Representation Theory

CT, Lent 2005

1 What is Representation Theory?

Groups arise in nature as "sets of symmetries (of an object), which are closed under composition and under taking inverses". For example, the *symmetric group* S_n is the group of all permutations (symmetries) of $\{1, \ldots, n\}$; the *alternating group* A_n is the set of all symmetries preserving the parity of the number of ordered pairs (did you really remember that one?); the *dihedral group* D_{2n} is the group of symmetries of the regular n-gon in the plane. The *orthogonal group* $\mathrm{O}(3)$ is the group of distance-preserving transformations of Euclidean space which fix the origin. There is also the group of *all* distance-preserving transformations, which includes the translations along with $\mathrm{O}(3)$.[1]

The official definition is of course more abstract, a group is a set G with a binary operation $*$ which is associative, has a unit element e and for which inverses exist. Associativity allows a convenient abuse of notation, where we write gh for $g * h$; we have $ghk = (gh)k = g(hk)$ and parentheses are unnecessary. I will often write 1 for e, but this is dangerous on rare occasions, such as when studying the group $\mathbb{Z}$ under addition; in that case, $e = 0$.

The abstract definition notwithstanding, the interesting situation involves a group "acting" on a set. Formally, an action of a group G on a set X is an "action map" $a : G \times X \to X$ which is *compatible with the group law*, in the sense that

$$a(h, a(g, x)) = a(hg, x) \qquad \text{and } a(e, x) = x.$$

This justifies the abusive notation $a(g, x) = g.x$ or even gx, for we have $h(gx) = (hg)x$.

From this point of view, geometry asks, "Given a geometric object X, what is its group of symmetries?" Representation theory reverses the question to "Given a group G, what objects X does it act on?" and attempts to answer this question by classifying such X up to isomorphism.

Before restricting to the linear case, our main concern, let us remember another way to describe an action of G on X. Every $g \in G$ defines a map $a(g) : X \to X$ by $x \mapsto gx$. This map is a bijection, with inverse map $a(g^{-1})$: indeed, $a(g^{-1}) \circ a(g)(x) = g^{-1}gx = ex = x$ from the properties of the action. Hence $a(g)$ belongs to the set $\mathrm{Perm}(X)$ of bijective self-maps of X. This set forms a group under composition, and the properties of an action imply that

1.1 Proposition. *An action of G on X "is the same as" a group homomorphism $\alpha : G \to \mathrm{Perm}(X)$.*

1.2 Remark. There is a logical abuse here, clearly an action, defined as a map $a : G \times X \to X$ is not the same as the homomorphism α in the Proposition; you are meant to read that specifying one is completely equivalent to specifying the other, unambiguously. But the definitions are designed to allow such abuse without much danger, and I will frequently indulge in that (in fact I denoted α by a in lecture).[2]

[1] This group is isomorphic to the *semi-direct product* $\mathrm{O}(3) \ltimes \mathbb{R}^3$ — but if you do not know what this means, do not worry.

[2] With respect to abuse, you may wish to err on the side of caution when writing up solutions in your exam!

The reformulation of Prop. 1.1 leads to the following observation. For any action a H on X and group homomorphism $\varphi : G \to H$, there is defined a *restricted* or *pulled-back* action $\varphi^* a$ of G on X, as $\varphi^* a = a \circ \varphi$. In the original definition, the action sends (g, x) to $\varphi(g)(x)$.

(1.3) Example: Tautological action of $\mathrm{Perm}(X)$ *on* X

This is the obvious action, call it T, sending (f, x) to $f(x)$, where $f : X \to X$ is a bijection and $x \in X$. Check that it satisfies the properties of an action! In this language, the action a of G on X is $\alpha^* T$, with the homomorphism α of the proposition — the pull-back under α of the tautological action.

(1.4) Linearity.

The question of classifying all possible X with action of G is hopeless in such generality, but one should recall that, in first approximation, mathematics is linear. So we shall take our X to a *vector space* over some ground *field*, and ask that the action of G be linear, as well, in other words, that it should preserve the vector space structure. Our interest is mostly confined to the case when the field of scalars is $\mathbb{C}$, although we shall occasional mention how the picture changes when other fields are studied.

1.5 Definition. A linear representation ρ of G on a complex vector space V is a set-theoretic action on V which preserves the linear structure, that is,

$$\rho(g)(\mathbf{v}_1 + \mathbf{v}_2) = \rho(g)\mathbf{v}_1 + \rho(g)\mathbf{v}_2, \forall \mathbf{v}_{1,2} \in V,$$
$$\rho(g)(k\mathbf{v}) = k \cdot \rho(g)\mathbf{v}, \forall k \in \mathbb{C}, \mathbf{v} \in V$$

Unless otherwise mentioned, *representation* will mean *finite-dimensional complex representation.*

(1.6) Example: The general linear group

Let V be a complex vector space of dimension $n < \infty$. After choosing a basis, we can identify it with $\mathbb{C}^n$, although we shall avoid doing so without good reason. Recall that the *endomorphism algebra* $\mathrm{End}(V)$ is the set of all linear maps (or *operators*) $L : V \to V$, with the natural addition of linear maps and the composition as multiplication. (If you do not remember, you should verify that the sum and composition of two linear maps is also a linear map.) If V has been identified with $\mathbb{C}^n$, a linear map is uniquely representable by a matrix, and the addition of linear maps becomes the entry-wise addition, while the composition becomes the matrix multiplication. (Another good fact to review if it seems lost in the mists of time.)

Inside $\mathrm{End}(V)$ there is contained the group $\mathrm{GL}(V)$ of *invertible* linear operators (those admitting a multiplicative inverse); the group operation, of course, is composition (matrix multiplication). I leave it to you to check that this is a group, with unit the identity operator Id. The following should be obvious enough, from the definitions.

1.7 Proposition. *V is naturally a representation of* $\mathrm{GL}(V)$.

It is called the *standard* representation of $\mathrm{GL}(V)$. The following corresponds to Prop. 1.1, involving the same abuse of language.

1.8 Proposition. *A representation of G on V "is the same as" a group homomorphism from G to* $\mathrm{GL}(V)$.

Proof. Observe that, to give a linear action of G on V, we must assign to each $g \in G$ a linear self-map $\rho(g) \in \mathrm{End}(V)$. Compatibility of the action with the group law requires

$$\rho(h)\left(\rho(g)(\mathbf{v})\right) = \rho(hg)(\mathbf{v}), \qquad \rho(1)(\mathbf{v}) = \mathbf{v}, \qquad \forall \mathbf{v} \in V,$$

whence we conclude that $\rho(1) = \mathrm{Id}$, $\rho(hg) = \rho(h) \circ \rho(g)$. Taking $h = g^{-1}$ shows that $\rho(g)$ is invertible, hence lands in $\mathrm{GL}(V)$. The first relation then says that we are dealing with a group homomorphism. □

1.9 Definition. An *isomorphism* ϕ between two representations (ρ_1, V_1) and (ρ_2, V_2) of G is a linear isomorphism $\phi : V_1 \to V_2$ which intertwines with the action of G, that is, satisfies

$$\phi(\rho_1(g)(\mathbf{v})) = \rho_2(g)(\phi(\mathbf{v})).$$

Note that the equality makes sense even if ϕ is not invertible, in which case it is just called an *intertwining operator* or *G-linear map.* However, if ϕ is invertible, we can write instead

$$\rho_2 = \phi \circ \rho_1 \circ \phi^{-1}, \tag{1.10}$$

meaning that we have an equality of linear maps after inserting any group element g. Observe that this relation determines ρ_2, if ρ_1 and ϕ are known. We can finally formulate the

Basic Problem of Representation Theory: Classify all representations of a given group G, up to isomorphism.

For arbitrary G, this is very hard! We shall concentrate on finite groups, where a very good general theory exists. Later on, we shall study some examples of topological compact groups, such as U(1) and SU(2). The general theory for compact groups is also completely understood, but requires more difficult methods.

I close with a simple observation, tying in with Definition 1.9. Given any representation ρ of G on a space V of dimension n, a choice of basis in V identifies this linearly with $\mathbb{C}^n$. Call the isomorphism ϕ. Then, by formula (1.10), we can define a new representation ρ_2 of G on $\mathbb{C}^n$, which is isomorphic to (ρ, V). So any n-dimensional representation of G is isomorphic to a representation on $\mathbb{C}^n$. The use of an abstract vector space does not lead to 'new' representation, but it does free us from the presence of a distinguished basis.

2 Lecture

Today we discuss the representations of a cyclic group, and then proceed to define the important notions of irreducibility and complete reducibility

(2.1) Concrete realisation of isomorphism classes

We observed last time that every m-dimensional representation of a group G was isomorphic to a representation on $\mathbb{C}^m$. This leads to a concrete realisation of the set of m-dimensional isomorphism classes of representations.

2.2 Proposition. *The set of m-dimensional isomorphism classes of G-representations is in bijection with the quotient*

$$\mathrm{Hom}\,(G; \mathrm{GL}(m; \mathbb{C}))\,/\mathrm{GL}(m; \mathbb{C})$$

of the set of group homomorphism to $\mathrm{GL}(m)$ *by the overall conjugation action on the latter.*

Proof. Conjugation by $\phi \in \mathrm{GL}(m)$ sends a homomorphism ρ to the new homomorphism $g \mapsto \phi \circ \rho(g) \circ \phi^{-1}$. According to Definition 1.9, this has exactly the effect of identifying isomorphic representations. □

2.3 Remark. The proposition is not as useful (for us) as it looks. It can be helpful in understanding certain infinite discrete groups — such as $\mathbb{Z}$ below — in which case the set Hom can have interesting geometric structures. However, for finite groups, the set of isomorphism classes is finite so its description above is not too enlightening.

(2.4) Example: Representations of $\mathbb{Z}$.

We shall classify all representations of the group $\mathbb{Z}$, with its additive structure. We must have $\rho(0) = \mathrm{Id}$. Aside from that, we must specify an invertible matrix $\rho(n)$ for every $n \in \mathbb{Z}$. However, given $\rho(1)$, we can recover $\rho(n)$ as $\rho(1 + \ldots + 1) = \rho(1)^n$. So there is no choice involved. Conversely, for any invertible map $\rho(1) \in \mathrm{GL}(m)$, we obtain a representation of $\mathbb{Z}$ this way.

Thus, m-dimensional isomorphism classes of representations of $\mathbb{Z}$ are in bijection with the conjugacy classes in $\mathrm{GL}(m)$. These can be parametrised by the *Jordan canonical form* (see the next example). We will have m continuous parameters — the eigenvalues, which are non-zero complex numbers, and are defined up to reordering — and some discrete parameters whenever two or more eigenvalues coincide, specifying the Jordan block sizes.

(2.5) Example: the cyclic group of order n

Let $G = \{1, g, \ldots, g^{n-1}\}$, with the relation $g^n = 1$. A representation of G on V defines an invertible endomorphism $\rho(g) \in \mathrm{GL}(V)$. As before, $\rho(1) = \mathrm{Id}$ and $\rho(g^k) = \rho(g)^k$, so all other images of ρ are determined by the single operator $\rho(g)$.

Choosing a basis of V allows us to convert $\rho(g)$ into a matrix A, but we shall want to be careful with our choice. Recall from general theory that there exists a *Jordan basis* in which $\rho(g)$ takes its block-diagonal *Jordan normal form*

$$A = \begin{bmatrix} J_1 & 0 & \ldots & 0 \\ 0 & J_2 & \ldots & 0 \\ & \ldots & \ldots & \\ 0 & 0 & \ldots & J_m \end{bmatrix}$$

where the *Jordan blocks* J_k take the form

$$J = \begin{bmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ & & \cdots & & \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda \end{bmatrix}.$$

However, we must impose the condition $A^n = \mathrm{Id}$. But A^n itself will be block-diagonal, with blocks J_k^n, so we must have $J_k^n = 1$. To compute that, let N be the Jordan matrix with $\lambda = 0$; we then have $J = \lambda\mathrm{Id} + N$, so

$$J^n = (\lambda\mathrm{Id} + N)^n = \lambda^n\mathrm{Id} + \binom{n}{1}\lambda^{n-1}N + \binom{n}{2}\lambda^{n-2}N^2 + \ldots;$$

but notice that N^p, for any p, is the matrix with zeroes and ones only, with the ones in index position (i, j) with $i = j + k$ (a line parallel to the diagonal, k steps above it). So the sum above can be Id only if $\lambda^n = 1$ and $N = 0$. In other words, J is a 1×1 block, and $\rho(g)$ is *diagonal* in this basis. We conclude the following

2.6 Proposition. *If V is a representation of the cyclic group G of order n, there exists a basis in which the action of every group element is diagonal, the with nth roots of unity on the diagonal.*

In particular, the m-dimensional representations of C_n are classified up to isomorphism by unordered m-tuples of nth roots of unity.

(2.7) Example: Finite abelian groups

The discussion for cyclic groups generalises to any finite Abelian group A. (The resulting classification of representations is more or less explicit, depending on whether we are willing to use the classification theorem for finite abelian groups; see below.) We recall the following fact from linear algebra:

2.8 Proposition. *Any family of commuting, separately diagonalisable $m \times m$ matrices can be simultaneously diagonalised.*

The proof is delegated to the example sheet; at any rate, an easier treatment of finite abelian groups will emerge from Schur's Lemma in Lecture 4.

This implies that any representation of A is isomorphic to one where every group element acts diagonally. Each diagonal entry then determines a *one-dimensional* representation of A. So the classification reads: m-dimensional isomorphism classes of representations of A are in bijection with unordered m-tuples of 1-dimensional representations. Note that for 1-dimensional representations, viewed as homomorphisms $\rho : A \to \mathbb{C}^\times$, there is no distinction between identity and isomorphism (the conjugation action of $\mathrm{GL}(1;\mathbb{C})$ on itself is trivial).

To say more, we must invoke the classification of finite abelian groups, according to which A is isomorphic to a direct product of cyclic groups. To specify a 1-dimensional representation of A we must then specify a root of unity of the appropriate order independently for each generator.

(2.9) Subrepresentations and Reducibility

Let $\rho : G \to \mathrm{GL}(V)$ be a representation of G.

2.10 Definition. A *subrepresentation* of V is a *G-invariant subspace* $W \subseteq V$; that is, we have

$$\forall \mathbf{w} \in W, g \in G \Rightarrow \rho(g)(\mathbf{w}) \in W.$$

W becomes a representation of G under the action $\rho(g)$.

Recall that, given a subspace $W \subseteq V$, we can form the *quotient space* V/W, the set of W-cosets $\mathbf{v} + W$ in V. If W was G-invariant, the G-action on V *descends to* (=defines) an action on V/W by setting $g(\mathbf{v} + W) := \rho(g)(\mathbf{v}) + W$. If we choose another $\mathbf{v}'$ in the same coset as $\mathbf{v}$, then $\mathbf{v} - \mathbf{v}' \in W$, so $\rho(g)(\mathbf{v} - \mathbf{v}') \in W$, and then the cosets $\rho(\mathbf{v}) + W$ and $\rho(\mathbf{v}') + W$ agree.

2.11 Definition. With this action, V/W is called the *quotient representation* of V under W.

2.12 Definition. The *direct sum* of two representations (ρ_1, V_1) and (ρ_2, V_2) is the space $V_1 \oplus V_2$ with the block-diagonal action $\rho_1 \oplus \rho_2$ of G.

(2.13) Example

In the direct sum $V_1 \oplus V_2$, V_1 is a sub-representation and V_2 is isomorphic to the associated quotient representation. Of course the roles of 1 and 2 can be interchanged. However, one should take care that for an *arbitrary* group, it need not be the case that any representation V with subrepresentation W decomposes as $W \oplus W/V$. This will be proved for complex representations of *finite* groups.

2.14 Definition. A representation is called *irreducible* if it contains no proper invariant subspaces. It is called *completely reducible* if it decomposes as a direct sum of irreducible subrepresentations.

In particular, irreducible representations are completely reducible.

For example, 1-dimensional representations of any group are irreducible. Earlier, we thus proved that finite-dimensional complex representations of a finite abelian group are completely reducible: indeed, we decomposed V into a direct sum of lines $L_1 \oplus \ldots \oplus L_{\dim V}$, along the vectors in the diagonal basis. Each line is preserved by the action of the group. In the cyclic case, the possible actions of C_n on a line correspond to the n eligible roots of unity to specify for $\rho(g)$.

2.15 Proposition. *Every complex representation of a finite abelian group is completely reducible, and every irreducible representation is 1-dimensional.*

It will be our goal to establish an analogous proposition for every finite group G. The result is called the *Complete Reducibility Theorem.* For non-abelian groups, we shall have to give up on the 1-dimensional requirement, but we shall still salvage a canonical decomposition.

3 Complete Reducibility and Unitarity

In the homework, you find an example of a complex representation of the group $\mathbb{Z}$ which is *not* completely reducible, and also of a representation of the cyclic group of prime order p over the finite field $\mathbb{F}_p$ which is not completely reducible. This underlines the importance of the following *Complete Reducibility Theorem* for finite groups.

3.1 Theorem. *Every complex representation of a finite group is completely reducible.*

The theorem is so important that we shall give two proofs. The first uses inner products, and so applies only to $\mathbb{R}$ or $\mathbb{C}$, but generalises to *compact* groups. The more algebraic proof, on the other hand, extends to any fields of scalars *whose characteristic does not divide the order of the group* (equivalently, the order of the group should not be 0 in the field).

Beautiful as it is, the result would have limited value without some supply of irreducible representations. It turns out that the following example provides an adequate supply.

(3.2) Example: the regular representation

Let $\mathbb{C}[G]$ be the vector space of complex functions on G. It has a basis $\{\mathbf{e}_g\}_{g\in G}$, with $\mathbf{e}_g$ representing the function equal to 1 at g and 0 elsewhere. G acts on this basis as follows:

$$\lambda(g)(\mathbf{e}_h) = \mathbf{e}_{gh}.$$

This set-theoretic action extends by linearity to the vector space:

$$\lambda(g)\left(\textstyle\sum_{h\in G} v_h \cdot \mathbf{e}_h\right) = \textstyle\sum_{h\in G} v_h \cdot \lambda(g)\mathbf{e}_h = \textstyle\sum_{h\in G} v_h \cdot \mathbf{e}_{gh}.$$

(Exercise: check that this defines a linear action.) On coordinates, the action is opposite to what you might expect: namely, the h-coordinate of $\lambda(g)(\mathbf{v})$ is $v_{g^{-1}h}$. The result is the *left regular representation* of G. Later we will decompose λ into irreducibles, and we shall see that *every* irreducible isomorphism class of G-reps occurs in the decomposition.

3.3 Remark. If G acts on a set X, let $\mathbb{C}[X]$ be the vector space of functions on X, with obvious basis $\{\mathbf{e}_x\}_{x\in X}$. By linear extension of the permutation action $\rho(g)(\mathbf{e}_x) = \mathbf{e}_{gx}$, we get a linear action of G on $\mathbb{C}[X]$; this is the *permutation representation* associated to X.

(3.4) Unitarity

Inner products are an important aid in investigating real or complex representations, and lead to a first proof of Theorem 3.1.

3.5 Definition. A representation ρ of G on a complex vector space V is *unitary* if V has been equipped with a hermitian inner product $\langle\,|\,\rangle$ which is preserved by the action of G, that is,

$$\langle \mathbf{v}|\mathbf{w}\rangle = \langle \rho(g)(\mathbf{v})|\rho(g)(\mathbf{w})\rangle, \quad \forall \mathbf{v},\mathbf{w}\in V,\ g\in G.$$

It is *unitarisable* if it can be equipped with such a product (even if none has been chosen).

For example, the regular representation of a finite group is unitarisable: it is made unitary by declaring the standard basis vectors $\mathbf{e}_g$ to be orthonormal.

The representation is unitary iff the homomorphism $\rho : G \to \mathrm{GL}(V)$ lands inside the *unitary group* $\mathrm{U}(V)$ (defined with respect to the inner product). We can restate this condition in the form $\rho(g)^* = \rho(g^{-1})$. (The latter is also $\rho(g)^{-1}$).

3.6 Theorem (Unitary Criterion). *Finite-dimensional unitary representations of any group are completely reducible.*

The proof relies on the following

3.7 Lemma. *Let V be a unitary representation of G and let W be an invariant subspace. Then, the orthocomplement $W^{\perp}$ is also G-invariant.*

Proof. We must show that, $\forall \mathbf{v} \in W^{\perp}$ and $\forall g \in G$, $g\mathbf{v} \in W^{\perp}$. Now, $\mathbf{v} \in W^{\perp} \Leftrightarrow \langle \mathbf{v}|\mathbf{w}\rangle = 0$, $\forall \mathbf{w} \in W$. If so, then $\langle g\mathbf{v}|g\mathbf{w}\rangle = 0$, for any $g \in G$ and $\forall \mathbf{w} \in W$. This implies that $\langle g\mathbf{v}|\mathbf{w}'\rangle = 0$, $\forall \mathbf{w}' \in W$, since we can choose $\mathbf{w} = g^{-1}\mathbf{w}' \in W$, using the invariance of W. But then, $g\mathbf{v} \in W^{\perp}$, and $g \in G$ was arbitrary. $\square$

We are now ready to prove the following stronger version of the unitary criterion (3.6).

