of bilinear forms splits in the same way; say

$$(\alpha^*)^2 = \alpha^Q + \alpha^S,$$

where α^Q is the representation of G in the space V^Q of symmetric forms on V, and α^S is the representation in the space V^S of skew-symmetric forms.

Now the dimensions of the spaces of invariant symmetric and skew-symmetric space are

$$I(1, \alpha^Q) \text{ and } I(1, \alpha^S),$$

respectively. Thus Proposition 6 can be reworded as follows:

$\mathbb{R}$: If α is real then

$$I(1, \alpha^Q) = 1 \text{ and } I(1, \alpha^S) = 0.$$

$\mathbb{C}$: If α is complex then

$$I(1, \alpha^Q) = 0 \text{ and } I(1, \alpha^S) = 0.$$

$\mathbb{H}$: If α is quaternionic then

$$I(1, \alpha^Q) = 0 \text{ and } I(1, \alpha^S) = 1.$$

Thus all (!) we have to do is to compute these 2 intertwining numbers. In fact it suffices to find one of them, since

$$I(1, \alpha^Q) + I(1, \alpha^S) = I(1, (\alpha^*)^2) = I(\alpha, \alpha^*)$$

which we already know to be 1 if α is real or quaternionic, and 0 if α is complex.

To compute $I(1, \alpha^Q)$, choose a basis $e_1, ..., e_n$ for V; and let the corresponding coordinates be $x_1, ..., x_n$. Then the $n(n+1)/2$ quadratic forms

$$x_i^2 \quad (1 \le i \le n), \qquad 2x_i x_j \quad (1 \le i < j \le n)$$

form a basis for V^Q. Let g_{ij} denote the matrix defined by $\alpha(g)$. Thus if $v = (x_1, ..., x_n) \in V$, then the coordinates of gv are

$$(gv)_i = \sum_j g_{ij} x_j.$$

Hence

$$g(x_i^2) = \sum_{j,k} g_{ij} x_j g_{ik} x_k.$$

In particular, the coefficient of x_i^2 in this (which is all we need to know for the trace) is g_{ii}^2. Similarly, the coefficient of $2x_i x_j$ in $g(2x_i x_j)$ is

$$g_{ii} g_{jj} + g_{ij} g_{ji}.$$

We conclude that

$$\chi_{\alpha^Q}(g) = \sum_i g_{ii}^2 + \sum_{i,j: i<j} (g_{ii} g_{jj} + g_{ij} g_{ji}).$$

But

$$\chi_\alpha(g) = \sum_i g_{ii}, \ \chi_\alpha(g^2) = \sum_{i,j} g_{ij} g_{ji}.$$

Thus

$$\chi_{\alpha^Q}(g) = \frac{1}{2} \left(\chi_\alpha(g)^2 + \chi_\alpha(g^2) \right).$$

Since

$$I(1, \alpha^Q) = \frac{1}{|G|} \sum_{g \in G} \chi_{\alpha^{[2]}}(g),$$

it follows that

$$2I(1, \alpha^Q) = \frac{1}{|G|} \sum_{g \in G} \left((\chi_\alpha(g))^2 + \chi_\alpha(g^2) \right).$$

But

$$\frac{1}{|G|} \sum_g \chi_\alpha(g)^2 = I(\alpha, \alpha^*).$$

Thus

$$2I(1, \alpha^Q) = I(\alpha, \alpha^*) + \frac{1}{|G|} \sum_g \chi_\alpha(g^2).$$

The result follows, since $I(\alpha, \alpha^*) = 1$ in the real and quaternionic cases, and 0 in the complex case. ◀

Appendix A

Linear algebra over the quaternions

The basic ideas of linear algebra carry over with the quaternions $\mathbb{H}$ (or indeed any skew-field) in place of $\mathbb{R}$ or $\mathbb{C}$.

A *vector space* W *over* $\mathbb{H}$ is an abelian group (written additively) together with an operation

$$\mathbb{H} \times W \to W : \ (q, w) \mapsto qw,$$

which we cann scalar multiplication, such that

1. $q(w_1 + w_2) = qw_1 + qw_2$,
2. $(q_1 + q_2)w = q_1 w + q_2 w$,
3. $(q_1 q_2)w = q_1(q_2 w)$,
4. $1w = w$.

The notions of *basis* and *dimension* (together with *linear independence* and *spanning*) carry over without change. Thus $e_1, \ldots, e_n$ are said to be linearly independent if

$$q_1 e_1 + \cdots + q_n e_n = 0 \Longrightarrow q_1 = \cdots = q_n = 0.$$