3.8 Theorem. *A finite-dimensional unitary representation of a group admits an orthogonal decomposition into irreducible unitary sub-representations.*

Proof. Clearly, any sub-representation is unitary for the restricted inner product, so we must merely produce the decomposition into irreducibles. Assume that V is not irreducible; it then contains a proper invariant subspace $W \subsetneq V$. By Lemma 3.7, $W^{\perp}$ is another sub-representation, and we clearly have an orthogonal decomposition $V = W \oplus W^{\perp}$, which is G-invariant. Continuing with W and $W^{\perp}$ must terminate in an irreducible decomposition, by finite-dimensionality. $\square$

3.9 Remark. The assumption $\dim V < \infty$ is in fact unnecessary for *finite* G, but cannot be removed without changes to the statement, for general G. For representations of infinite, non-compact groups on infinite-dimensional (Hilbert) spaces, the most we can usually hope for is a decomposition into a *direct integral* of irreducibles. For example, this happens for the translation action of the group $\mathbb{R}$ on the Hilbert space $L^2(\mathbb{R})$. The irreducible "components" are the *Fourier modes* $\exp(\mathrm{i}kx)$, labelled by $k \in \mathbb{R}$; note that they are not quite in L^2. Nonetheless, there results an integral decomposition of any vector $f(x) \in L^2(\mathbb{R})$ into Fourier modes,

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(k)\exp(-\mathrm{i}k \cdot x)dk,$$

known to you as the *Fourier inversion formula*; $\hat{f}(k)/2\pi$ should be regarded as the coordinate value of f along the "basis vector" $\exp(-\mathrm{i}kx)$. Even this milder expectation of integral decomposition can fail for more general groups, and leads to delicate and difficult problems of functional analysis (von Neumann algebras).

3.10 Theorem (Weyl's unitary trick). *Finite-dimensional representations of finite groups are unitarisable.*

Proof. Starting with a hermitian, positive definite inner product $\langle\,|\,\rangle$ on your representation, construct an new one $\langle\,|\,\rangle'$ by *averaging over* G,

$$\langle \mathbf{v}|\mathbf{w}\rangle' := \frac{1}{|G|}\sum\nolimits_{g\in G}\langle g\mathbf{v}|g\mathbf{w}\rangle.$$

Check invariance and positivity (homework). $\square$

3.11 Remark. Weyl's unitary trick applies to continuous representations of *compact* groups. In that case, we use integration over G to average. The key point is the existence of a measure on G which is invariant for the translation action of the group on itself (the Haar measure). For groups of geometric origin, such as U(1) (and even U(n) or SU(n)) the existence of such a measure is obvious, but in general it is a difficult theorem.

The complete reducibility theorem (3.1) for finite groups follows from Theorem 3.8 and Weyl's unitary trick.

(3.12) Alternative proof of Complete Reducibility

The following argument makes no appeal to inner products, and so has the advantage of working over more general ground fields.

3.13 Lemma. *Let V be a finite-dimensional representation of the finite group G over a field of characteristic not dividing $|G|$. Then, every invariant subspace $W \subseteq V$ has an invariant complement $W' \subseteq V$.*

Recall that the subspace W' is a *complement* of W if $W \oplus W' = V$.

Proof. The complementing condition can be broken down into two parts,

- $W \cap W' = \{0\}$
- $\dim W + \dim W' = \dim V$.

We'd like to construct the complement using the same trick as with the inner product, by "averaging a complement over G" to produce an invariant complement. While we cannot average subspaces, note that any complement of W is completely determined as the *kernel of the projection operator* $P : V \to W$, which sends a vector $\mathbf{v} \in V$ to its W-component $P(\mathbf{v})$. Choose now an arbitrary complement $W" \subseteq V$ (not necessarily invariant) and call P the projection operator. It satisfies

- $P = \mathrm{Id}$ on W
- $\mathrm{Im} P = W$.

Now let $Q := \frac{1}{|G|} \sum_{g \in G} g \circ P \circ g^{-1}$. I claim that:

- Q commutes with G, that is, $Q \circ h = h \circ Q$, $\forall h \in G$;
- $Q = \mathrm{Id}$ on W;
- $\mathrm{Im} Q = W$.

The first condition implies that $W' := \ker Q$ is a G-invariant subspace. Indeed, $\mathbf{v} \in \ker Q \Rightarrow Q(\mathbf{v}) = 0$, hence $h \circ Q(\mathbf{v}) = 0$, hence $Q(h\mathbf{v}) = 0$ and $h\mathbf{v} \in \ker Q$, $\forall h \in G$. The second condition implies that $W' \cap W = \{0\}$, while the third implies that $\dim W' + \dim W = \dim V$; so we have constructed our invariant complement. □

4 Schur's Lemma

To understand representations, we also need to understand the *automorphisms* of a representation. These are the invertible self-maps commuting with the G-action. To see the issue, assume given a representation of G, and say you have performed a construction within it (such as a drawing of your favourite Disney character; more commonly, it will be a distinguished geometric object). If someone presents you with another representation, claimed to be isomorphic to yours, how uniquely can you repeat your construction in the other copy of the representation? Schur's Lemma answers this for irreducible representations.

4.1 Theorem (Schur's lemma over $\mathbb{C}$). *If V is an irreducible complex G-representation, then every linear operator $\phi : V \to V$ commuting with G is a scalar.*

Proof. Let λ be an eigenvalue of ϕ. I claim that the eigenspace E_λ is G-invariant. Indeed, $\mathbf{v} \in E_\lambda \Rightarrow \phi(\mathbf{v}) = \lambda\mathbf{v}$, whence

$$\phi(g\mathbf{v}) = g\phi(\mathbf{v}) = g(\lambda\mathbf{v}) = \lambda \cdot g\mathbf{v},$$

so $g\mathbf{v} \in E_\lambda$, and g was arbitrary. But then, $E_\lambda = V$ by irreducibility, so $\phi = \lambda\mathrm{Id}$. □

Given two representations V and W, we denote by $\mathrm{Hom}^G(V, W)$ the vector space of *intertwiners* from V to W, meaning the linear operators which commute with the G-action.

4.2 Corollary. *If V and W are irreducible, the space $\mathrm{Hom}^G(V, W)$ is either 1-dimensional or $\{0\}$, depending on whether or not the representations are isomorphic. In the first case, any non-zero map is an isomorphism.*

Proof. Indeed, the kernel and the image of an intertwiner ϕ are invariant subspaces of V and W, respectively (proof as for the eigenspaces above). Irreducibility leaves $\ker\phi = 0$ or V and $\Im\phi = 0$ or W as the only options. So if ϕ is not injective, $\ker\phi = V$ and $\phi = 0$. If ϕ is injective and $V \neq 0$, then $\Im\phi = W$ and so ϕ is an isomorphism. Finally, to see that two intertwiners ϕ, ψ differ by a scalar factor, apply Schur's lemma to $\phi^{-1} \circ \psi$. □

(4.3) Schur's Lemma over other fields

The correct statement over other fields (even over $\mathbb{R}$) requires some preliminary definitions.

4.4 Definition. An *algebra* over a field $\mathbf{k}$ is an associative ring with unit, containing a distinguished copy of $\mathbf{k}$, commuting with every algebra element, and with $1 \in \mathbf{k}$ being the algebra unit. A *division ring* is a ring where every non-zero element is invertible, and a *division algebra* is a division ring which is also a $\mathbf{k}$-algebra.

4.5 Definition. Let V be a G-representation over $\mathbf{k}$. The *endomorphism algebra* $\mathrm{End}^G(V)$ is the space of linear self-maps $\phi : V \to V$ which commute with the group action, $\rho(g)\circ\phi = \phi\circ\rho(g)$, $\forall g \in G$. The addition is the usual addition of linear maps, and the multiplication is composition.

I entrust it to your care to check that $\mathrm{End}^G(V)$ is indeed a $\mathbf{k}$-algebra. In earlier notation, $\mathrm{End}^G(V) = \mathrm{Hom}^G(V, V)$; however, $\mathrm{Hom}^G(V, W)$ is only a vector space in general.

4.6 Theorem (Schur's Lemma). *If V is an irreducible finite-dimensional G-representation over $\mathbf{k}$, then $\mathrm{End}^G(V)$ is a finite-dimensional division algebra over $\mathbf{k}$.*

Proof. For any $\phi : V \to V$ commuting with G, $\ker\phi$ and $\mathrm{Im}\phi$ are G-invariant subspaces. Irreducibility implies that either $\ker\phi = \{0\}$, so ϕ is injective and then an isomorphism (for dimensional reasons, or because $\mathrm{Im}\phi = V$), or else $\ker\phi = V$ and then $\phi = 0$. □

A second proof of Schur's Lemma over $\mathbb{C}$ can now be deduced from the following.

4.7 Proposition. *The only finite-dimensional division algebra over $\mathbb{C}$ is $\mathbb{C}$ itself.*

Proof. Let, indeed, A be the algebra and $\alpha \in A$. The elements $1, \alpha, \alpha^2, \ldots$ of the finite-dimensional vector space A over $\mathbb{C}$ must be linearly dependent. So we have a relation $P(\alpha) = 0$ in A, for some polynomial P with complex coefficients. But such a polynomial factors into linear terms $\prod_k (\alpha - \alpha_k)$, for some complex numbers α_k, the roots of P. Since we are in a division ring and the product is zero, one of the factors must be zero, therefore $\alpha \in \mathbb{C}$ and so $A = \mathbb{C}$. □

4.8 Remark. Over the reals, there are three finite-dimensional division algebras, namely $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$. All of these can occur as endomorphism algebras of irreducibles of finite groups (see Example Sheet 1).

(4.9) Application: Finite abelian groups revisited
Schur's Lemma gives a shorter proof of the 1-dimensionality of irreps for finite abelian groups.

4.10 Theorem. *Any irreducible complex representation of an abelian group G is one-dimensional.*

Proof. Let $g \in G$ and call ρ the representation. Then, $\rho(g)$ commutes with every ρh. By irreducibility and Schur's lemma, it follows that $\rho(g)$ is a scalar. As each group element acts by a scalar, every line in V is invariant, so irreducibility implies that V itself is a line. $\square$

As one might expect from the failure of Schur's lemma, this proposition is utterly false over other ground fields. For example, over the reals, the two irreducible representations of $\mathbb{Z}/3$ have dimensions 1 and 2. (They are the trivial and the rotation representations, respectively). However, one can show in general that the irreps are 1-dimensional over their endomorphism ring, which is an extension field of $\mathbf{k}$.

The structure theorem asserts that every finite abelian G splits as $\prod_k C_{n_k}$, for cyclic groups C_{n_k} of orders n_k. Choose a generator g_k in each factor; this must act on any line by a root of unity of order n_k. So, to specify a one-dimensional representation, one must choose such a root of unity for each k. In particular, *the number of irreducible isomorphism classes of representations of a finite abelian group equals the order of the group.* Nonetheless, there is no *canonical* correspondence between group elements and representations; one must choose generators in each cyclic factor for that.

4.11 Remark. The equality above generalises to non-abelian finite groups, but not as naively as one might think. One correct statement is that the number of irreducible isomorphism classes equals the number of conjugacy classes in the group. Another generalisation asserts that the squared dimensions of the irreducible isomorphism classes add up to the order of the group.

5 Isotypical Decomposition

To motivate the result we prove today, recall that any diagonalisable linear endomorphism $A : V \to V$ leads to an *eigenspace decomposition* of the space V as $\bigoplus_\lambda V(\lambda)$, where $V(\lambda)$ denotes the subspace of vectors verifying $A\mathbf{v} = \lambda\mathbf{v}$. The decomposition is *canonical*, in the sense that it depends on A alone and no other choices. (By contrast, there is no canonical *eigenbasis* of V: we must choose a basis in each $V(\lambda)$ for that.) To show how useful this is, we now classify the irreducible representations of D_6 "by hand", also proving complete reducibility in the process.

(5.1) Example: Representations of D_6
The symmetric group S_3 on three letters, also isomorphic to the dihedral group D_6, has two generators g, r satisfying $g^3 = r^2 = 1$, $gr = rg^{-1}$. It is easy to spot three irreducible representations:

- The trivial 1-dimensional representation $\mathbf{1}$
- The sign representation S
- The geometric 2-dimensional representation W.

All these representations are naturally defined on real vector spaces, but we wish to work with the complex numbers so we view them as complex vector spaces instead. For instance, the 2-dimensional rep. is defined by $W = \mathbb{C}^2$ by letting g act on $W = \mathbb{C}^2$ as the diagonal matrix $\rho(g) = \operatorname{diag}\left[\omega, \omega^2\right]$, and r as the matrix $\rho(r) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$; here, $\omega = \exp\left(\frac{2\pi i}{3}\right)$ is the basic cube root of unity. (*Exercise:* Relate this to the geometric action of D_6. You must change bases, to diagonalise g.)

Let now (π, V) be any complex representation of D_6. We can diagonalise the action of $\pi(g)$ and split V into eigenspaces,

$$V = V(1) \oplus V(\omega) \oplus V(\omega^2).$$

The relation $rgr^{-1} = g^{-1}$ shows that $\pi(r)$ must preserve $V(1)$ and interchange the other summands: that is, $\pi(r) : V(\omega) \to V(\omega^2)$ is an isomorphism, with inverse $\pi(r)$ (since $r^2 = 1$). Decompose $V(1)$ into the $\pi(r)$-eigenspaces; the eigenvalues are ± 1, and since g acts trivially it follows that $V(1)$ splits into a sum of copies of $\mathbf{1}$ and S. Further, choose a basis $\mathbf{e}_1, \dots \mathbf{e}_n$ of $V(\omega)$, and let $\mathbf{e}'_k = \pi(r)\mathbf{e}_k$. Then, $\pi(r)$ acts on the 2-dimensional space $\mathbb{C}\mathbf{e}_k \oplus \mathbb{C}\mathbf{e}'_k$ as the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, while $\pi(g)$ acts as diag $\left[\omega, \omega^2\right]$. It follows that we have split $V(\omega) \oplus V(\omega^2)$ into n copies of the 2-dimensional representation W.

(5.2) Decomposition into isotypical components

Continuing the general theory, the next task is to understand completely reducible representations in terms of their irreducible constituents. A completely reducible V decomposes into a sum of irreducibles. Grouping isomorphic irreducible summands into blocks, we write $V \cong \bigoplus_k V_k$, where each V_k is isomorphic to a sum of n_k copies of the irreducible representation W_k; the latter are assumed non-isomorphic for different k's. Another notation we might employ is

$$V \cong \bigoplus\nolimits_k W_k^{\oplus n_k}.$$

We call n_k the *multiplicity of* W_k *in* V, and the decomposition of V just described the *canonical decomposition* of V, or *decomposition into isotypical components* V_k. This terminology must now be justified: we do not know yet whether this decomposition is unique, and not even whether the n_k are unambiguously defined.

5.3 Lemma. *Let $V = \bigoplus_k V_k$, $V' = \bigoplus_k V'_k$ be canonical decompositions of two representations V and V'. Then, any G-invariant homomorphism $\phi : V' \to V$ maps V'_k to V_k.*

Note that, for any completely reducible V and V', we can arrange the decompositions to be labelled by the same indexes, adding zero blocks if some W_k is not present in the decomposition. Obviously, ϕ will be zero on those. We are not assuming uniqueness of the decompositions; rather, this will follow from the Lemma.

Proof. We consider the "block decomposition" of ϕ with respect to the direct sum decompositions of V and V'. Specifically, restricting ϕ to the summand V'_l and projecting to V_k leads to the "block" $\phi_{kl} : V'_l \to V_k$, which is a linear G-map. Consider now an *off-diagonal block* ϕ_{kl}, $k \neq l$. If we decompose V_k and V'_l further into copies of irreducibles W_k and W_l, the resulting blocks of ϕ_{kl} give G-linear maps between non-isomorphic irreducibles, and such maps are all null by Corollary 4.2. So $\phi_{kl} = 0$ if $k \neq l$, as asserted. □

5.4 Theorem. *Let $V = \bigoplus V_k$ be a canonical decomposition of V.*

(i) Every sub-representation of V which is isomorphic to W_k is contained in V_k.
(ii) The canonical decomposition is unique*, that is, it does not depend on the original decomposition of V into irreducibles.*
(iii) The endomorphism algebra $\mathrm{End}^G(V_k)$ *is isomorphic to a matrix algebra $M_{n_k}(\mathbb{C})$, a choice of isomorphism coming from a decomposition of V_k into a direct sum of the irreducible W_k.*
(iv) $\mathrm{End}^G(V)$ *is isomorphic to the direct sum of matrix algebras $\bigoplus_k M_{n_k}(\mathbb{C})$, block-diagonal with respect to the canonical decomposition of V.*

Proof. Part (i) follows from the Lemma by taking $V' = W_k$. This gives an intrinsic description of V_k, as the sum of *all* copies of W_k contained in V, leading to part (ii).

Let $\phi \in \mathrm{End}^G(V_k)$ write V_k as a direct sum $\cong W_k^{\oplus n_k}$ of copies of W_k. In the resulting block-decomposition of ϕ, each block ϕ_{pq} is a G-invariant linear map between copies of W_k. By Schur's lemma, all blocks are scalar multiples of the identity: $\phi_{pq} = f_{pq} \cdot \mathrm{Id}$, for some constant f_{pq}. Assigning to ϕ the matrix $[f_{pq}]$ identifies $\mathrm{End}^G(V_k)$ with $M_{n_k}(\mathbb{C})$.[3] Part (iv) follows from (iii) and the Lemma. □

5.5 Remark. When the ground field is not $\mathbb{C}$, we must allow the appropriate division rings to replace $\mathbb{C}$ in the theorem.

(5.6) Operations on representations

We now discuss some operations on group representations; they can be used to enlarge the supply of examples, once a single interesting representation of a group is known. We have already met the direct sum of two representations in Lecture 2; however, that cannot be used to produce new irreducibles.

5.7 Definition. The *dual representation* of V is the representation ρ^* on the dual vector space $V^* := \mathrm{Hom}(V;\mathbb{C})$ of linear maps $V \to \mathbb{C}$ defined by $\rho^*(g)(L) = L \circ \rho(g)^{-1}$.

That is, a linear map $L : V \to \mathbb{C}$ is mapped to the linear map $\rho^*(g)(L)$ sending $\mathbf{v}$ to $L\left(\rho(g)^{-1}(\mathbf{v})\right)$. Note how this definition preserves the *duality pairing* between V and V^*,

$$L(\mathbf{v}) = \rho^*(g)(L)\left(\rho(g)(\mathbf{v})\right).$$

There is actually little choice in the matter, because the more naive option $\rho^*(g)(L) = L \circ \rho(g)$ does *not*, in fact, determine an action of G. The following is left as an exercise.

5.8 Proposition. *The dual representation ρ^* is irreducible if and only if ρ was so.*

Exercise. Study duality on irreducible representations of abelian groups.

The dual representation is a special instance of the Hom space of two representations: we can replace the 1-dimensional space of scalars $\mathbb{C}$, the target space of our linear map, by an arbitrary representation W.

5.9 Definition. For two vector spaces V, W, $\mathrm{Hom}(V, W)$ will denote the vector space of linear maps from V to W. If V and W carry linear G-actions ρ and σ, then $\mathrm{Hom}(V, W)$ carries a natural G-action in which $g \in G$ sends $\phi : V \to W$ to $\sigma(g) \circ \phi \circ \rho(g)^{-1}$.

In the special case $W = \mathbb{C}$ with the trivial G-action, we recover the construction of V^*. Note also that the *invariant vectors* in the Hom space are precisely the *G-linear maps* from V to W. When $V = W$, this Hom space and its invariant part are algebras, but of course not in general.

6 Tensor products

We now construct the *tensor product* of two vector spaces V and W. For simplicity, we will take them to be finite-dimensional. We give two definitions; the first is more abstract, and contains all the information you need to work with tensor products, but gives little idea of the true size of the resulting space. The second is quite concrete, but requires a choice of basis, and in this way breaks any symmetry we might wish to exploit in the context of group actions.

[3] A different decomposition of V_k is related to the old one by a G-isomorphism $S : W_k^{\oplus n_k} \to W_k^{\oplus n_k}$; but this G-isomorphism is itself block-decomposable into scalar multiples of Id, so the effect on ϕ is simply to conjugate the associated matrix $[f_{pq}]$ by S.

6.1 Definition. The *tensor product* $V \otimes_{\mathbf{k}} W$ of two vector spaces V and W over $\mathbf{k}$ is the $\mathbf{k}$-vector space based on elements $\mathbf{v} \otimes \mathbf{w}$, labelled by pairs of vectors $\mathbf{v} \in V$ and $\mathbf{w} \in W$, modulo the following relations, for all $k \in \mathbb{C}$, $\mathbf{v} \in V$, $\mathbf{w} \in W$:

$$\begin{aligned}(\mathbf{v}_1 + \mathbf{v}_2) \otimes \mathbf{w} &= \mathbf{v}_1 \otimes \mathbf{w} + \mathbf{v}_2 \otimes \mathbf{w};\\ \mathbf{v} \otimes (\mathbf{w}_1 + \mathbf{w}_2) &= \mathbf{v} \otimes \mathbf{w}_1 + \mathbf{v} \otimes \mathbf{w}_2;\\ (k \cdot \mathbf{v}) \otimes \mathbf{w} &= \mathbf{v} \otimes (k \cdot \mathbf{w}) = k \cdot (\mathbf{v} \otimes \mathbf{w}).\end{aligned}$$

In other words, $V \otimes W$ is the quotient of the vector space with basis $\{\mathbf{v} \otimes \mathbf{w}\}$ by the subspace spanned by the differences of left– and right-hand sides in each identity above. When the ground field is understood, it can be omitted from the notation.

6.2 Proposition. *Let* $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ *and* $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ *be bases of* V *and* W*. Then,* $\{\mathbf{e}_p \otimes \mathbf{f}_q\}$*, with* $1 \leq p \leq m$*,* $1 \leq q \leq n$*, is a vector space basis of* $V \otimes W$*.*

Proof. First, we show that the vectors $\mathbf{e}_p \otimes \mathbf{f}_q$ *span* the tensor product $V \otimes W$. Indeed, for any $\mathbf{v} \otimes \mathbf{w}$, we can express the factors $\mathbf{v}$ and $\mathbf{w}$ linearly in terms of the basis elements. The tensor relations then express $\mathbf{v} \otimes \mathbf{w}$ as a linear combination of the $\mathbf{e}_p \otimes \mathbf{f}_q$. To check linear independence, we construct, for each pair (p, q), a linear functional $L : V \otimes W \to \mathbf{k}$ equal to 1 on $\mathbf{e}_p \otimes \mathbf{f}_q$ and vanishing on all other $\mathbf{e}_r \otimes \mathbf{f}_s$. This shows that no linear relation can hold among our vectors.

Let $\varepsilon : V \to \mathbf{k}$ and $\varphi : W \to \mathbf{k}$ be the linear maps which are equal to 1 on $\mathbf{e}_p$, resp. $\mathbf{f}_q$ and vanish on the other basis vectors. We let $L(\mathbf{v} \otimes \mathbf{w}) = \varepsilon(\mathbf{v}) \cdot \varphi(\mathbf{w})$. We note by inspection that L takes equal values on the two sides of each of the identities in (6.1), and so it descends to a well- defined linear map on the quotient space $V \otimes W$. This completes the proof. □

The tensor product behaves like a multiplication with respect to the direct sum:

6.3 Proposition. *There are natural isomorphisms* $U \otimes (V \otimes W) \cong (U \otimes V) \otimes W$*,* $(U \oplus V) \otimes W \cong (U \otimes W) \oplus (V \otimes W)$

Proof. "Naturality" means that no choice of basis is required to produce an isomorphism. Indeed, send $\mathbf{u} \otimes (\mathbf{v} \otimes \mathbf{w})$ to $(\mathbf{u} \otimes \mathbf{v}) \otimes \mathbf{w}$ for the first map, and $(\mathbf{u}, \mathbf{v}) \otimes \mathbf{w}$ to $(\mathbf{u} \otimes \mathbf{w}, \mathbf{v} \otimes \mathbf{w})$ in the second, extend by linearity to the spaces and check that the tensor relations (6.1) are preserved by the maps. □

6.4 Definition. The tensor product of two representations ρ on V and σ on W of a group G is the representation $\rho \otimes \sigma$ on $V \otimes W$ defined by the condition

$$(\rho \otimes \sigma)(g)(\mathbf{v} \otimes \mathbf{w}) = \rho(g)(\mathbf{v}) \otimes \sigma(g)(\mathbf{w}),$$

and extended to all vectors in $V \otimes W$ by linearity.

By inspection, this construction preserves the linear relations in Def. 6.1, so it does indeed lead to a linear action of G on the tensor product. We shall write this action in matrix form, using a basis as in Proposition 6.2, but first let us note a connection between these constructions.

6.5 Proposition. *If* $\dim V, \dim W < \infty$*, there is a natural isomorphism of vector spaces (preserving G-actions, if defined) from* $W \otimes V^*$ *to* $\mathrm{Hom}(V, W)$*.*

Proof. There is a natural map from the first space to the second, defined by sending $\mathbf{w} \otimes L \in W \otimes V^*$ to the rank one linear map $\mathbf{v} \mapsto \mathbf{w} \cdot L(\mathbf{v})$. We extend this to all of $W \otimes V^*$ by linearity. Clearly, this preserves the G-actions, as defined earlier, so we just need to check the map is an isomorphism. In the basis $\{\mathbf{f}_i\}$ of W, the dual basis $\{\mathbf{e}_j^*\}$ of V^* and the basis $\{\mathbf{f}_i \otimes \mathbf{e}_j^*\}$ of $W \otimes V^*$ constructed in Def. 6.1.ii, we can check that the basis vector $\mathbf{f}_i \otimes \mathbf{e}_j^*$ is sent to the elementary matrix E_{ij}; and the latter form a basis of all linear maps from V to W. □

(6.6) The tensor product of two linear maps

Consider two linear operators $A \in \mathrm{End}(V)$ and $B \in \mathrm{End}(W)$. We define an operator $A \otimes B \in \mathrm{End}(V \otimes W)$ by $(A \otimes B)(\mathbf{v} \otimes \mathbf{w}) = (A\mathbf{v}) \otimes (B\mathbf{w})$, extending by linearity to all vectors in $V \otimes W$. In bases $\{\mathbf{e}_i\}, \{\mathbf{f}_k\}$ of V, W, our operators are given by matrices A_{ij}, $1 \le i, j \le m$, and B_{kl}, $1 \le k, l \le n$. Let us write the matrix form of $A \otimes B$ in the basis $\{\mathbf{e}_i \otimes \mathbf{f}_k\}$ of $V \otimes W$. Note that a basis index for $V \otimes W$ is a pair (i, k) with $1 \le i \le m$, $1 \le k \le n$. We have the simple formula

$$(A \otimes B)_{(i,k)(j,l)} = A_{ij} \cdot B_{kl},$$

which is checked by observing that the application of $A \otimes B$ to the basis vector $\mathbf{e}_j \otimes \mathbf{f}_l$ contains $\mathbf{e}_i \otimes \mathbf{f}_k$ with the advertised coefficient $A_{ij} \cdot B_{kl}$.

Here are some useful properties of $A \otimes B$.

6.7 Proposition. *(i) The eigenvalues of $A \otimes B$ are $\lambda_i \cdot \mu_k$.*
(ii) $\mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A)\mathrm{Tr}(B)$.
(iii) $\det(A \otimes B) = \det(A)^n \det(B)^m$.

Proof. $\mathrm{Tr}(A \otimes B) = \sum_{i,k}(A \otimes B)_{(i,k)(i,k)} = \sum_{i,k} A_{ii}B_{kk} = \sum_i A_{ii} \cdot \sum_k B_{kk} = \mathrm{Tr}(A) \cdot \mathrm{Tr}(B)$, proving part (ii). When A and B are diagonalisable, parts (i) and (iii) are seen as follows. Choose the $\{\mathbf{e}_i\}, \{\mathbf{f}_j\}$ to be eigenbases for A, B. Then, $A \otimes B$ is also diagonal, with entries $\lambda_i \cdot \mu_k$. This deals with the generic case, because generic matrices are diagonalisable. For any matrix, we can choose a diagonalisable matrix arbitrarily close-by, and the formulae (i) and (iii) apply. If you dislike this argument, choose instead bases in which A and B are upper-triangular (this can always be done), and check that $A \otimes B$ is also upper-triangular in the tensor basis. □

6.8 Remark. A concrete proof of (iii) can be given using row-reduction. Applied to A and B, this results in upper-triangular matrices S and T, whose diagonal entries multiply to give the two determinants. Now, ordering the index pairs (i, k) lexicographically, the matrix $A \otimes B$ acquires a decomposition into blocks of size $n \times n$

$$\begin{bmatrix} A_{11} \cdot B & A_{12} \cdot B & \cdots & A_{1m} \cdot B \\ A_{21} \cdot B & \cdots & \cdots & A_{2m} \cdot B \\ & \cdots & \cdots & \\ A_{m1} \cdot B & A_{m2} \cdot B & \cdots & A_{mm} \cdot B \end{bmatrix}$$

We can apply the row-reduction procedure for A treating the blocks in this matrix as entries; this results in the matrix $[S_{ij} \cdot B]$. Applying now the row-reduction procedure for B to each block of n consecutive rows results in the matrix $[S_{ij} \cdot T]$, which is $S \otimes T$. Its determinant is $\prod_{i,k} S_{ii}T_{kk} = (\det S)^n \cdot (\det T)^m$. The statement about eigenvalues follows from the determinant formula by considering the characteristic polynomial.

7 Examples and complements

The formula for $A \otimes B$ permits us to write the matrix form of the tensor product of two representations, in the basis (6.2). However, the matrices involved are often too large to give any insight. In the following example, we take a closer look at the *tensor square* $V^{\otimes 2} := W \otimes W$ of the irreducible 2-dimensional representation of D_6; we will see that it decomposes into the three distinct irreducibles listed in Lecture 5.

(7.1) Example: Decomposing the tensor square of a representation

Let $G = D_6$. It has the three irreps $\mathbf{1}, S, W$ listed in Lecture 5. It is easy to check that $\mathbf{1}\otimes S = S$, $S \otimes S = \mathbf{1}$. To study $W \otimes W$, we could use the matrix to write the 4×4 matrices of $\rho \otimes \rho$, but there is a better way to understand what happens.

Under the action of g, W decomposes as a sum $L_1 \oplus L_2$ of eigen-lines. We also call L_0 the line with trivial action of the group C_3 generated by g. We get a C_3-isomorphism

$$W \otimes W = (L_1 \otimes L_1) \oplus (L_2 \otimes L_2) \oplus (L_1 \otimes L_2) \oplus (L_2 \otimes L_1).$$

Now, as C_3-reps, $L_1 \otimes L_1 \cong L_2$, $L_2 \otimes L_2 \cong L_1$, $L_1 \otimes L_2 \cong L_0$, so our decomposition is isomorphic to $L_2 \oplus L_1 \oplus L_0 \oplus L_0$. We can also say something about r: since it swaps L_1 and L_2 in each factor of the tensor product, we see that, in the last decomposition, it must swap L_1 with L_2 and also the two factors of L_0 among themselves. We suspect that the first two lines add up to an isomorphic copy of W; to confirm this, choose a basis vector $\mathbf{e}_1$ in L_1, and let $\mathbf{e}_2 = \rho^{\otimes 2}(\mathbf{e}_1) \in L_2$. Because $r^2 = 1$, we see that in this basis of $L_2 \oplus L_1$, G acts by the same matrices as on W.

On the last two summands, G acts trivially, but r swaps the two lines. Choosing a basis of the form $\{\mathbf{f}, \rho^{\otimes 2}(\mathbf{f})\}$ leads again to the matrix expression $\rho^{\otimes 2}(r) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$; this diagonalises to $\operatorname{diag}[1, -1]$. So we recognise here the sum of the trivial representation and the *sign representation* S of S_3. All in all, we have

$$W \otimes W \cong W \oplus L_0 \oplus S.$$

The appearance of the 'new' irreducible sign representation shows that there cannot be a straightforward universal formula for the decomposition of tensor products of irreducibles. One substantial accomplishment of the *theory of characters*, which we will study next, is a simple calculus which permits the decomposition of representations into irreducibles.

For your amusement, here is the matrix form of $\rho \otimes \rho$, in a basis matching the decomposition into lines given earlier. (Note: this is not the lexicographically ordered basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$.)

$$\rho^{\otimes 2}(g) = \begin{bmatrix} \omega^2 & 0 & 0 & 0 \\ 0 & \omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \qquad \rho^{\otimes 2}(r) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

(7.2) The symmetric group action on tensor powers

There was no possibility of $W \otimes W$ being irreducible: switching the two factors in the product gives an action of C_2 which commutes with D_6, and the ± 1 eigenspaces of the action must then be invariant under D_6. The $+1$-eigenspace, the part of $W \otimes W$ invariant under the switch, is called the *symmetric square*, and the (-1)-part is the *exterior square*. These are special instances of the following more general decomoposition.

Let V be a representation of G, $\rho : G \to \mathrm{GL}(V)$, and let $V^{\otimes n} := V \otimes \ldots \otimes V$ (n times). If $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a basis of V, then a basis of $V^{\otimes n}$ is the collection of vectors $\mathbf{v}_{k_1} \otimes \cdots \otimes \mathbf{v}_{k_n}$, where the indexes $k_1, \ldots, k_n$ range over $\{1 \cdots m\}^n$. So $V^{\otimes n}$ has dimension m^n.

More generally, any n-tuple of vectors $\mathbf{u}_1, \cdots, \mathbf{u}_n \in V$ defines a vector $\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_n \in V^{\otimes n}$. A general vector in $V^{\otimes n}$ is a linear combination of these.

7.3 Proposition. *The symmetric group acts on $V^{\otimes n}$ by permuting the factors: $\sigma \in S_n$ sends $\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_n$ to $\mathbf{u}_{\sigma(1)} \otimes \cdots \otimes \mathbf{u}_{\sigma(n)}$.*

Proof. This prescription defines a permutation action on our standard basis vectors, which extends by linearity to the entire space. By expressing a general $\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_n$ in terms of the standard basis vectors, we can check that σ maps it to $\mathbf{u}_{\sigma(1)} \otimes \cdots \otimes \mathbf{u}_{\sigma(n)}$, as claimed. $\square$

7.4 Proposition. *Setting* $\rho^{\otimes n}(g)(\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_n) = \rho(g)(\mathbf{u}_1) \otimes \cdots \otimes \rho(g)(\mathbf{u}_n)$ *and extending by linearity defines an action of* G *on* $V^{\otimes n}$*, which commutes with the action of* S_n.

Proof. Assuming this defines an action, commutation with S_n is clear, as applying $\rho^{\otimes n}(g)$ before or after a permutation of the factors has the same effect. We can see that we get an action from the first (abstract) definition of the tensor product: $\rho^{\otimes n}(g)$ certainly defines an action on the (huge) vector space based on all the symbols $\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_n$, and this action preserves the linear relations defining $V^{\otimes n}$; so it descends to a well-defined linear operator on the latter. $\square$

7.5 Corollary. *Every* S_n*-isotypical component of* $V^{\otimes n}$ *is a* G*-sub-representation.*

indeed, since $\rho^{\otimes n}(g)$ commutes with S_n, it must preserve the canonical decomposition of $V^{\otimes n}$ under S_n (proved in Lecture 4).

Recall two familiar representations of the symmetric group, the trivial 1-dimensional representation $\mathbf{1}$ and the sign representation $\varepsilon : S_n \to \{\pm 1\}$ (defined by declaring that every transposition goes to -1). The two corresponding blocks in $V^{\otimes n}$ are called $\mathrm{Sym}^n V$, the *symmetric powers*, and $\Lambda^n V$, the *exterior* or *alternating* powers of V. They are also called the *invariant/anti-invariant* subspaces, or the *symmetric/anti-symmetric parts* of $V^{\otimes n}$.

7.6 Proposition. *Let* $P_\pm$ *be the following linear self-maps on* $V^{\otimes n}$,

$$\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_n \mapsto \begin{cases} \dfrac{1}{n!} \displaystyle\sum_{\sigma \in S_n} \mathbf{u}_{\sigma(1)} \otimes \cdots \otimes \mathbf{u}_{\sigma(n)} & \text{for } P_+ \\ \dfrac{1}{n!} \displaystyle\sum_{\sigma \in S_n} \varepsilon(\sigma) \mathbf{u}_{\sigma(1)} \otimes \cdots \otimes \mathbf{u}_{\sigma(n)} & \text{for } P_- \end{cases}$$

Then:

- $\mathrm{Im}(P_\pm) \subset V^{\otimes n}$ *belongs to the invariant/anti-invariant subspace;*
- $P_\pm$ *acts as the identity on invariant/anti-invariant vectors;*
- $P_\pm^2 = P_\pm$*, and so the two operators are the projections onto* $\mathrm{Sym}^n V$ *and* $\Lambda^n V$.

Proof. The first part is easy to see: we are averaging the transforms of a vector under the symmetric group action, so the result is clearly invariant (or anti-invariant, when the sign is inserted). The second part is obvious, and the third follows from the first two. $\square$

7.7 Proposition (Basis for $\mathrm{Sym}^n V, \Lambda^n V$**).** *If* $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ *is a basis for* V*, then: a basis for* $\mathrm{Sym}^n V$ *is*

$$\left\{ \frac{1}{n!} \sum_{\sigma \in S_n} \mathbf{v}_{k_{\sigma(1)}} \otimes \cdots \otimes \mathbf{v}_{k_{\sigma(n)}} \right\}, \quad \text{as } 1 \le k_1 \le \cdots \le k_n \le m,$$

and a basis for $\Lambda^n V$ *is*

$$\left\{ \frac{1}{n!} \sum_{\sigma \in S_n} \varepsilon(\sigma) \mathbf{v}_{k_{\sigma(1)}} \otimes \cdots \otimes \mathbf{v}_{k_{\sigma(n)}} \right\}, \quad \text{as } 1 \le k_1 < \cdots < k_n \le m.$$

Proof. (Sketch) The previous proposition and our basis for $V^{\otimes n}$ make it clear that the vectors above span Sym^n and Λ^n, if we range over all choices of k_i. However, an out-of-order collection of k_i can be ordered by a permutation, resulting in the same basis vector (up to a sign, in the alternating case). So the out-of-order indexes merely repeat some of the basis vectors. Also, in the alternating case, the sum is zero as soon as two indexes have equal values. Finally, distinct sets of indexes k_i lead to linear combinations of *disjoint* sets of standard basis vectors in $V^{\otimes n}$, which shows the linear independence of our collections. $\square$

7.8 Corollary. *$\Lambda^n V = 0$ if $n > m$. More generally, $\dim \Lambda^m V = 1$ and $\dim \Lambda^k V = \binom{m}{k}$.*

Notation: In the symmetric power, we often use $\mathbf{u}_1 \cdot \ldots \cdot \mathbf{u}_n$ to denote the symmetrised vectors in Prop. 7.4, while in the exterior power, one uses $\mathbf{u}_1 \wedge \ldots \wedge \mathbf{u}_n$.

8 The character of a representation

We now turn to the core results of the course, the *character theory of group representations*. This allows an effective calculus with group representations, including their tensor products, and their decomposition into irreducibles.

We want to attach invariants to a representation ρ of a finite group G on V. The matrix coefficients of $\rho(g)$ are basis-dependent, hence not true invariants. Observe, however, that g generates a finite cyclic subgroup of G; this implies the following (see Lecture 2).

8.1 Proposition. *If G and $\dim V$ are finite, then every $\rho(g)$ is diagonalisable.* $\square$

More precisely, all eigenvalues of $\rho(g)$ will be roots of unity of orders dividing that of G. (Apply Lagrange's theorem to the cyclic subgroup generated by g.)

To each $g \in G$, we can assign its eigenvalues on V, up to order. Numerical invariants result from the elementary symmetric functions of these eigenvalues, which you also know as the coefficients of the characteristic polynomial $\det_V[\lambda - \rho(g)]$. Especially meaningful are

- the constant term, $\det \rho(g)$ (up to sign);
- the sub-leading term, $\mathrm{Tr}\rho(g)$ (up to sign).

The following is clear from multiplicativity of the determinant:

8.2 Proposition. *The map $g \mapsto \det_V \rho(g) \in \mathbb{C}$ defines a 1-dimensional representation of G.* $\square$

This is a nice, but "weak" invariant. For instance, you may know that the alternating group A_5 is *simple*, that is, it has no proper normal subgroups. Because it is not abelian, any homomorphism to $\mathbb{C}$ must be trivial, so the determinant of any of its representations is 1. Even for abelian groups, the determinant is too weak to distinguish isomorphism classes of representations. The winner turns out to be the other interesting invariant.

8.3 Definition. The character of the representation ρ is the complex-valued function on G defined by $\chi_\rho(g) := \mathrm{Tr}_V(\rho(g))$.

The Big Theorem of the course is that *this is a complete invariant*, in the sense that it determines ρ up to isomorphism.

For a representation $\rho : G \to \mathrm{GL}(V)$, we have defined $\chi_V : G \to \mathbb{C}$ by $\chi_V(g) = \mathrm{Tr}_V(\rho(g))$.

8.4 Theorem (First properties).

1. *χ_V is conjugation-invariant, $\chi_V(hgh^{-1}) = \chi_V(g), \forall g, h \in G$;*
2. *$\chi_V(1) = \dim V$;*
3. *$\chi_V(g^{-1}) = \overline{\chi_V(g)}$;*
4. *For two representations V, W, $\chi_{V\oplus W} = \chi_V + \chi_W$, and $\chi_{V\otimes W} = \chi_V \cdot \chi_W$;*
5. *For the dual representation V^*, $\chi_{V^*}(g) = \chi_V(g^{-1})$.*

Proof. Parts (1) and (2) are clear. For part (3), choose an invariant inner product on V; unitarity of $\rho(g)$ implies that $\rho(g^{-1}) = \rho(g)^{-1} = \overline{\rho(g)}^T$, whence (3) follows by taking the trace.[4] Part (4) is clear from

$$\mathrm{Tr}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \mathrm{Tr}A + \mathrm{Tr}B, \qquad \text{and} \quad \rho_{V\oplus W} = \begin{bmatrix} \rho_V & 0 \\ 0 & \rho_W \end{bmatrix}.$$

The product formula follows form the identity $\mathrm{Tr}(A \otimes B) = \mathrm{Tr}A \cdot \mathrm{Tr}B$ from Lecture 6. Finally, Part (5) follows form the fact that the action of g on a linear map $L \in \mathrm{Hom}(V;\mathbb{C}) = V^*$ is composition with $\rho(g)^{-1}$ (Lecture 5). □

8.5 Remark. conjugation-invariant functions on G are also called *central functions* or *class functions*. Their value at a group element g depends only on the conjugacy class of g. We can therefore view them as functions on the set of conjugacy classes.

(8.6) The representation ring

We can rewrite property (4) more professionally by introducing an algebraic construct.

8.7 Definition. The *representation ring* R_G of the finite group G is the free abelian group based on the set of isomorphism classes of irreducible representations of G, with multiplication reflecting the tensor product decomposition of irreducibles:

$$[V] \cdot [W] = \sum\nolimits_k n_k \cdot [V_k] \quad \text{iff} \quad V \otimes W \cong \bigoplus\nolimits_k V_k^{\oplus n_k},$$

where we write $[V]$ for the conjugacy class of an irreducible representation V, and the tensor product $V \otimes W$ has been decomposed into irreducibles V_k, with multiplicities n_k.

Thus defined, the representation ring is associative and commutative, with identity the trivial 1-dimensional representation: $[\mathbb{C}] \cdot [W] = [W]$ for any W, because $\mathbb{C} \otimes W \cong W$.

Example: $G = C_n$, the cyclic group of order n

Our irreducibles are $L_0, \ldots, L_{n-1}$, where the generator of C_n acts on L_k as $\exp\left(\frac{2\pi i k}{n}\right)$. We have $L_p \otimes L_q \cong L_{p+q}$, subscripts being taken modulo n. It follows that $R_{C_n} \cong \mathbb{Z}[X]/(X^n - 1)$, identifying L_k with X^k.

The professional restatement of property (4) in Theorem 8.4 is the following.

8.8 Proposition. *The character is a ring homomorphism from R_G to the class functions on G. It takes the involution $V \leftrightarrows V^*$ to complex conjugation.*

Later, we shall return to this homomorphism and list some other good properties.

(8.9) Orthogonality of characters

For two complex functions φ, ψ on G, we let

$$\langle \varphi | \psi \rangle := \frac{1}{|G|} \sum_{g \in G} \overline{\varphi(g)} \cdot \psi(g).$$

In particular, this defines an inner product on class functions φ, ψ, where we sum over conjugacy classes $C \subset G$:

$$\langle \varphi | \psi \rangle = \frac{1}{|G|} \sum |C| \cdot \overline{\varphi(C)} \psi(C).$$

We are now ready for the first part of our main theorem. There will be a complement in the next lecture.

[4]For an alternative argument, recall that the eigenvalues of $\rho(g)$ are roots of unity, and those of $\rho(g)^{-1}$ will be their inverses; but these are also their conjugates. Recall now that the trace is the sum of the eigenvalues.

8.10 Theorem (Orthogonality of characters).

1. *If V is irreducible, then $\|\chi_V\|^2 = 1$.*
2. *If V, W are irreducible and not isomorphic, then $\langle \chi_V | \chi_W \rangle = 0$.*

Before proving the theorem, let us list some consequences to illustrate its power.

8.11 Corollary. *The number of times an irreducible representation V appears in an irreducible decomposition of some W is $\langle \chi_V | \chi_W \rangle$.*

8.12 Corollary. *The above number (called the multiplicity of V in W) is independent of the irreducible decomposition*

We had already proved this in Lecture 4, but we now have a second proof.

8.13 Corollary. *Two representations are isomorphic iff they have the same character.*

(Use complete reducibility and the first corollary above).

8.14 Corollary. *The multiplicity of the trivial representation in W is $\frac{1}{|G|}\sum_{g\in G}\chi_W(g)$.*

8.15 Corollary (Irreducibility criterion). *V is irreducible iff $\|\chi_V\|^2 = 1$.*

Proof. Decompose V into irreducibles as $\bigoplus_k V_k^{\oplus n_k}$; then, $\chi_V = \sum_k n_k \cdot \chi_k$, and $\|\chi_V\|^2 = \sum_k n_k^2$. So $\|\chi_V\|^2 = 1$ iff all the n_k's vanish but one, whose value must be 1. $\square$

In preparation for the proof of orthogonality, we establish the following Lemma. For two representations V, W of G and any linear map $\phi : V \to W$, define

$$\phi_0 = \frac{1}{|G|}\sum_{g\in G}\rho_W(g)\circ\phi\circ\rho_V(g)^{-1}.$$

8.16 Lemma.

1. *ϕ_0 intertwines the actions G.*
2. *If V and W are irreducible and not isomorphic, then $\phi_0 = 0$.*
3. *If $V = W$ and $\rho_V = \rho_W$, then $\phi_0 = \dfrac{\mathrm{Tr}\phi}{\dim V}\cdot \mathrm{Id}$.*

Proof. For part (1), we just examine the result of conjugating by $h \in G$:

$$\begin{aligned}\rho_W(h)\circ\phi_0\circ\rho_V(h)^{-1} &= \frac{1}{|G|}\sum_{g\in G}\rho_W(h)\rho_W(g)\circ\phi\circ\rho_V(g)^{-1}\rho_V(h)^{-1}\\ &= \frac{1}{|G|}\sum_{g\in G}\rho_W(hg)\circ\phi\circ\rho_V(hg)^{-1} = \frac{1}{|G|}\sum_{g\in G}\rho_W(g)\circ\phi\circ\rho_V(g)^{-1} = \phi_0.\end{aligned}$$

Part (2) now follows from Schur's lemma. For part (3), Schur's lemma again tells us that ϕ_0 is a scalar; to find it, it suffices to take the trace over V: then, $\phi_0 = \mathrm{Tr}(\phi_0)/\dim V$. But we have

$$\mathrm{Tr}(\phi_0) = \frac{1}{|G|}\sum_{g\in G}\mathrm{Tr}\left(\rho_V(g)\circ\phi\circ\rho_V(g)^{-1}\right) = \mathrm{Tr}(\phi),$$

as claimed. $\square$

Proof of orthogonality. Choose invariant inner products on V, W and orthonormal bases $\{\mathbf{v}_i\}$ and $\{\mathbf{w}_j\}$ for the two spaces. When $V = W$ we shall use the same inner product and basis in both. We shall use the "bra·ket" vector notation explained in the handout. Then we have

$$\langle \chi_W | \chi_V \rangle = \frac{1}{|G|}\sum_{g\in G} \bar{\chi}_W(g)\chi_V(g) = \frac{1}{|G|}\sum_{g\in G} \chi_W(g^{-1})\chi_V(g)$$
$$= \frac{1}{|G|}\sum_{i,j}\sum_{g\in G} \langle \mathbf{w}_i | \rho_W(g)^{-1} | \mathbf{w}_i\rangle\langle \mathbf{v}_j|\rho_V(g)|\mathbf{v}_j\rangle.$$

We now interpret each summand as follows. Recall first that $|\mathbf{w}_i\rangle\langle\mathbf{v}_j|$ designates the linear map $V \to W$ which sends a vector $\mathbf{v}$ to the vector $\mathbf{w}_i \cdot \langle \mathbf{v}_j|\mathbf{v}\rangle$. The product $\langle \mathbf{w}_i | \rho_W(g)^{-1} | \mathbf{w}_i\rangle\langle \mathbf{v}_j|\rho_V(g)|\mathbf{v}_j\rangle$ is then interpretable as the result of applying the linear map $\rho_W(g)^{-1} \circ |\mathbf{w}_i\rangle\langle\mathbf{v}_j| \circ \rho_V(g)$ to the vector $|\mathbf{v}_j\rangle$, and then taking dot product with $\mathbf{w}_j$.

Fix now i and j and sum over $g \in G$. Lemma 8.16 then shows that the sum of linear maps

$$\rho_W(g)^{-1} \circ |\mathbf{w}_i\rangle\langle\mathbf{v}_j| \circ \rho_V(g) = \begin{cases} 0 & \text{if } V \not\cong W, \\ \operatorname{Tr}(|\mathbf{w}_i\rangle\langle\mathbf{v}_j|)\,/\dim V & \text{if } \rho_V = \rho_W. \end{cases} \tag{8.17}$$

In the first case, summing over i, j still leads to zero, and we have proved that $\chi_V \perp \chi_W$. In the second case,

$$\operatorname{Tr}(|\mathbf{w}_i\rangle\langle\mathbf{v}_j|) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases}$$

and summing over i, j leads to a factor of $\dim V$, cancelling the denominator in the second line of (8.17) and completing our proof. □

8.18 Remark. The inner product of characters can be defined over any ground field k of characteristic not dividing $|G|$, by using $\chi(g^{-1})$ in lieu of $\overline{\chi(g)}$. We have used Schur's Lemma over $\mathbb{C}$ in establishing Part (3) of Lemma 8.16, although not for Parts 1 and 2. Hence, the orthogonality of characters of non-isomorphic irreducibles holds over any such field k, but *orthonormality* $\|\chi\|^2 = 1$ can fail if k is not algebraically closed. The value of the square depends on the decomposition of the representation in the algebraic closure $\bar{k}$. This can be determined from the division algebra $\mathbb{D}_V = \operatorname{End}_k^G(V)$ and the Galois theory of its centre, which will be a finite extension field of k. See, for instance, Serre's book for more detail.

(8.19) The representation ring again

We defined the representation ring R_G of G as the free abelian group based on the isomorphism classes of irreducible representations of G. Using complete reducibility, we can identify the linear combinations of these basis elements with non-negative coefficients with isomorphism classes of G-representations: we simply send a representation $\bigoplus_i V_i^{\oplus n_i}$, decomposed into irreducibles, to the linear combination $\sum_i n_i[V_i] \in R_G$ (the bracket denotes the isomorphism class). General elements of R_G are sometimes called *virtual representations* of G. We defined the multiplication in R_G using the tensor product of representations. We sometimes denote the product by $\otimes$, to remember its origin. The one-dimensional trivial representation $\mathbf{1}$ is the unit.

We then defined the *character* χ_V of a representation ρ on V to be the complex-valued function on G with $\chi_V(g) = \operatorname{Tr}_V \rho(g)$, the operator trace in the representation ρ. This function is conjugation-invariant, or a class function on G. We denote the space of complex class functions on G by $\mathbb{C}[G]^G$. By linearity, χ becomes a map $\chi : R_G \to \mathbb{C}[G]^G$. We then checked that

- $\chi_{V\oplus W} = \chi_V + \chi_W$, $\chi_{V\otimes W} = \chi_V \cdot \chi_W$. Thus, $\chi : R_G \to \mathbb{C}[G]^G$ is a *ring homomorphism.*
- Duality corresponds to complex conjugation: $\chi(V^*) = \bar{\chi}_V$.

- The linear functional $\mathrm{Inv} : R_G \to \mathbb{Z}$, sending a virtual representation to the multiplicity (=coefficient) of the trivial representation $\mathbf{1}$ corresponds to averaging over G:

$$\mathrm{Inv}(V) = \frac{1}{|G|} \sum_{g \in G} \chi_V(g).$$

- We can define a positive definite *inner product* on R_G by $\langle V|W\rangle := \mathrm{Inv}(V^* \otimes W)$. This corresponds to the obvious inner product on characters, $\frac{1}{|G|}\sum_{g\in G} \bar{\chi}_V(g)\chi_W(g)$.

8.20 Remark. The structures detailed above for R_G is that of a *Frobenius ring with involution.* (In a ring without involution, we would just require that the bilinear pairing $V \times W \mapsto \mathrm{Inv}(V \otimes W)$ should be non-degenerate). Similarly, $\mathbb{C}[G]^G$ is a complex Frobenius algebra with involution, and we are saying that the character χ is a homomorphism of such structures. It fails to be an isomorphism "only in the obvious way": the two rings share a basis of irreducible representations (resp. characters), but the R_G coefficients are integral, not complex.

9 The Regular Representation

Recall that the regular representation of G has basis $\{\mathbf{e}_g\}$, labelled by $g \in G$; we let G permute the basis vectors according to its left action, $\lambda(h)(\mathbf{e}_g) = \mathbf{e}_{hg}$. This defines a linear action on the span of the $\mathbf{e}_g$.

We can identify the span of the $\mathbf{e}_g$ with the space $\mathbb{C}[G]$ of complex functions on G, by matching $\mathbf{e}_g$ with the function sending g to $1 \in \mathbb{C}$ and all other group elements to zero. Under this correspondence, elements $h \in G$ act ("on the left") on $\mathbb{C}[G]$, by sending the function φ to the function $\lambda(h)(\varphi)(g) := \varphi(h^{-1}g)$.

We can see that the character of the regular representation is

$$\chi_{\mathrm{reg}}(g) = \begin{cases} |G| & \text{if } g = 1 \\ 0 & \text{otherwise} \end{cases}$$

in particular, $\|\chi_{\mathrm{reg}}\|^2 = |G|^2/|G| = |G|$, so this is far from irreducible. Indeed, every irreducible representation appears in the regular representation, as the following proposition shows.

9.1 Proposition. *The multiplicity of any irreducible representation in the regular representation equals its dimension.*

Proof. $\langle\chi_V|\chi_{\mathrm{reg}}\rangle = \frac{1}{|G|}\chi_V(1)\chi_{\mathrm{reg}}(1) = \chi_V(1) = \dim V.$ □

For instance, the trivial representation appears exactly once, which we can also confirm directly: the only translation-invariant functions on G are constant.

9.2 Proposition. $\sum_V \dim^2 V = |G|$*, the sum ranging over all irreducible isomorphism classes.*

Proof. From $\chi_{\mathrm{reg}} = \sum \dim V \cdot \chi_V$ we get $\|\chi_{\mathrm{reg}}\|^2 = \sum \dim^2 V$, as desired. □

We now proceed to complete our main orthogonality theorem.

9.3 Theorem (Completeness of characters). *The irreducible characters form a basis for the space of complex class functions.*

Thus, they form an *orthonormal basis* for that space. The proof consists in showing that if a class function φ is orthogonal to every irreducible character, then its value at every group element is zero. For this, we need the following.

Construction. To every function $\varphi : G \to \mathbb{C}$, and to every representation (V, ρ) of G, we assign the following linear self-map $\rho(\varphi)$ of V:

$$\rho(\varphi) = \sum_{g \in G} \varphi(g)\rho(g).$$

We note, by applying $\rho(h)$ $(h \in G)$ and relabelling the sum, that $\rho(h)\rho(\varphi) = \rho\left(\lambda(h)\varphi\right)$.

9.4 Lemma. *φ is a class function iff $\rho(\varphi)$ commutes with G, in every representation V.*

Proof. If φ is a class function, then $\varphi(h^{-1}gh) = \varphi(g)$ for all g, h and so

$$\rho(h)\rho(\varphi)\rho(h^{-1}) = \sum_{g \in G} \varphi(g)\rho(hgh^{-1}) = \sum_{k \in G} \varphi(h^{-1}kh)\rho(k) = \sum_{k \in G} \varphi(k)\rho(k).$$

In the other direction, let V be the regular representation; then $\rho(\varphi)(\mathbf{e}_1) = \sum_{g \in G} \varphi(g)\mathbf{e}_g$, and

$$\begin{aligned} \lambda(h)\rho(\varphi)(\mathbf{e}_1) &= \sum_{g \in G} \varphi(g)\mathbf{e}_{hg} = \sum_{g \in G} \varphi(h^{-1}g)\mathbf{e}_g, \qquad \text{whereas} \\ \rho(\varphi)\lambda(h)(\mathbf{e}_1) &= \rho(\varphi)(\mathbf{e}_h) = \sum_{g \in G} \varphi(g)\mathbf{e}_{gh} = \sum_{g \in G} \varphi(gh^{-1})\mathbf{e}_g; \end{aligned}$$

equating the two shows that $\varphi(h^{-1}g) = \varphi(gh^{-1})$, as claimed. □

When φ is a class function and V is irreducible, Schur's Lemma and the previous result show that $\rho(\varphi)$ is a scalar. To find it, we compute the trace:

$$\mathrm{Tr}\left[\rho(\varphi)\right] = \sum_{g \in G} \varphi(g)\chi_V(g) = |G| \cdot \langle \chi_{V^*} | \varphi \rangle$$

(recall that $\overline{\chi_V} = \chi_{V^*}$). We obtain that, when V is irreducible,

$$\rho(\varphi) = \frac{|G|}{\dim V} \cdot \langle \chi_{V^*} | \varphi \rangle \cdot \mathrm{Id}. \tag{9.5}$$

Proof of Completeness. Assume that the class function φ is orthogonal to all irreducible characters. By (9.5), $\rho(\varphi) = 0$ in every irreducible representation V, hence (by complete reducibility) also in the regular representation. But, as we saw in the proof of Lemma 9.4, we have $\rho(\varphi)(\mathbf{e}_1) = \sum_{g \in G} \varphi(g)\mathbf{e}_g$ in the regular representation; so $\varphi = 0$, as desired. □

(9.6) The Group algebra

We now study the regular representation in greater depth. Define a multiplication on $\mathbb{C}[G]$ by setting $\mathbf{e}_g \cdot \mathbf{e}_h = \mathbf{e}_{gh}$ on the basis elements and extending by linearity:

$$\sum_g \varphi_g \mathbf{e}_g \cdot \sum_h \psi_h \mathbf{e}_h = \sum_{g,h} \varphi_g \psi_h \mathbf{e}_{gh} = \sum_g \left(\sum_h \varphi_{gh^{-1}} \psi_h \right) \mathbf{e}_g.$$

This is associative and distributive for addition, but it is not commutative unless G was so. (Associativity follows from the same property in the group.) It contains $\mathbf{e}_1$ as the multiplicative identity, and the copy $\mathbb{C}\mathbf{e}_1$ of $\mathbb{C}$ commutes with everything. This makes $\mathbb{C}[G]$ into a $\mathbb{C}$-algebra.

The group G embeds into the group of multiplicatively invertible elements of $\mathbb{C}[G]$ by $g \mapsto \mathbf{e}_g$. For this reason, we will often replace $\mathbf{e}_g$ by g in our notation, and think of elements of $\mathbb{C}[G]$ as linear combinations of group elements, with the obvious multiplication.

9.7 Proposition. *There is a natural bijection between modules over $\mathbb{C}[G]$ and complex representations of G.*

Proof. On each representation (V, ρ) of G, we let the group algebra act in the obvious way,

$$\rho\left(\sum_g \varphi_g \cdot g\right)(\mathbf{v}) = \sum_g \varphi_g \rho(g)(\mathbf{v}).$$

Conversely, given a $\mathbb{C}[G]$-module M, the action of $\mathbb{C}\mathbf{e}_1$ makes it into a complex vector space and a group action is defined by embedding G inside $\mathbb{C}[G]$ as explained above. □

We can reformulate the discussion above (perhaps more obscurely) by saying that a group homomorphism $\rho : G \to \mathrm{GL}(V)$ extends naturally to an algebra homomorphism from $\mathbb{C}[G]$ to $\mathrm{End}(V)$; V is a module over $\mathrm{End}(V)$ and the extended homomorphism makes it into a $\mathbb{C}[G]$-module. Conversely, any such homomorphism defines by restriction a representation of G, from which the original map can then be recovered by linearity.

Choosing a complete list (up to isomorphism) of irreducible representations V of G gives a homomorphism of algebras,

$$\oplus\rho_V : \mathbb{C}[G] \to \bigoplus\nolimits_V \mathrm{End}(V), \quad \varphi \mapsto (\rho_V(\varphi)). \tag{9.8}$$

Now each space $\mathrm{End}(V)$ carries two commuting actions of G: these are left composition with $\rho_V(g)$, and right composition with $\rho_V(g)^{-1}$. For an alternative realisation, $\mathrm{End}(V)$ is isomorphic to $V \otimes V^*$ and G acts separately on the two factors. There are also two commuting actions of G on $\mathbb{C}[G]$, by multiplication on the left and on the right, respectively:

$$(\lambda(h)\varphi)(g) = \varphi(h^{-1}g), \quad (\rho(h)\varphi)(g) = \varphi(gh).$$

It is clear from our construction that the two actions of $G \times G$ on the left and right sides of (9.8) correspond, hence our algebra homomorphism is also a map of $G \times G$-representations. The main result of this subsection is the following much more precise version of Proposition 9.1.

9.9 Theorem. *The homomorphism* (9.8) *is an isomorphism of algebras, and hence of $G \times G$-representations. In particular, we have an isomorphism of $G \times G$-representations,*

$$\mathbb{C}[G] \cong \bigoplus\nolimits_V V \otimes V^*,$$

with the sum ranging over the irreducible representations of G.

For the proof, we require the following

9.10 Lemma. *Let V, W be complex irreducible representations of the finite groups G and H. Then, $V \otimes W$ is an irreducible representation of $G \times H$. Moreover, every irrep of $G \times H$ is of that form.*

The reader should be cautioned that the statement can fail if the field is not algebraically closed. Indeed, the proof uses the orthonormality of characters.

Proof. Direct calculation of the square norm shows that the character $\chi_v(g) \times \chi_W(h)$ of $G \times H$ is irreducible. (Homework: do this!). Moreover, I claim that these characters span the class functions on $G \times H$. Indeed, conjugacy classes in $G \times H$ are Cartesian products of classes in the factors, and we can write every indicator class function in G and H as a linear combination of characters; so we can write every indicator class function on $G \times H$ as a linear combination of the $\chi_V \times \chi_W$. □

In particular, it follows that the representation $V \otimes V^*$ of $G \times G$ is irreducible. Moreover, no two such reps for distinct V can be isomorphic; indeed, on pairs $(g, 1) \in G \times G$, the character reduces to that of V.

Proof of Theorem 9.8. By dimension-count, it suffices to show that our map $\oplus\rho_V$ is *surjective.* The $G \times G$-structure will be essential for this. Note from our Lemma that the decomposition $S := \bigoplus_V V \otimes V^*$ splits the sum S into pairwise non-isomorphic, irreducible representations. From the results in Lecture 4, we know that any $G \times G$-map into S must be block-diagonal for the isotypical decompositions. *In particular, if a map surjects onto each factor separately, it surjects onto the sum S.* Since the summands are irreducible, it suffices then to show that each projection to $\mathrm{End}(V)$ is *non-zero.* But the group identity $\mathbf{e}_1$ maps to the identity in each $\mathrm{End}(V)$. $\square$

9.11 Remark. (i) It can be shown that the inverse map to $\oplus\rho_V$ assigns to $\phi \in \mathrm{End}(V)$ the function $g \mapsto \mathrm{Tr}_V(\rho_V(g)\phi)/\dim V$. (Optional Homework.)
(ii) Schur's Lemma ensures that the invariants on the right-hand side for the action of the *diagonal* copy $G \subset G \times G$ are the multiples of the identity in each summand. On the left side, we get the conjugation-invariant functions, or class functions.

10 The character table

In view of our theorem about completeness and orthogonality of characters, the goal of complex representation theory for a finite group is to produce the *character table.* This is the square matrix whose rows are labelled by the irreducible characters, and whose columns are labelled by the conjugacy classes in the group. The entries in the table are the values of the characters.

A representation can be regarded as "known" as soon as its character is known. Indeed, one can extract the irreducible multiplicities from the dot products with the rows of the character table. You will see in the homework (Q. 2.5) that you can even construct the isotypical decomposition of your representation, once the irreducible characters are known.

Orthonormality of characters implies a certain orthonormality of the rows of the character table, when viewed as a matrix A. "A certain" reflects to the fact that the inner product between characters is not the dot product of the rows; rather, the column of a class C must be weighed by a factor $|C|/|G|$ in the dot product (as the sum goes over group elements and not conjugacy classes). In other words, if B is the matrix obtained from A by multiplying each column by $\sqrt{|C|/|G|}$, then the orthonormality relations imply that B is a *unitary matrix.* Orthonormality of its columns leads to the

10.1 Proposition (The second orthogonality relations). *For any conjugacy classes C, C', we have, summing over the irreducible characters χ of G,*

$$\sum_\chi \overline{\chi(C)} \cdot \chi(C') = \begin{cases} |G|/|C| & \text{if } C = C', \\ 0 & \text{otherwise.} \end{cases}$$

10.2 Remark. The orbit-stabiliser theorem, applied to the conjugation action of G on itself and some element $c \in C$, shows that the prefactor $|G|/|C|$ is the order of $Z_G(c)$, the *centraliser* of c in G (the subgroup of G-elements commuting with c).

As a reformulation, we can express the indicator function E_C of a conjugacy class $C \subset G$ in terms of characters:

$$E_C(g) = \frac{|C|}{|G|} \sum_\chi \overline{\chi(C)} \cdot \chi(g).$$

Note that the coefficients need *not* be integral, so the indicator function of a conjugacy class is not the character of a representation.

(10.3) Example: The cyclic group C_n of order n

Call g a generator, and let L_k be the one-dimensional representation where g acts as ω^k, where $\omega = \exp\left(\frac{2\pi i}{n}\right)$. Then we know that $L_0, \ldots, L_{n-1}$ are all the irreducibles and the character table is the following:

	$\{1\}$	$\{g\}$	$\ldots$	$\{g^q\}$	$\ldots$
$\mathbf{1}$	1	1	$\ldots$	1	$\ldots$
L_0	1	ω	$\ldots$	ω^q	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$
L_p	1	ω^p	$\ldots$	ω^{pq}	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	

Moving on to D_6, for which we had found the one-dimensional irreducibles $\mathbf{1}$, S and the 2-dimensional irreducible V, we note that $1 + 1 + 2^2 = 6$ so there are no other irreducibles. We thus expect 3 conjugacy classes and, indeed, we find $\{1\}, \{g, g^2\}, \{r, rg, rg^2\}$, the three reflections being all conjugate ($g^{-1}rg = rg^2$, etc). From our matrix construction of the representation we fill in the entries of the character table:

	$\{1\}$	$\{g, g^2\}$	$\{rg, \ldots\}$
$\mathbf{1}$	1	1	1
S	1	1	-1
V	2	-1	0

Let us check one of the orthogonality relations: $\chi_V^2 = (2^2 + 1^2)/6 = 5/6$, which does not match the theorem very well; that is, of course, because we forgot to weigh the conjugacy classes by their size. It is helpful to make a note of the weights somewhere in the character table. The correct computation is $\chi_V^2 = (2^2 + 2 \cdot 1^2)/6 = 1$.

The general dihedral group D_{2n} works very similarly, but there is a distinction, according to the parity of n. Let us first take $n = 2m + 1$, odd, and let $\omega = \exp(2\pi i/n)$. We can again see the trivial and sign representations $\mathbf{1}$ and S; additionally, we have m two-dimensional representations V_k, labelled by $0 < k \leq m$, in which the generator g of the cyclic subgroup C_n acts as the diagonal matrix with entries ω^k, ω^{-k}, and the reflection r switches the basis vectors. Irreducibility of V_k follows as in the case $n = 3$, but we will of course be able to check that from the character. The conjugacy classes are $\{1\}$, the pairs of opposite rotations $\{g^k, g^{-k}\}$ and all n reflections, $m+2$ total, matching the number of representations we found. Here is the character table:

	$\{1\}$	$\ldots$	$\{g^q, g^{-q}\}$	$\ldots$	refs.
$\mathbf{1}$	1	$\ldots$	1	$\ldots$	1
S	1	$\ldots$	1	$\ldots$	-1
$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$
V_p	2	$\ldots$	$2\cos\frac{2\pi pq}{n}$	$\ldots$	0
$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$

The case of even $n = 2m$ is slightly different. We can still find the representations $\mathbf{1}$, S and V_k, for $1 \leq k \leq m$, but the sum of squared dimensions $1 + 1 + m \cdot 2^2 = 2m + 2$ is now incorrect. The problem is explained by observing that, in the representation V_m, g acts as $-\mathrm{Id}$, so g and r are simultaneously diagonalisable and V_m splits into two lines $L_+ \oplus L_-$, distinguished by the eigenvalues of r. The sum of squares is now the correct $1 + 1 + 4(m-1) + 1 + 1 = 4m$. Observe also that there are *two* conjugacy classes of reflections, $\{r, rg^2, \ldots, rg^{n-2}\}$ and $\{rg, rg^3, \ldots\}$.

The character table of D_{4m} is thus the following ($0 < p, q < m$):

	$\{1\}$	$\ldots$	$\{g^q, g^{-q}\}$	$\ldots$	$\{r, rg^2, \ldots\}$	$\{rg, \ldots\}$
$\mathbf{1}$	1	$\ldots$	1	$\ldots$	1	1
S	1	$\ldots$	1	$\ldots$	-1	-1
$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$	
V_p	2	$\ldots$	$2\cos\frac{2\pi pq}{n}$	$\ldots$	0	0
$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
$L_\pm$	1	$\ldots$	$(-1)^q$	$\ldots$	± 1	∓ 1

For the *Group algebra* see Lecture 13.

11 Groups of order pq

Generalising the dihedral croups, we now construct the character table of certain groups of order pq, where p and q are two primes such that $q|(p-1)$. We define a special group $F_{p,q}$ by generators and relations. It can be shown that, other than the cyclic group of the same order, this is the *only* group of order pq.

(11.1) The group $F_{p,q}$.
A theorem from algebra asserts that the group $\mathbb{Z}/p^\times$ of non-zero residue classes under multiplication is *cyclic*. A proof of this fact can be given as follows: one shows (from the structure theorem, or directly) that any abelian group of order d, in which the equation $x^d = 1$ has no more than d solutions, is cyclic. For the group $\mathbb{Z}/p^\times$, this property follows because $\mathbb{Z}/p$ is a field, so the polynomial cannot split into more than d linear factors.

Let then u be an element of multiplicative order q. (Such elements exist if $q|(p-1)$, and there are precisely $\varphi(q) = q - 1$ of them). Consider the group $F_{p,q}$ with two generators a, b satisfying the relations $a^p = b^q = 1$, $bab^{-1} = a^u$. When $q = 2$, we must take $u = -1$ and we obtain the dihedral group D_{2p}.

11.2 Proposition. *$F_{p,q}$ has exactly pq elements, namely $a^m b^n$, for $0 \le m < p$, $0 \le n < q$.*

Proof. Clearly, these elements exhaust $F_{p,q}$, because the commutation relation allows us to move all b's to the right of the a's in any word. To show that there are no further identifications, the best way is to realise the group concretely. We shall in fact embed $F_{p,q}$ within the group $\mathrm{GL}_2(\mathbb{Z}/p)$ of invertible 2×2 matrices with coefficients in $\mathbb{Z}/p$. Specifically, we consider the matrices of the form

$$\begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} : \quad x \text{ a power of } u, \quad y \text{ arbitrary.}$$

There are pq such matrices and they are closed under multiplication. We realise $F_{p,q}$ by sending $a \mapsto \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $b \mapsto \begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix}$. □

11.3 Remark. It can be shown (you'll do this in the homework) that a different choice of u leads to an isomorphic group. So omitting u from the notation of $F_{p,q}$ is in fact justified.

Let now $r = (p-1)/q$, and let $v_1, \ldots, v_r$ be a set of representatives for the cosets of $\mathbb{Z}/p^\times$ modulo the subgroup generated by u. We view the v_i as residue classes mod p; thus, $\{v_i u^j\}$ is a complete set of non-zero residue classes mod p, as $1 \le i \le r$ and $0 \le j < q$.

11.4 Proposition. *$F_{p,q}$ has $q + r$ conjugacy classes. A complete list of representatives is $\{1, a^{v_m}, b^n\}$, with $1 \le m \le r$ and $1 < n < q$.*

Proof. More precisely, the class of a^{v_m} contains the elements $a^{v_m u^i}$, $0 \le i < q$ and that of b^n the elements $a^j b^n$, $0 \le j < p$. We obtain the other powers of a by successive b-conjugation—conjugation by a does not change those elements—and the $a^j b^n$ by a-conjugation: $ab^n a^{-1} = a^{1-u^n} b$, and $u^n \ne 1$ if $n < q$, so $1 - u^n \ne 0 \mod p$ and repeated conjugation produces all powers of a. Moreover, it is clear that further conjugation by b does not enlarge this set. □

(11.5) Representations.

The subgroup generated by a is normal in $F_{p,q}$, and the quotient group is cyclic of order q. We can thus see q 1-dimensional representations of $F_{p,q}$, on which a acts trivially, but b acts by the various roots of unity of order q. We need to find r more representations, whose squared dimensions should add up to $(p-1)q = rq^2$. The obvious guess is to look for q-dimensional representations. In fact, much as we did for the dihedral group, we can find them and spot their irreducibility without too much effort.

Note that, on any representation, the eigenvalues of a are pth roots of unity, powers of $\omega = \exp \frac{2\pi i}{p}$. Assume now that ω^k is an eigenvalue, $k \ne 0$. I claim that the powers of exponent $ku, ku^2, \dots ku^{q-1}$ are also present among the eigenvalues. Indeed, let $\mathbf{v}$ be the ω^k-eigenvector; we have

$$\rho(a)\rho(b)^{-1}(\mathbf{v}) = \rho(b)^{-1}\rho(a^u)(\mathbf{v}) = \rho(b)^{-1}\omega^{ku}(\mathbf{v}) = \omega^{ku}\rho(b)^{-1}(\mathbf{v}),$$

and so $\rho(b)^{-1}(\mathbf{v})$ is an eigenvector with eigenvalue ω^{ku}. Thus, the non-trivial eigenvalues of a cannot appear in isolation, but rather in the r families $\{\omega^{v_i u^j}\}$, in each of which $0 \le j < q$. Fixing i, the smallest possible matrix is then

$$\rho_i(a) = \mathrm{diag}[\omega^{v_i}, \omega^{v_i u}, \dots, \omega^{v_i u^{q-1}}].$$

Note now that the matrix B which cyclically permutes the standard basis vectors satisfies $B\rho_i(a)B^{-1} = \rho_i(a^u)$, so we can set $\rho_i(b) = B$ to get a q-dimensional representation of $F_{p,q}$. This is irreducible, for we saw that no representation where $a \ne 1$ can have dimension lower than q.

(11.6) The character table

Let $\eta = \exp \frac{2\pi i}{q}$. Call L_k the 1-dimensional representation in which b acts as η^k and V_m the representation ρ_m above. The character table for $F_{p,q}$ looks as follows. Note that $0 < k < q$, $0 < m \le r$, $1 \le l \le r$, $0 < n < q$, and the sum in the table ranges over $0 \le s < q$.

	1	a^{v_l}	b^n
$\mathbf{1}$	1	1	1
L_k	1	1	η^{kn}
V_m	q	$\sum_s \omega^{v_m v_l u^s}$	0

The orthogonality relations are not completely obvious but require a short calculation. The key identity is $\sum \omega^{x v_m u^s} = 0$, where $x \mod p$ is a fixed, non-zero residue and m and s in the sum range over all possible choices.

12 The alternating group A_5

In this lecture we construct the character table of the alternating group A_5 of even permutations on five letters. Unlike the previous examples, we will now exploit the properties of characters in finding representations and proving their irreducibility.

We start with the conjugacy classes. Recall that every permutation is a product of disjoint *cycles*, uniquely up to their order, and that two permutations are conjugate *in the symmetric* *group* iff they have the same *cycle type*, meaning the cycle lengths (with their multiplicities). Thus, a complete set of representatives for the conjugacy classes in S_5 is

$$\mathrm{Id}, (12), (12)(34), (123), (123)(45), (1234), (12345),$$

seven in total. Of these, only $\mathrm{Id}, (12)(34), (123)$ and (12345) are in A_5.

Now, a conjugacy class in S_5 may or may not break up into several classes in A_5. The reason is that we may or may not be able to relate two given permutations in A_5 which are conjugate in A_5 by an *even* permutation. Now the first three permutations in our list all commute with some transposition. This ensures that we can implement the effect of any conjugation in S_5 by a conjugation in A_5: conjugate by the extra transposition if necessary. We do not have such an argument for the 5-cycle, so there is the possibility that the cycles (12345) and (21345) are *not* conjugate in A_5. To make sure, we count the size of the conjugacy class by the orbit-stabiliser theorem. The size of the conjugacy class containing $g \in G$ is $|G|$ divided by the order of the centraliser of g. The centraliser of (12345) in S_5 is the cyclic subgroup of order 5 generated by this cycle, and is therefore the same as the centraliser in A_5. So the order of the conjugacy class in S_5 is 24, but in A_5 only 12. So our suspicions are confirmed, and there are *two* classes of 5-cycles in A_5.

12.1 Remark. The same orbit-stabiliser argument shows that the conjugacy class in S_n of an even permutation σ is a single conjugacy class in A_n precisely when σ commutes with some odd permutation; else, it breaks up into two classes of equal size .

For use in computations, we note that the orders of the conjugacy classes are 1 for Id, 15 for $(12)(34)$, 20 for (123) and 12 each for (12345) and (21345). We have, as we should,

$$1 + 15 + 20 + 12 + 12 = 60 = |A_5|.$$

We now start producing the rows of the character table, listing the conjugacy classes in the order above. The trivial rep $\mathbf{1}$ gives the row $1, 1, 1, 1, 1$. Next, we observe that A_5 acts naturally on $\mathbb{C}^5$ by permuting the basis vectors. This representation is not irreducible; indeed, the line spanned by the vector $\mathbf{e}_1 + \ldots + \mathbf{e}_5$ is invariant. The character $\chi_{\mathbb{C}^5}$ is $5, 1, 2, 0, 0$ and its dot product $\langle \chi_{\mathbb{C}^5} | \chi_1 \rangle = 5/60 + 1/4 + 2/3 = 1$ with the trivial rep shows that $\mathbf{1}$ appears in $\mathbb{C}^5$ with multiplicity one. The complement V has character $\chi_V = 4, 0, 1, -1, -1$, whose square norm is

$$\|\chi_V\|^2 = \frac{16}{60} + \frac{1}{3} + \frac{1}{5} + \frac{1}{5} = 1,$$

showing that V is irreducible.

We seek more representations, and consider for that the tensor square $V^{\otimes 2}$. We already know this is reducible, because it splits as $\mathrm{Sym}^2 V \oplus \Lambda^2 V$, but we try extract irreducible components. For this we need the character formulae for $\mathrm{Sym}^2 V$ and $\Lambda^2 V$.

12.2 Proposition. *For any $g \in G$,*

$$\begin{aligned} \chi_{\mathrm{Sym}^2 V}(g) &= \left[\chi_V(g)^2 + \chi_V(g^2)\right]/2, \\ \chi_{\Lambda^2 V}(g) &= \left[\chi_V(g)^2 - \chi_V(g^2)\right]/2. \end{aligned}$$

Proof. Having fixed g, we choose a basis of V on which g acts diagonally, with eigenvalues $\lambda_1, \ldots, \lambda_d$, repeated as necessary. From our bases of $\mathrm{Sym}^2 V, \Lambda^2 V$, we see that the eigenvalues of g on the two spaces are $\lambda_p \lambda_q$, with $1 \le p \le q \le d$ on Sym^2 and with $1 \le p < q \le d$ on Λ^2. Then, the proposition reduces to the obvious identities

$$\begin{aligned} 2\sum_{p \le q} \lambda_p \lambda_q &= \left(\sum \lambda_p\right)^2 + \sum \lambda_p^2, \\ 2\sum_{p < q} \lambda_p \lambda_q &= \left(\sum \lambda_p\right)^2 - \sum \lambda_p^2. \quad \square \end{aligned}$$

The class function $g \mapsto \chi_V(g^2)$ is denoted by $\Psi^2\chi_V$. In our case, it is seen to give the values $(4, 4, 1, -1, -1)$, whereas χ_V^2 is $(16, 0, 1, 1, 1)$. So our characters are $\chi_{\mathrm{Sym}^2 V} = (10, 2, 1, 0, 0)$ and $\chi_{\Lambda^2 V} = (6, -2, 0, 1, 1)$. We can see that $\mathrm{Sym}^2 V$ will have a positive dot product with $\mathbf{1}$; specifically,

$$\langle \chi_{\mathrm{Sym}^2 V} | \chi_{\mathbf{1}} \rangle = \frac{1}{6} + \frac{2}{4} + \frac{1}{3} = 1,$$

so $\mathrm{Sym}^2 V$ contains $\mathbf{1}$ exactly once. We also compute

$$\langle \chi_{\mathrm{Sym}^2 V} | \chi_V \rangle = \frac{4}{6} + \frac{1}{3} = 1,$$

showing that V appears once as well. The complement W of $\mathbf{1} \oplus V$ has character $(5, 1, -1, 0, 0)$, with norm-square

$$\|\chi_W\|^2 = \frac{25}{60} + \frac{1}{4} + \frac{1}{3} = 1,$$

showing that W is a *new, 5-dimensional irreducible representation.*

We need two more representations. The squared dimensions we have so far add up to $1 + 4^2 + 5^2 = 42$, missing 18. The only decomposition into two squares is $3^2 + 3^2$, so we are looking for two 3-dimensional ones. With luck, they will both appear in $\Lambda^2 V$. This will happen precisely if

- $\|\chi_{\Lambda^2 V}\|^2 = 2$,
- $\chi_{\Lambda^2 V}$ is orthogonal to the other three characters.

We check the first condition, leaving the second as an exercise:

$$\|\chi_{\Lambda^2 V}\|^2 = \frac{36}{60} + \frac{4}{4} + \frac{1}{5} + \frac{1}{5} = 1 + 1 = 2.$$

We are left with the problem of decomposing $\chi_{\Lambda^2 V}$ into two irreducible characters. These must be unit vectors orthogonal to the other three characters. If we were sure the values are real, there would be a unique solution, discoverable by easy linear algebra. With complex values allowed, there is a one-parameter family of solutions; but it turns out only one solution will take the value 3 at the conjugacy class Id, so a bit of work will produce the answer. However, there is a better way (several, in fact).

One method is to understand $\Lambda^2 V$ from the point of view of S_5. (In the homework, you get to check that it is an irreducible representation of S_5; this is why this make sense). Notice that conjugation by your favourite transposition τ defines an automorphism of A_5, hence a permutation of the conjugacy classes. Thereunder, the first three classes stay fixed, but the two types of 5-cycles get interchanged. Therefore, this automorphism will switch the last two columns of the character table. Now, the same automorphism also acts on the set of representations, because the formula

$${}^\tau\rho(\sigma) := \rho(\tau\sigma\tau^{-1})$$

defines a new representation ${}^\tau\rho$ from any given representation ρ, acting on the same vector space. It is easy to see that ${}^\tau\rho$ is irreducible iff ρ is so (they will have the same invariant subspaces in general). So, the same automorphism also permutes the *rows* of the character table. Now, the three rows we found have dimensions $1, 4, 5$ and are the unique irreducible reps of those dimensions. So, at most, τ can switch the last two rows; else, it does nothing.

Now, if switching the last two columns did nothing, it would follow that *all characters of A_5 take equal values on the last two conjugacy classes*, contradicting the fact that the characters span all class functions. So we must be in the first case and τ must switch the last two rows. But then it follows that the bottom two entries in the first three columns (those not moved by

τ) are always equal; as we know they sum to $\chi_{\Lambda^2 V}$, they are $(3, -1, 0)$. For the last two rows and two columns, swapping the rows must have the same effect as swapping the columns, so the last 2×2 block is

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix};$$

we have $a + b = 1$ and, by orthogonality, $ab = -1$; whence $a, b = \frac{1 \pm \sqrt{5}}{2}$.

All in all, the character table for A_5 is the following:

	{Id}	(12)(34)	(123)	(12345)	(21345)
1	1	1	1	1	1
V	4	0	1	-1	-1
W	5	1	-1	0	0
Λ'	3	-1	0	a	b
Λ''	3	-1	0	b	a

12.3 Remark. There are two other ways to fill in the last two rows. One is to argue *a priori* that the character values are real. This can be done using the little result in Question 2.6.b. Knowing that, observe that the character value of a 5-cycle in a 3-dimensional representation must be a sum of three fifth roots of unity. The only real values of such sums are 3 and the two values a, b above. The value 3 could only be attained if the 5-cycles acted as the identities; but in that case, they would commute with everything in the group. This is not possible, as A_5 is simple, hence has no normal subgroups, hence is isomorphic to its image under any non-trivial representation; so any relation holding in the representation would have to hold in the group. So the only possible character values are a and b and you can easily check that the table above is the only option compatible with orthogonality.

Another way to fill the last rows, of course, would be to "spot" the representations. They actually come from the two identifications of A_5 with the group of isometries of the icosahedron. The five-cycles become rotations by multiples of $2\pi/5$, whose traces are seen to be a and b from the matrix description of these rotations.

13 Integrality in the group algebra

The goal of this section is to study the properties of *integral elements* in the group algebra, resulting in the remarkable theorem that *the degree of an irreducible representation divides the order of the group*. We recall first the notion of *integrality.*

13.1 Definition. Let $S \subset R$ be a subring of the commutative ring R. An element $x \in R$ is called *integral over S* if the following equivalent conditions are satisfied:
(a) x satisfies some monic polynomial equation with coefficients in S;
(b) The ring $S[x]$ generated by x over S within R is a finitely generated S-module.
If $S = \mathbb{Z}$, x is simply called *integral.* An integral element of $\mathbb{C}$ is called an *algebraic integer.*

Recall that an S-module M is *finitely generated* if it contains elements $m_1, \ldots, m_k$ such that every $m \in M$ is expressible as $s_1 m_1 + \cdots + s_k m_k$, for suitable $s_i \in S$. Both conditions are seen to be equivalent to the statement that, for some large N, x^N is expressible in terms of the x^n, $n < N$, with coefficients in S. Here are two general facts that we will use.

13.2 Proposition. *If $x, y \in R$ are both integral over S, then so are $x + y$ and xy.*

Proof. We know that x is integral over S and that y is integral over $S[x]$. Hence $S[x]$ is a finitely generated S-module and $S[x, y]$ is a finitely generated $S[x]$-module. All in all, $S[x, y]$ is a finitely generated S-module. (The products of one generator for $S[x]$ over S and of $S[x, y]$ over $S[x]$ give a set of generators of $S[x, y]$ over S.) □

13.3 Proposition. *The image of an integral element under any ring homomorphism is integral.*

This is clear from the first definition. Finally, we note the following fact about algebraic integers.

13.4 Proposition. *Any rational algebraic integer is a 'genuine' integer (in $\mathbb{Z}$).*

Sketch of proof. Assume that the reduced fraction $x = p/q$ satisfies a monic polynomial equation of degree n with integer coefficients and get a contradiction by showing that the denominator q^n in the leading term cannot be cancelled by any of the other terms in the sum. $\square$

13.5 Remark. The proposition is a special case of *Gauss' Lemma*, which says that any monic factor with rational coefficients of a monic polynomial with integral coefficients must in fact have integral coefficients.

Our goal is the following

13.6 Theorem. *The dimension of any complex irreducible representation of a finite group divides the order of the group.*

Calling G the group and V the representation, we will prove that $|G|/\dim V$ is an *algebraic integer.* The proposition above will then imply the theorem. Observe the identity

$$\frac{|G|}{\dim V} = \sum_{g \in G} \chi_V(g^{-1}) \cdot \frac{\chi_V(g)}{\dim V} = \sum_C \overline{\chi_V}(C)\chi_V(C) \cdot \frac{|C|}{\dim V};$$

where the last sum ranges over the conjugacy classes C. The theorem then follows from the following two propositions.

13.7 Proposition. *The values $\chi(g)$ of the character of any representation of G are algebraic integers.*

Proof. Indeed, the eigenvalues of $\rho(g)$ are roots of unity. $\square$

13.8 Proposition. *For any conjugacy class $C \subset G$ and irreducible representation V, the number $\chi_V(C) \cdot |C|/\dim V$ is an algebraic integer.*

To prove the last proposition, we need to recall the group algebra $\mathbb{C}[G]$, the space of formal linear combinations $\sum \varphi_g \cdot g$ of group elements with multiplication prescribed by the group law and the distributive law. It contains the subspace of class functions. We need to know that this is a commutative subalgebra; in fact we have the following, more precise

13.9 Proposition. *The space of class functions is the centre $Z(\mathbb{C}[G])$ of $\mathbb{C}[G]$.*

Proof. Class functions are precisely those elements $\varphi \in \mathbb{C}[G]$ for which $\lambda(g)\varphi\rho(g)^{-1} = \varphi$, or $\lambda(g)\varphi = \varphi\rho(g)$, in terms of the left and right multiplication actions of G. Rewriting in terms of the multiplication law in $\mathbb{C}[G]$, this says $g \cdot \varphi = \varphi \cdot g$, so class functions are those commuting with every $g \in G$. But that is equivalent to their commutation with every element of $\mathbb{C}[G]$. $\square$

Let now $E_C \in \mathbb{C}[G]$ denote the indicator function of the conjugacy class $C \subset G$. Note that $E_{\{1\}}$ is the identity in $\mathbb{C}[G]$.

13.10 Proposition. *$\bigoplus_C \mathbb{Z} \cdot E_C$ is a subring of $Z(\mathbb{C}[G])$. In particular, every E_C is integral in $Z(\mathbb{C}[G])$.*

Proof. Writing $E_C = \sum_{g \in C} g$ makes it clear that a product of two E's is a sum of group elements (with integral coefficients). Grouping them by conjugacy class leads to an integral combination of the other E's. Integrality of E_C follows, because the ring $\mathbb{Z}[E_C]$ is a subgroup of the free, finitely generated abelian group $\bigoplus_C \mathbb{Z} \cdot E_C$, and so it is a finitely generated abelian group as well. $\square$

Recall now that the group homomorphism $\rho_V : G \to \mathrm{GL}(V)$ extends by linearity to an algebra homomorphism $\rho_V : \mathbb{C}[G] \to \mathrm{End}(V)$; namely, $\rho_V(\varphi) = \sum_g \varphi_g \cdot \rho_V(g)$. This is compatible with the conjugation action of G, and so class functions must go to the subalgebra $\mathrm{End}^G(V)$ of G-invariant endomorphisms. When V is irreducible, the latter is isomorphic to $\mathbb{C}$, by Schur's lemma. Since $\mathrm{Tr}_V(\alpha \cdot \mathrm{Id}) = \alpha \dim V$, the isomorphism is realised by taking the trace, and dividing by the dimension. We obtain the following

13.11 Proposition. *When V is irreducible,* $\mathrm{Tr}_V / \dim V$ *is an algebra homomorphism:*

$$\frac{\mathrm{Tr}_V(\,)}{\dim V} : Z(\mathbb{C}[g]) \to \mathbb{C}, \quad \sum\nolimits_g \varphi_g \cdot g \mapsto \sum\nolimits_g \varphi_g \cdot \frac{\chi_V(g)}{\dim V}. \qquad \square$$

Applying this homomorphism to E_C results in $\chi_V(C)\frac{|C|}{\dim V}$. Propositions 13.3 and 13.10 imply (13.8) and thus our theorem.

14 Induced Representations

The study of A_5 illustrates the need for more methods to construct representations. Now, for any group, we have the regular representation, which contains every irreducible; but there is no good method to break it up in general. We need some smaller representations.

One tool is suggested by the first of the representations of A_5, the permutation representation $\mathbb{C}^5$. In general, whenever G acts on X, it will act linearly on the space $\mathbb{C}[X]$ of complex-valued functions on X. This is called the *permutation representation* of X. Some of its basic properties are listed in Q. 2.3. For example, the decomposition $X = \coprod X_k$ into orbits of the G-action leads to a direct sum decomposition $\mathbb{C}[X] = \bigoplus \mathbb{C}[X_k]$. So if we are interested in irreducible representations, we might as well confine ourselves to transitive group actions. In this case, X can be identified with the space G/H of left cosets of G for some subgroup H (the stabiliser of some point). You can check that, if $G = A_5$ and $H = A_4$, then $\mathbb{C}[X] \cong \mathbb{C}^5$, the permutation representation studied in the last lecture.

We will generalise this construction. Let then $H \subset G$ be a subgroup and $\rho : H \to \mathrm{GL}(W)$ a representation thereof.

14.1 Definition. The *induced representation* $\mathrm{Ind}_H^G W$ is the subspace of the space of all maps $G \to W$, with the left action of G on G, which are invariant under the action of H, on the right on G and simultaneously by ρ on W.

Under the *left action* of an element $k \in G$, a map $\varphi : G \to W$ gets sent to the map $g \mapsto \varphi(k^{-1}g) \in W$. The *simultaneous action* of $h \in H$ on G and on W sends it to the map $g \mapsto \rho(h)\,(\varphi(gh))$. This makes it clear that the two actions, of G and of H, commute, and so the subspace of H-invariant maps is a sub-representation of G.

We can write the actions a bit more explicitly by observing that the space of W-valued maps on G is isomorphic to $\mathbb{C}[G] \otimes W$: using the natural basis $\mathbf{e}_g$ of functions on G and any basis $\{\mathbf{f}_i\}$ of W, we identify the basis element $\mathbf{e}_g \otimes \mathbf{f}_i$ of $\mathbb{C}[G] \otimes W$ with the map taking $g \in G$ to $\mathbf{f}_i$ and every other group element to zero. In this basis, the left action of $k \in G$ sends $\mathbf{e}_g \otimes \mathbf{w}$ to $\mathbf{e}_{kg} \otimes \mathbf{w}$. The right action of $h \in H$ sends it to $\mathbf{e}_{gh^{-1}} \otimes \mathbf{w}$, whereas the simultaneous action on the right and on W sends it to $\mathbf{e}_{kg} \otimes \rho(h)(\mathbf{w})$. A general function $\sum_{g \in G} \mathbf{e}_g \otimes \mathbf{w}_g \in \mathbb{C}[G] \otimes W$ gets sent by h to

$$\sum_{g \in G} \mathbf{e}_{gh^{-1}} \otimes \rho(h)(\mathbf{w}_g) = \sum_{g \in G} \mathbf{e}_g \otimes \rho(h)(\mathbf{w}_{gh}).$$

Thus, the H-invariance condition is

$$\rho(h)(\mathbf{w}_{gh}) = \mathbf{w}_g, \quad \forall h \in H. \tag{*}$$

Let $R \subset G$ be a set of representatives for the cosets G/H. For $r \in R$, let $W_r \subset \mathrm{Ind}_H^G W$ be the subspace of *functions supported on* rH, that is, the functions vanishing everywhere else.

14.2 Proposition. *Each vector space W_r is isomorphic to W, and we have a decomposition, as a vector space,* $\mathrm{Ind}_H^G W \cong \bigoplus_r W_r$.

In particular, $\dim \mathrm{Ind}_H^G W = |G/H| \cdot \dim W$.

Proof. For each r, we construct an isomorphism from W to W_r by

$$\mathbf{w} \in W \mapsto \sum_{h \in H} \mathbf{e}_{rh} \otimes \rho(h^{-1})(\mathbf{w}).$$

Clearly, this function is supported on rH, and, in view of (*), it is H-invariant; moreover, the same condition makes it clear that every invariant function on rH has this form, for some $\mathbf{w} \in W$. Also, every function $\sum_{g \in G} \mathbf{e}_g \otimes \mathbf{w}_g$ can be re-written as

$$\sum_{r \in R} \sum_{h \in H} \mathbf{e}_{rh} \otimes \mathbf{w}_{rh} = \sum_{r \in R} \sum_{h \in H} \mathbf{e}_{rh} \otimes \rho(h^{-1})(\mathbf{w}_r),$$

in view of the same H-invariance condition. This gives the desired decomposition $\bigoplus_r W_r$. □

Examples

1. If $H = \{1\}$ and W is the trivial representation, $\mathrm{Ind}_H^G W$ is the regular representation of G.
2. If H is any subgroup and W is the trivial representation, then $\mathrm{Ind}_H^G W \cong \mathbb{C}[G/H] \otimes W$, the permutation representation on G/H times the vector space W with the trivial G-action.
3. (This is not obvious) More generally, if the action of H on W extends to an action of the ambient group G, then $\mathrm{Ind}_H^G W \cong \mathbb{C}[G/H] \otimes W$, with the extended G-action on W.

Let us summarise the basic properties of induction.

14.3 Theorem (Basic Properties).

(i) $\mathrm{Ind}(W_1 \oplus W_2) \cong \mathrm{Ind}W_1 \oplus \mathrm{Ind}W_2$.
(ii) $\dim \mathrm{Ind}_H^G W = |G/H| \dim W$.
(iii) $\mathrm{Ind}_{\{1\}}^G \mathbf{1}$ *is the regular representation of* G.
(iv) Let $H \subset K \subset G$. *Then,* $\mathrm{Ind}_K^G \mathrm{Ind}_H^K W \cong \mathrm{Ind}_H^G W$.

Proof. Part (i) is obvious from the definition, and we have seen (ii) and (iii). Armed with sufficient patience, you can check property (iv) from the character formula for the induced representation that we will give in the next lecture. Here is the proof from first definitions. $\mathrm{Ind}_H^G W$ is the space of maps from G to W that are invariant under H. An element of $\mathrm{Ind}_K^G \mathrm{Ind}_H^K W$ is a map from G into the space of maps from K into W, invariant under K and under H. This is also a map $\phi : G \times K \to W$, satisfying

$$\phi(gk', k) = \phi(g, k'k) \text{ and } \phi(g, kh^{-1}) = \rho_W(h)\phi(g, k)$$

Such a map is determined by its restriction to $G \times \{1\}$, and the resulting restriction $\psi : G \to W$ satisfies $\psi(gh^{-1}) = \rho_W(\psi)(g, k)$ for $h \in H$, and as such determines an element of $\mathrm{Ind}_H^G W$. Conversely, such a ψ is extended to a ϕ on all of $G \times K$ using the K-invariance condition. □

15 Induced characters and Frobenius Reciprocity

(15.1) Character formula for $\mathrm{Ind}_H^G W$.

Induction is used to construct new representations, in the hope of producing irreducibles. It is quite rare that an induced rep will be irreducible (but we will give a useful criterion for that, due to Mackey); so we have to extract the 'known' irreducibles with the correct multiplicities. The usual method to calculate multiplicities involves characters, so we need a *character formula for the induced representation.*

We begin with a procedure which assigns to any class function on H a class function on G. The obvious extension by zero is usually not a class function—the complement of H in G is usually not a union of conjugacy classes. The most natural fix would be to take the transforms of the extended function under conjugation by all elements of G, and add them up. This, however, is a bit redundant, because many of these conjugates coincide:

15.2 Lemma. *If φ is a class function on H, extended by zero to all of G, and $x \in G$ is fixed, then the "conjugate function" $g \in G \mapsto \varphi(x^{-1}gx)$ depends only on the coset xH.*

Proof. Indeed, $\varphi(h^{-1}x^{-1}gxh) = \varphi(x^{-1}gx)$, because φ was invariant under conjugation in H. □

This justifies the following; recall that $R \subset G$ is a system of representatives for the cosets G/H.

15.3 Definition. Let φ be a class function on H. The *induced function* $\mathrm{Ind}_H^G(\varphi)$ is the class function on G given by

$$\mathrm{Ind}_H^G(\varphi)(g) = \sum_{r\in R} \varphi(r^{-1}gr) = \frac{1}{|H|}\sum_{x\in G}\varphi(x^{-1}gx),$$

where we declare φ to be zero at the points of $G \setminus H$.

The sum thus ranges over those r for which $r^{-1}gr \in H$. We are ready for the main result.

15.4 Theorem. *The character of the induced representation* $\mathrm{Ind}_H^G W$ *is* $\mathrm{Ind}_H^G(\chi_W)$.

Proof. Recall from the previous lecture that we can decompose $\mathrm{Ind}W$ as

$$\mathrm{Ind}_H^G W = \bigoplus\nolimits_{r\in R} W_r, \tag{15.5}$$

where $W_r \cong W$ as a vector space. Moreover, as each W_r consists of the functions supported only on the coset rH, the action of G on $\mathrm{Ind}W$ permutes the summands, in accordance to the action of G on the left cosets G/H.

We want the trace of the action of $g \in G$. We can confine ourselves to the summands in (15.5) which are preserved by g; the others will contribute purely off-diagonal blocks to the action. Now, $grH = rH$ iff $r^{-1}gr = k \in H$. For each such r, we need the trace of g on W_r. Now the isomorphism $W \cong W_r$ was defined by

$$\mathbf{w} \in W \mapsto \sum\nolimits_{h\in H} \mathbf{e}_{rh} \otimes h^{-1}\mathbf{w},$$

which transforms under g to give

$$g\sum_{h\in H}\mathbf{e}_{rh}\otimes h^{-1}\mathbf{w} = \sum_{h\in H}\mathbf{e}_{grh}\otimes h^{-1}\mathbf{w} = \sum_{h\in H}\mathbf{e}_{rkh}\otimes h^{-1}\mathbf{w} = \sum_{h\in H}\mathbf{e}_{rh}\otimes h^{-1}k\mathbf{w},$$

and so the action of g on W_r corresponds, under this isomorphism, to the action of $k = r^{-1}gr$ on W, and so the block W_r contributes $\chi_W(r^{-1}gr)$ to the trace of g. Summing over the relevant r gives our theorem. □

(15.6) Frobenius Reciprocity

Surprisingly, one can compute multiplicities of G-irreducibles in Ind$W$ which does *not* require knowledge of its character on G, but only the character of W in H. The practical advantage is that a dot product of characters on G, involving one induced rep, is reduced to one on H. This is the content of the *Frobenius reciprocity theorem* below. To state it nicely, we introduce some notation. Write $\mathrm{Res}_G^H V$, or just Res$V$, when regarding a G-representation V merely as a representation of the subgroup H. (Res$V$ is the "restricted" representation). For two G-representations U, V, let

$$\langle U|V\rangle_G = \langle \chi_U|\chi_V\rangle;$$

recall that this is also the dimension of the space $\mathrm{Hom}^G(U,V)$ of G-invariant linear maps. (Check it by decomposing into irreducibles.) Because the isomorphism $\mathrm{Hom}(U,V) \cong V \otimes U^*$ is compatible with the actions of G, the same number is also the dimension of the subspace $(V \otimes U^*)^G \subset V \otimes U^*$ of G-invariant vectors.

15.7 Theorem (Frobenius Reciprocity). $\langle V|\mathrm{Ind}_H^G W\rangle_G = \langle \mathrm{Res}_G^H V|W\rangle_H$.

15.8 Remark. Viewing the new $\langle\,|\,\rangle_G$ as an inner product between representations, the theorem says, in more abstract language, that Ind_H^G and Res_G^H are *adjoint functors*.

15.9 Corollary. *The multiplicity of the G-irreducible representation V in the induced representation* $\mathrm{Ind}_H^G W$ *is equal to the multiplicity of the irreducible H-representation W in the restricted representation* $\mathrm{Res}_G^H V$.

Proof of Frobenius reciprocity. One can check directly (you will do so in the homework) that the linear map Ind, from class functions on H to class functions on G, is the hermitian adjoint of the restriction of class functions from G to H; the Reciprocity formula then follows from Theorem 15.4. Here, we give a proof which applies more generally (even when character theory does not).

We must show that $\dim \mathrm{Hom}^G(V;\mathrm{Ind}W) = \dim \mathrm{Hom}^H(V;W)$. We prove this by constructing a *natural* isomorphism between the two spaces (independent of arbitrary choices),

$$\mathrm{Hom}^G(V;\mathrm{Ind}W) \cong \mathrm{Hom}^H(V;W). \qquad (*)$$

The chain of isomorphisms leading to (*) is the following:

$$\begin{aligned}\mathrm{Hom}^G(V;\mathrm{Ind}W) = \mathrm{Hom}^G\left(V;(\mathbb{C}[G]\otimes W)^H\right) &= \mathrm{Hom}^{G\times H}(V;\mathbb{C}[G]\otimes W)\\ &= (\mathbb{C}[G]\otimes \mathrm{Hom}(V;W))^{G\times H} = \mathrm{Hom}^H(V;W),\end{aligned}$$

where the superscripts G, H indicate that we take the subspaces of invariant vectors under the respective groups, and it remains to explain the group actions and the isomorphisms.

The first step is the definition of the induced representation. Now, because the G-action (on V and on the left on $\mathbb{C}[G]$) commutes with the H-action (on the right on $\mathbb{C}[G]$ and on W), we get an action of $G \times H$ on $\mathrm{Hom}(V;\mathbb{C}[G]\otimes W)$, and the invariant vectors under $G \times H$ are precisely those invariant under *both* G and H; which accounts for our second isomorphism. In the third step, we have used the isomorphism of vector spaces

$$\mathrm{Hom}(V;\mathbb{C}[G]\otimes W) = \mathbb{C}[G]\otimes \mathrm{Hom}(V;W),$$

which equates the linear maps from V to W-valued functions on G with the functions on G with values in $\mathrm{Hom}(V;W)$: indeed, both sides are maps from $G \times V$ to W which are linear in V. We keep note of the fact that G acts on the left on $\mathbb{C}[G]$ and on V, while H acts on the right on $\mathbb{C}[G]$ and on W.

The last step gets rid of G altogether. It is explained by the fact that a map on G with values *anywhere* which is invariant under left multiplication of G is completely determined by its value at the identity; and this value can be prescribed without constraint. To be more precise, let $\phi : G \times V \to W$ be a map linear on V, and denote by $\phi_g \in \mathrm{Hom}(V, W)$ its value at G. Invariance under g forces us to have $\phi_{gk} = \phi_k \circ \rho_V(g^{-1})$, as seen by acting on ϕ by g and evaluating at $k \in G$. In particular, $\phi_g = \phi_1 \circ \rho_V(g^{-1})$. Conversely, for any $\psi_1 = \psi \in \mathrm{Hom}(V; W)$, the map on G sending k to $\psi_k := \psi_1 \circ \rho_V(k)^{-1}$ is invariant, because

$$\psi_{gk} = \psi_1 \circ \rho_V(gk)^{-1} = \psi_1 \circ \rho_V(k)^{-1}\rho_V(g)^{-1} = \psi_k \circ \rho_V(g)^{-1},$$

as needed for invariance under any $g \in G$.

We have thus shown that sending $\phi : G \to \mathrm{Hom}(V; W)$ to ϕ_1 gives a linear isomorphism from the subspace of G-invariants in $\mathbb{C}[G] \otimes \mathrm{Hom}(V; W)$ to $\mathrm{Hom}(V; W)$. It remains to compare the H-invariance conditions on the two spaces, to establish the fourth and last isomorphism in the chain. It turns out, in fact, that the H-action on the first space, on the right on G and on W, agrees with the H-action on $\mathrm{Hom}(V, W)$, simultaneously on V and on W. Indeed, acting by $h \in H$ on ϕ results in the map $g \mapsto \rho_W(h) \circ \phi_{gh}$. In particular, 1 goes to $\rho_W(h) \circ \phi_h$. But we know that $\phi_h = \phi_1 \circ \rho_V(h)^{-1}$, whence it follows that the transformed map sends $1 \in G$ to $\rho_W(h) \circ \phi_1 \circ \rho_V(h)^{-1}$. So the effect on $\phi_1 \in \mathrm{Hom}(V; W)$ is that of the simultaneous action of h on V and W, as claimed. ☐

16 Mackey theory

The main result of this section describes the restriction to a subgroup $K \subset G$ of an induced representation $\mathrm{Ind}_H^G W$. No relation is assumed between K and H, but the most useful application of this rather technical formula comes in when $K = H$: in this case, we obtain a neat irreducibility criterion for $\mathrm{Ind}_H^G W$. The formulae become much cleaner when H is a normal subgroup of G, but this plays no role in their proof.

16.1 Definition. The *double coset space* $K\backslash G/H$ is the set of $K \times H$-orbits on G, for the left action of K and the right action of H.

Note that left and right multiplication commute, so we get indeed an action of the product group $K \times H$. There are equivalent ways to describe the set, as the set of orbits for the (left) action of K on the coset space G/H, or the set of orbits for the right action of H on the space $K\backslash G$. An important special case is when $H = K$ and is *normal* in G: in this case, $K\backslash G/H$ equals the simple coset space G/H. (Why?)

Let $S \in G$ be a set of representatives for the double cosets (one point in each orbit); we thus have $G = \coprod_s KsH$. For each $s \in S$, let

$$H_s = sHs^{-1} \cap K.$$

Then, H_s is a subgroup of K, but it also embeds into H by the homomorphism $H_s \ni x \mapsto s^{-1}hs$. To see its meaning, note that $H_s \subset K$ is the stabiliser of the coset sH under the action of K on G/H, while the embedding in H identifies H_s with the stabiliser of Ks under the (right) action of H on $K\backslash G$.

A representation $\rho : H \to \mathrm{GL}(W)$ leads to a representation of H_s on the same space by $\rho_s(x) = \rho(s^{-1}xs)$. We write W_s for the space W, when this action of H_s is understood.

16.2 Theorem (Mackey's Restriction Formula).

$$\mathrm{Res}_G^K \mathrm{Ind}_H^G(W) \cong \bigoplus_{s \in S} \mathrm{Ind}_{H_s}^K(W_s)$$

Proof. We have $G = \coprod_s KsH$, and the left K-action and right H-action permute these double cosets. We can then decompose the space of W-valued functions on G into a sum of functions supported on each coset. H-invariance of the decomposition means that the invariant subspace $\mathrm{Ind}_H^G(W) \subset \mathbb{C}[G] \otimes W$ decomposes accordingly, into the sum of H-invariants in $\bigoplus_s \mathbb{C}[KsH] \otimes W$.

We claim now that the H-invariant part of $\mathbb{C}[KsH] \otimes W$ is isomorphic to $\mathrm{Ind}_{H_s}^K(W_s)$, as a representation of K. Let then $\phi : KsH \to W$ be a map invariant for the action of H (on the right on KsH, and on W). Restricting to the subset Ks gives a map $\psi : K \to W$ invariant under the same action of H_s, by setting $\psi(k) := \phi(ks)$. But this is a vector in $\mathrm{Ind}_{H_s}^K(W_s)$, so we do get a natural linear map $\phi \mapsto \psi$ from the first space to the second.

Conversely, let now $\psi : K \to W$ represent a vector in the induced representation $\mathrm{Ind}_{H_s}^K(W_s)$; we define a map $\phi : KsH \to W$ by $\phi(ksh) := \rho(h)^{-1}\psi(k)$. If $x \in H$, then

$$\phi(ksh \cdot x^{-1}) = \rho(xh^{-1})\psi(k) = \rho(x)\phi(ksh),$$

showing that ϕ is invariant under the action on H (on the right on KsH and simultaneously on W). There is a problem, however: the same group element may have several expressions as ksh, with $k \in K$ and $h \in H$, and we must check that the value $\phi(ksh)$ is independent of that. Let $ksh = k'sh'$; then, $k^{-1}k' = s \cdot h'h^{-1} \cdot s^{-1}$, and so their common value y is in H_s. But then,

$$\phi(ksh) = \rho(h)^{-1}\psi(k) = \rho(h)^{-1}\rho(y)^{-1}\psi(ky) = \rho(yh)^{-1}\psi(kk^{-1}k') = \rho(h')^{-1}\psi(k') = \phi(k'sh'),$$

as needed, where for the second equality we have used invariance of ψ under $y \in H_s$. The maps $\phi \mapsto \psi$ and $\psi \mapsto \phi$ are clearly inverse to each other, completing the proof of the theorem. □

16.3 Remark. One can give a computational proof of Mackey's theorem using the formula for the character of an induced representation. However, the proof given above applies even when character theory does not (e.g. when complete reducibility fails).

(16.4) Mackey's irreducibility criterion

We will now investigate irreducibility of the induced representation $\mathrm{Ind}_H^G(W)$ by applying Mackey's formula to the case $K = H$.

Note that, when $K = H$, we have for each $s \in S$ *two* representations of the group $H_s = sHs^{-1} \cap H$ on the space W: ρ_s defined as above and the restricted representation $\mathrm{Res}_H^{H_s} W$. The two actions stem from the two embeddings of H_s inside H, by s^{-1}-conjugation and by the natural inclusion, respectively. Call two representations of a group *disjoint* if they have no irreducible components in common. This is equivalent to orthogonality of their characters.

16.5 Theorem (Mackey's Irreducibility Criterion). $\mathrm{Ind}_H^G W$ *is irreducible if and only if*

(i) W is irreducible

(ii) For each $s \in S \setminus H$, the representations W_s and $\mathrm{Res}_H^{H_s} W$ are disjoint.

16.6 Remark. The set of representatives S for the double cosets was arbitrary, so we might as well impose condition (ii) on any $g \in G \setminus H$; but it suffices to check it for $g \in S$.

The criterion becomes especially effective when H is normal in G. In this case, the double cosets are the same as the simple cosets (left or right). Moreover, $sHs^{-1} = H$, so $H_s = H$ for all s; and the representation W_s of H is irreducible, if W was so. Hence the following

16.7 Corollary. *Let $H \subset G$ be a normal subgroup. Then, $\mathrm{Ind}_H^G W$ is irreducible iff W is irreducible and the representations W and W_s are non-isomorphic, for all $s \in G \setminus H$.*

Again, it suffices to check the condition on a set of representatives; in fact, it turns out that the isomorphism class of W_g, for $g \in G$, depends only on the coset gH.

Proof of Mackey's criterion. We have, in the notation $\langle U|V\rangle := \langle \chi_U|\chi_V\rangle$,

$$\begin{aligned}\langle \mathrm{Ind}_H^G W|\mathrm{Ind}_H^G W\rangle_G &= \langle \mathrm{Res}_G^H \mathrm{Ind}_H^G W|W\rangle_H \\ &= \sum_s \langle \mathrm{Ind}_{H_s}^H W_s|W\rangle_H = \sum_s \langle W|\mathrm{Ind}_{H_s}^H W_s\rangle_H = \sum_s \langle \mathrm{Res}_H^{H_s} W|W_s\rangle_{H_s}.\end{aligned}$$

For the first and last equality, we have used the Frobenius reciprocity formula; the second equality is Mackey's formula, while the third uses the fact that the dot product of characters is an integer (hence real).

Now every term in the final sum is a non-negative integer, so the the answer is 1, as required by irreducibility, only if all terms but one vanish, and the non-zero term is 1 itself. Now the choice of representatives S was arbitrary, so we can assume that 1 was chosen to represent the identity coset $HH = H$. If so, $H_1 = H$ and $W_1 = W$, and the $s = 1$ term is $\|\chi_W\|^2$. So this is the non-vanishing term, and it must equal 1, and so W must be irreducible; while all other terms must vanish, which is part (ii) of Mackey's criterion. □

(16.8) Examples

1. If $G = D_{2n}$, $H = C_n$ and W is the representation L_k with generator eigenvalue $\exp\left(\frac{2\pi i k}{n}\right)$, the induced representation is V_k. The transform of L_k under reflection is the representation L_{-k}, and Mackey's criterion holds whenever $\exp\left(\frac{2\pi i k}{n}\right) \neq \exp\left(\frac{-2\pi i k}{n}\right)$, that is, when $k \neq 0, n/2$ (when n is even); and indeed, in each of those cases, V_k breaks up into two lines.

2. If $G = S_5$ and $H = A_5$, Mackey's criterion fails for the first three irreducible representations $\mathbf{1}, V$ and W, but holds for the two 'halves' of $\Lambda^2 V$. So the first three representations induce reducible ones, but the last two induce irreducible representations. One can check (by characters, but also directly) that the first three induced representations split as $\mathbf{1} \oplus S$, $V \oplus (V \otimes S)$ and $W \oplus (W \otimes S)$, while Λ' and Λ'' each induce $\Lambda^2 V$, which is irreducible under S_5. As a matter of fact, this leads to the complete list of irreducibles of S_5.

17 Nilpotent groups and p-groups

As an application of induced representations, we now show that for a certain class of groups (the nilpotent ones) all irreducible representations arise by induction from a 1-dimensional representation of an appropriate subgroup. Except for the definition that we are about to give, this section contains *optional material.*

17.1 Definition. The group G is *solvable* if there is a finite chain

$$G = G^{(n)} \supset G^{(n-1)} \supset \cdots \supset G^{(0)} = \{1\}$$

of subgroups such that $G^{(k)}$ is normal in $G^{(k+1)}$ and the quotient group $G^{(k+1)}/G^{(k)}$ is abelian. G is *nilpotent* if the $G^{(k)}$ are normal in all of G and each $G^{(k+1)}/G^{(k)}$ is central in $G/G^{(k)}$.

(17.2) Examples

(i) Every abelian group is nilpotent.
(ii) The group B of upper-triangular invertible matrices with entries in $\mathbb{Z}/p$ is solvable, but not nilpotent; the subgroups $B^{(k)}$ have 1's on the diagonal and zeroes on progressively more super-diagonals.
(iii) The subgroup $N \subset B$ of upper-triangular matrices with 1's on the diagonal is nilpotent.
(iv) Every subgroup and quotient group of a solvable (nilpotent) group is solvable (nilpotent): just use the chain of intersections with the $G^{(k)}$ or the quotient chain, as appropriate.

The following proposition underscores the importance of nilpotent groups. Recall that a *p-group* is a group whose order is a power of the prime number p.

17.3 Proposition. *Every p-group is nilpotent.*

Proof. We will show that the centre Z of a p-group G is non-trivial. Since G/Z is also a p-group, we can repeat the argument and produce the chain in 17.1.

Consider the conjugacy classes in G. Since they are orbits of the conjugation action, the orbit-stabiliser theorem implies that their orders are powers of p. If 1 was the only central element, then the order of every other conjugacy class would be $0 \pmod p$. They would then add up to $1 \pmod p$. However, the order of G is $0 \pmod p$, contradiction. $\square$

17.4 Remark. The same argument can be applied to the (linear) action of G on a vector space over a finite field of characteristic p, and implies that *every such action has non-zero invariant vectors.* This has the remarkable implication that *the only irreducible representation of a p-group in characteristic p is the trivial one.*

17.5 Theorem. *Every irreducible representation of a nilpotent group is induced from a one-dimensional representation of an appropriate subgroup.*

The subgroup will of course depend on the representation. Still, this is rather a good result, especially for p-groups, whose structure is usually sensible enough that an understanding of all subgroups is not a hopeless task. The proof will occupy the rest of the lecture and relies on the following

17.6 Lemma (Clever Lemma). *Let G be a finite group, $A \subset G$ a normal subgroup and V an irreducible representation of G. Then, $V = \mathrm{Ind}_H^G W$ for some intermediate group $A \subset H \subset G$ and an irreducible representation of H whose restriction to A is* isotypical.

Recall that "W is isotypical for A" means that it is a direct sum of many copies of the same irreducible representation of A. Note that we may have $H = G$ and $W = V$, if its restriction to A is already isotypical. We are interested in the special case when A is abelian, in which case the fact that the irreducible reps are 1-dimensional implies that the isotypical representations are precisely the *scalar* ones.

17.7 Corollary. *With the same assumptions, if $A \subset G$ is abelian, then either its action on V is scalar, or else V is induced from a proper subgroup $H \subset G$.*

Proof of the Clever Lemma. Let $V = \bigoplus_i V_i$ be the isotypical decomposition of V with respect to A. If there is a single summand, we take $H = G$ as indicated. Otherwise, observe that the action of G on V must permute the blocks V_i. Indeed, denoting the action by ρ, choose $g \in G$ and let ${}^g\rho(a) := \rho(gag^{-1})$. Conjugation by g defines an automorphism of the (normal) subgroup A, which must permute its irreducible characters.

Then, $\rho(g) : V \to V$ defines an isomorphism of A-representations, if we let A act via ρ on the first space and via ${}^g\rho$ on the second. Moreover, the same decomposition $V = \bigoplus_i V_i$ is also the isotypical decomposition under the action ${}^g\rho$ of A. Of course, the irreducible type assigned to V_i has been changed, according to the action of g-conjugation on characters of A; but the fact that all irreducible summands within each V_i are isomorphic, and distinct from those in V_j, $j \neq i$, cannot have changed.

It follows that $\rho(g)$ preserves the blocks in the isotypical decomposition; reverting to the action ρ on the second space, we conclude that $\rho(g)$ permutes the blocks V_i according to its permutation action on irreducible characters of A.

Note, moreover, that this permutation action of G is *transitive*, (if we ignore zero blocks). Indeed, the sum of blocks within an orbit will be a G-invariant subspace of V, which was assumed irreducible. It follows that all non-zero blocks have the same dimension.

Choose now a (non-zero) block V_0 and let H be its stabiliser within G. From the orbit-stabiliser theorem, $\dim V = |G/H| \cdot \dim V_0$. I claim that $V \cong \mathrm{Ind}_H^G V_0$. Indeed, Frobenius reciprocity gives

$$\dim \mathrm{Hom}^G(\mathrm{Ind}_H^G V_0, V) = \dim \mathrm{Hom}^H(V_0, Res_G^H V) \geq 1,$$

the inequality holding because V_0 is an H-subrepresentation of V. However, V is irreducible under G and its dimension equals that of $\mathrm{Ind}_H^G V_0$, so any non-zero G-map between the two spaces is an isomorphism. □

17.8 Lemma (Little Lemma). *If G is nilpotent and non-abelian, then it contains a normal, abelian and non-central subgroup.*

Proof. Let $Z \subset G$ be the centre and let $G' \subset G$ be the smallest subgroup in the chain (17.1) not contained in Z. Then, G'/Z is central in G/Z, and so for any $g' \in G'$, $gg'g^{-1} \in g'Z$ for any $g \in G$. Thus, any such g' generates, together with Z, a subgroup of G with the desired properties. □

Proof of Theorem 17.5. Let $\rho : G \to V$ be an irreducible representation of the nilpotent group G. The subgroup $\ker\rho$ is normal and $G/\ker\rho$ is also nilpotent. If $G/\ker\rho$ is abelian, then the irreducible representation V must be one-dimensional and we are done. Else, let $A \subset G/\ker\rho$ be a normal, abelian, non-central subgroup. For a non-central element $a \in A$, $\rho(a)$ cannot be a scalar, or else it would commute with $\rho(g)$ for any $g \in$G, and then $\rho(gag^{-1}) = \mathrm{Id}$ for any g, so $gag^{-1} \in \ker\rho$ and then a would be central in $G/\ker\rho$.

It follows from Corollary 17.7 that there exists a proper subgroup $H \subset G$ with a representation W of $H/\ker\rho$, such that V is induced from $H/\ker\rho$ to $G/\ker\rho$. But that is also the representation $\mathrm{Ind}_H^G W$. The subgroup H is also nilpotent, and we can repeat the argument until we find a 1-dimensional W. □

(17.9) Example

A non-abelian exampleis the *Heisenberg group* of upper-triangular 3×3 matrices in $\mathbb{Z}/p$ with 1's on the diagonal. There are $(p-1)$ irreducible representations of dimension p, and they are all induced from 1-dimensional representations of the abelian subgroup of elements of the following form

$$\begin{bmatrix} 1 & * & * \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

18 Burnside's $p^a q^b$ theorem

(18.1) Additional comments

Induction from representations of "known" subgroups is an important tool in constructing the character table of a group. Frobenius' and Mackey's results can be used to (attempt to) break up induced reps into irreducibles. For arbitrary finite groups, there isn't a result as good as the one for p-groups — it need not be true that every irreducible representation is induced from a 1-dimensional one. However, the following theorem of Brauer's is a good substitute for that. I refer you to Serre's book for the proof, which is substantially more involved than that of Theorem 17.5.

If G is any finite group and $|G| = p^n m$, with m prime to p, then two theorems of Sylow assert the following:

1. G contains subgroups of order p^n;
2. all these subgroups are conjugate;
3. every p-subgroup of G is contained in one of these.

The subgroups in (1) are called the Sylow p-subgroups of G.

An *elementary p-subgroup* of G is one which decomposes as a direct product of a cyclic subgroup and a p-group. Up to conjugation, the elementary subgroups of G are of the form $\langle x\rangle \times S$, where $x \in G$ and S a subgroup of a fixed Sylow subgroup of the centraliser of x in G.

18.2 Theorem (Brauer). *The representation ring of G is spanned over $\mathbb{Z}$ by representations induced from 1-dimensional reps of elementary subgroups. In other words, every irreducible character is an integral linear combination of characters induced in this way.*

Knowing the integral span of characters within the space of class functions determines the irreducible characters, as the functions of square-norm 1 whose value at $1 \in G$ is positive. So, in principle, the irreducible characters can be found in finitely many steps, once the structure of the group is sufficiently well known.

(18.3) Burnside's theorem

One impressive application of the theory of characters to group theory is the following theorem of Burnside's, proved early in the 20th century. It was not given a purely group-theoretic proof until 1972. Let p, q be prime and a, b arbitrary natural numbers.

18.4 Theorem (Burnside). *Every group of order $p^a q^b$ is solvable.*

Noting that every subgroup and quotient group of such a group has order $p^{a'}q^{b'}$, it suffices to prove the following

18.5 Proposition. *Every group of order $p^a q^b \neq p, q$ contains a non-trivial normal subgroup.*

Indeed, we can continue applying the proposition to the normal subgroup and the quotient group, until we produce a chain where all subquotients have orders p or q (and thus are abelian). Groups containing no non-trivial normal subgroups are called *simple*, so we will prove that *the order of a simple group is contains at least three distinct prime factors*. The group A_5, of order 60, is simple, so this is the end of this road.

We will keep reformulating the desired property of our groups until we convert it into an integrality property of characters.

1. *Every non-trivial representation of a simple group is faithful.*
   (Indeed, the kernel must be trivial.)

2. *If some group element $g \neq 1$ acts as a scalar in a representation ρ, then either g is central or else ρ is not faithful. Either way, the group is not simple.*
   Indeed, $ugu^{-1} \in \ker \rho$ for any $u \in G$.

3. *With standard notations, $\rho(g)$ is a scalar iff $|\chi(g)| = \chi(1)$.*
   Indeed, $\chi(g)$ is a sum of its $\chi(1)$ eigenvalues, which are roots of unity. The triangle inequality for the norm of the sum is strict, unless all eigenvalues agree.

The following two facts now imply Proposition 18.5 and thus Burnside's theorem.

18.6 Proposition. *For any character χ of the finite group G, $|\chi(g)| = \chi(1)$ if and only if $\chi(g)/\chi(1)$ is a non-zero algebraic integer.*

18.7 Proposition. *If G has order $p^a q^b$, there exists an irreducible character χ and an element g such that $\chi(g)/\chi(1)$ is a non-zero algebraic integer.*

Proof of (18.6). Let d be the dimension of the representation; we have

$$\alpha := \frac{\chi(g)}{\chi(1)} = \frac{\omega_1 + \cdots + \omega_d}{d},$$

for suitable roots of unity ω_k. Now, the minimal polynomial of α over $\mathbb{Q}$ is

$$P_\alpha(X) = \prod\nolimits_\beta (X - \beta)$$

where β ranges over all distinct transforms of α under the action of the Galois group Γ of the algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$.[5] This is because P_α, having rational coefficients, is invariant under the action of Γ, and so it must admit all the β's as roots. On the other hand, the product above is clearly Γ-invariant, hence it has rational coefficients, and has α as a root; so it must agree with the minimal polynomial.

Gauss' Lemma asserts that any monic polynomial with integer coefficients is irreducible over $\mathbb{Q}$ as soon as it is so over $\mathbb{Z}$. Therefore, if α is an algebraic integer, then $P_\alpha(X)$ must have integer coefficients: indeed, being an algebraic integer, α must satisfy some monic equations with integral coefficients, among which the one of least degree must be irreducible over $\mathbb{Q}$, and hence must agree with the minimal polynomial P_α.

Now, if $\alpha \neq 0$, then no transform β can be zero. If so, the product of all β, which is integral, has modulus 1 or more. However, the Galois transform of a root of unity is also a root of unity, so every β is an average of roots of unity, and has modulus ≤ 1. We therefore get a contradiction from algebraic integrality, unless $\alpha = 0$ or $|\alpha| = 1$. $\square$

Proof of 18.7. We will find a conjugacy class $\{1\} \neq C \subset G$ whose order is a power of p, and an irreducible character χ, with $\chi(C) \neq 0$ and $\chi(1)$ not divisible by p. Assuming that, recall from Lecture 13 that the number $|C| \cdot \frac{\chi(C)}{\chi(1)}$ is an algebraic integer. As $|C|$ is a power of p, which does not divide the denominator, we are tempted to conclude that $\chi(C)/\chi(1)$ is already an algebraic integer. This would be clear from the prime factorisation, if $\chi(C)$ were an actual integer. As it is, we must be more careful and argue as follows.

Bézout's theorem secures the existence of integers m, n such that $m|C| + n\chi(1) = 1$. Then,

$$m|C|\frac{\chi(C)}{\chi(1)} + n\chi(C) = \frac{\chi(C)}{\chi(1)},$$

displaying $\chi(C)/\chi(1)$ as a sum of algebraic integers.

It remains to find the desired C and χ. Choose a q-Sylow subgroup $H \subset G$; this has non-trivial centre, as proved in Proposition 17.3. The centraliser of a non-trivial element $g \subset Z(H)$ contains H, and so the conjugacy class C of g has p-power order, from the orbit-stabiliser theorem. This gives our C.

The column orthogonality relations of the characters, applied to C and $\{1\}$, give

$$1 + \sum_{\chi \neq 1} \chi(C)\chi(1) = 0,$$

where the χ in the sum range over the non-trivial irreducible characters. Each $\chi(C)$ is an algebraic integer. If every $\chi(1)$ for which $\chi(C) \neq 0$ were divisible by p, then dividing the entire equation by p would show that $1/p$ was an algebraic integer, contradiction. So there exists a χ with $\chi(C) \neq 0$ and $\chi(1)$ not divisible by p, as required. $\square$

[5] It suffices to consider the extension field generated by the roots of unity appearing in α.

19 Topological groups

In studying the finite-dimensional complex representations of a finite group G, we saw that

1. they are all unitarisable;
2. they are completely reducible;
3. the characters of irreducible representations form an orthonormal basis of the space of class functions;
4. Representations are determined by their characters, up to isomorphisms;
5. The character map $\chi \to R_G$ is a ring homomorphism, compatible with involutions, $\chi \leftrightarrows \bar{\chi}$ corresponds to $V \leftrightarrows V^*$, and with inner products, $\langle \chi_V | \chi_W \rangle = \langle V | W \rangle := \dim \mathrm{Hom}^G(V; W)$.

Note that $1 \Rightarrow 2$ and $3 \Rightarrow 4$; part (5) is obvious from the definition of characters, save for the inner product bit, which follows from (1) and (3).

We will now study the representation theory of infinite *compact* groups, and will see that the same properties (1)–(5) hold. Point (3) will be amended to account for the fact that the space of class functions is *infinite-dimensional*. Correspondingly, there will be infinitely many isomorphism classes, but their characters will still form a *complete orthonormal set* in the (Hilbert) space of class functions. (This is also called a Hilbert space basis.)

There is a good general theory for all compact groups. In the special case of compact *Lie* groups — the groups of isometries of geometric objects, such as $\mathrm{SO}(n), \mathrm{SU}(n), \mathrm{U}(n)$ — the theory is due to Hermann Weyl[6] and is outrageously successful: unlike the finite group case, we can list all irreducible characters in explicit form! That is the *Weyl character formula*. However, the general theory requires two deeper theorems of analysis (existence of the *Haar measure* and the *Peter-Weyl theorem*, see below), that go a bit beyond our means. For simple groups such as $U(1)$, $SU(2)$ and related ones, these results are quite easy to establish directly, and so we shall study them completely rigorously. It is not much more difficult to handle $U(n)$ for general n, but we only have time for a quick survey.

19.1 Definition. A *topological group* is a group which is also a topological space, and for which the group operations are continuous. It is called *compact* if it is so as a topological space. A representation of a topological group G on a finite-dimensional vector space V is a *continuous* group homomorphism $\rho : G \to \mathrm{GL}(V)$, with the topology of $\mathrm{GL}(V)$ inherited from the space $\mathrm{End}(V)$ of linear self-maps.

19.2 Remark. One can define continuity for infinite-dimensional representations, but more care is needed when defining $\mathrm{GL}(V)$ and its topology. Typically, one starts with some topology on V, then one chooses a topology on the space $\mathrm{End}(V)$ of continuous linear operators (for instance, the norm topology on bounded operators, when V is a Hilbert space) and then one embeds $\mathrm{GL}(V)$ as a closed subset of $\mathrm{End}(V) \times \mathrm{End}(V)$ by the map $g \mapsto (g, g^{-1})$. This defines the structure of a topological group on $\mathrm{GL}(V)$.

(19.3) Orthonormality of characters

In establishing the key properties (1) and (3) above, the two ingredients were Schur's Lemma and Weyl's trick of averaging over G. The first one applies without change to any group. The second must be replaced by *integration over G*. Here, we take advantage of the continuity assumption in our representations, which ensures that all intervening functions in the argument are continuous on G, and can be integrated.

Even so, the key properties of the average over G were its invariance under left and right translation, and the fact that G has total volume 1 (so that the average of the constant function

[6] The unitary case was worked out by Schur.

1 is 1). The second property can always be fixed by normalising the volume, but the first imposed a strong constraint on the volume form we use for integration: the volume element dg on the group must be left- and right invariant. A (difficult) theorem of Haar asserts that the two constraints determine the volume element uniquely (subject to a reasonable technical constraint: specifically, any compact, Hausdorff group carries a unique bi-invariant regular Borel measure of total mass 1). For Lie groups, an easier proof of existence can be given; we'll see that for U(1) and SU(2).

Armed with an invariant integration measure, the main results and their consequences in Lecture 8 follow as before, replacing $\frac{1}{|G|}\sum_{g\in G}$ in the arguments with $\int_G dg$.

19.4 Proposition. *Every finite-dimensional representation of a compact group is unitarisable, and hence completely reducible.*

19.5 Theorem. *Characters of irreducible representations of a compact group have unit L^2 norm, and characters of non-isomorphic irreducibles are orthogonal.*

19.6 Corollary. *A representation is irreducible iff its character has norm* 1*.*

Completeness of characters is a different matter. Its formulation now is that the characters form a Hilbert space *basis* of the space of square-integrable class functions. In addition to the orthogonality properties, this asserts that any continuous class function which is orthogonal to every irreducible character is null. This asserts the existence of an "ample supply" of irreducible representations, and so it cannot be self-evident.

In the case of finite groups, we found enough representations by studying the regular representations. This method is also applicable to the compact case, but relies here on a fundamental theorem that we state without proof:

19.7 Theorem (Peter-Weyl). *The space of square-integrable functions on G is the Hilbert space direct sum over finite-dimensional irreducible representations V,*

$$L^2(G) \cong \hat{\bigoplus} \operatorname{End}(V).$$

The map from left to right sends a function f on G to the operators $\int_G f(g)\cdot\rho_V(g)dg$. The inverse map sends $\phi \in \operatorname{End}(V)$ to the function $g \mapsto \operatorname{Tr}_V(\rho_V(g)^*\phi)$. The inner product on $\operatorname{End}(V)$ corresponding to L^2 on G is given by $\langle\phi|\psi\rangle = \operatorname{Tr}_V(\phi^*\psi)\cdot\dim V$.

As opposed to the case of finite groups, the proof for compact groups is somewhat involved. (Lie groups allow for an easier argument.) However, we shall make no use of the Peter-Weyl theorem here; in the cases we study, the completeness of characters will be established by direct construction of representations.

(19.8) A compact group: U(1)

To illustrate the need for a topology and the restriction to continuous representations, we start off by ignoring these aspects. As an abelian group, U(1) is isomorphic to $\mathbb{R}/\mathbb{Z}$ via the map $x \mapsto \exp(2\pi i x)$. Now $\mathbb{R}$ is a vector space over $\mathbb{Q}$, and general nonsense (Hamel's basis theorem) asserts that is must have a basis $A = \{\alpha\} \subset \mathbb{R}$ as a $\mathbb{Q}$-vector space; moreover, we can choose our first basis element to be 1. Thus, we have isomorphisms of abelian groups,

$$\mathbb{R} \cong \mathbb{Q} \oplus \bigoplus_{\alpha\neq 1} \mathbb{Q}\alpha, \qquad \mathbb{R}/\mathbb{Z} \cong \mathbb{Q}/\mathbb{Z} \times \bigoplus_{\alpha\neq 1} \mathbb{Q}\alpha.$$

It's easy to see that the choice of an integer n and of a complex number z_α for each $A \ni \alpha \neq 1$ defines a one-dimensional complex representation of the second group, by specifying

$$(x, x_\alpha) \mapsto \exp(2\pi i n x)\cdot\prod_{\alpha\neq 1}\exp(2\pi i z_\alpha \cdot x_\alpha).$$

Notice that, for every element in our group, all but finitely many coordinates x_α must vanish, so that almost all factors in the product are 1.

A choice of complex number for each element of a basis of $\mathbb{R}$ over $\mathbb{Q}$ is not a sensible collection of data; and it turns out this does not even succeed in covering all 1-dimensional representations of $\mathbb{R}/\mathbb{Z}$. Things will improve dramatically when we impose the *continuity* requirement; we will see in the next lecture that the only surviving datum is the integer n.

19.9 Remark. The subgroup $\mathbb{Q}/\mathbb{Z}$ turns out to have a more interesting representation theory, even when continuity is ignored. The one-dimensional representations are classified by the *profinite completion* $\widehat{\mathbb{Z}}$ of $\mathbb{Z}$.

Henceforth, "representation" of a topological group will mean continuous representation. We now classify the finite-dimensional representations of $\mathrm{U}(1)$. The character theory is closely tied to the theory of Fourier series.

19.10 Theorem. *A continuous* 1*-dimensional representation* $\mathrm{U}(1) \to \mathbb{C}^\times$ *has the form* $z \mapsto z^n$*, for some integer* n*.*

For the proof, we will need the following

19.11 Lemma. *The closed proper subgroups of* $\mathrm{U}(1)$ *are the cyclic subgroups* μ_n *of* n*th roots of unity* $(n > 1)$*.*

Proof of the Lemma. If $q \in \mathrm{U}(1)$ is not a root of unity, then its powers are dense in $\mathrm{U}(1)$ (exercise, using the pigeonhole principle). So, any closed, proper subgroup of $\mathrm{U}(1)$ consists only of roots of 1. Among those, there must be one of smallest argument (in absolute value) or else there would be a sequence converging to 1; these would again generate a dense subgroup of $\mathrm{U}(1)$. The root of unity of smallest argument is then the generator. $\square$

Proof of Theorem 19.10. A continuous map $\varphi : \mathrm{U}(1) \to \mathbb{C}^\times$ has compact, hence bounded image. The image must lie on the unit circle, because the integral powers of any other complex number form an unbounded sequence. So φ is a continuous homomorphism from $\mathrm{U}(1)$ itself.

Now, $\ker\varphi$ is a closed subgroup of $\mathrm{U}(1)$. If $\ker\varphi$ is the entire group, then $\varphi \cong 1$ and the theorem holds with $n = 0$. If $\ker\varphi = \mu_n$ $(n \geq 1)$, we will now show that $\varphi(z) = z^{\pm n}$, with the same choice of sign for all z. To see this, define a *continuous* function

$$\psi : [0, 2\pi/n] \to \mathbb{R}, \qquad \psi(0) = 0, \qquad \psi(\theta) = \arg\varphi(e^{\mathrm{i}\theta});$$

in other words, we parametrise $\mathrm{U}(1)$ by the argument θ, start with $\psi(0) = 0$, which is one value of the argument of $\varphi(1) = 1$, choose the argument so as to make the function continuous.

Because $\ker\varphi = \mu_n$, ψ must be *injective* on $[0, 2\pi/n)$. By continuity, it must be monotonically increasing or decreasing (Intermediate Value Theorem), and we must have $\psi(2\pi/n) = \pm 2\pi$: the value zero is ruled out by monotonicity and any other multiple of 2π would lead to an intermediate value of θ with $\varphi(e^{\mathrm{i}\theta}) = 1$. Henceforth, $\pm$ denotes the sign of $\psi(2\pi/n)$.

Because φ is a homomorphism and $\varphi(e^{2\pi\mathrm{i}/n}) = 1$, $\varphi(e^{2\pi\mathrm{i}/mn})$ must be an mth root of unity, and so $\psi(\{2\pi k/mn\}) \subset \{\pm 2\pi k/m\}$, $k = 0, \ldots, m$. By monotonicity, these $m + 1$ values must be taken exactly once and in the natural order, so $\psi(2\pi k/mn) = \pm 2\pi k/m$, for all m and all $k = 0, \ldots, m$. But then, $\psi(\theta) = \pm n \cdot \theta$, by continuity, and $\varphi(z) = z^{\pm n}$, as claimed. $\square$

19.12 Remark. Complete reducibility now shows that continuous finite-dimensional representations of $\mathrm{U}(1)$ are in fact *algebraic*: that is, the entries in a matrix representations are Laurent polynomials in z. This is not an accident; it holds for a large class of topological groups (the compact Lie groups).

(19.13) Character theory

The homomorphisms $\rho_n : \mathrm{U}(1) \to \mathbb{C}^\times$, $\rho_n(z) = z^n$ form the complete list of irreducible representations of U(1). Clearly, their characters (which we abusively denote by the same symbol) are linearly independent; in fact, they are orthonormal in the inner product

$$\langle\varphi|\psi\rangle = \frac{1}{2\pi}\int_0^{2\pi} \overline{\varphi(\theta)}\psi(\theta)d\theta,$$

which corresponds to "averaging over U(1)" ($z = e^{\mathrm{i}\theta}$). The functions ρ_n are sometimes called *Fourier modes*, and their finite linear combinations are the *Fourier polynomials.*

As U(1) is abelian, it coincides with the space of its conjugacy classes. We can now state our main theorem.

19.14 Theorem. *(i) The functions ρ_n form a complete list of irreducible characters of* U(1). *(ii) Every finite-dimensional representation V of* U(1) *is isomorphic to a sum of the ρ_n. Its character χ_V is a Fourier polynomial. The multiplicity of ρ_n in V is the inner product $\langle\rho_n|\chi_V\rangle$.*

Recall that complete reducibility of representations follows by Weyl's unitary trick, averaging any given inner product by integration on U(1).

As the space of (continuous) class functions is infinite-dimensional, it requires a bit of care to state the final part of our main theorem, that the characters form a "basis" of the space of class functions. The good setting for that is the *Hilbert space of square-integrable functions*, which we discuss below. For now, let us just note an algebraic version of the result.

19.15 Proposition. *The ρ_n form a basis of the polynomial functions on* $\mathrm{U}(1) \subset \mathbb{R}^2$.

Indeed, on the unit circle, $\bar{z} = z^{-1}$ so the the Fourier polynomials are also the polynomials in z and $\bar{z}$, which can also be expressed as polynomials in x, y.

(19.16) Digression: Fourier series

The Fourier modes ρ_n form a complete orthonormal set (orthonormal basis) of the *Hilbert space* $L^2(\mathrm{U}(1))$. This means that every square-integrable function $f \in L^2$ has a (Fourier) series expansion

$$f(\theta) = \sum_{-\infty}^{\infty} f_n \cdot e^{\mathrm{i}n\theta} = \sum_{-\infty}^{\infty} f_n \cdot \rho_n, \tag{*}$$

which converges in *mean square*; the *Fourier coefficients* f_n are given by the formula

$$f_n = \langle\rho_n|f\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{-\mathrm{i}n\theta} f(\theta)d\theta. \tag{**}$$

Recall that mean-square convergence signifies that the partial sums approach f, in the distance defined by the inner product. The proof of most of this is fairly easy. Orthonormality of the ρ_n implies the Fourier expansion formula for the *Fourier polynomials*, the finite linear combinations of the ρ_n. (Of course, in this case the Fourier series is a finite sum, and no analysis is needed.) Furthermore, for any given finite collection S of indexes n, and for any $f \in L^2$, the sum of terms in (*) with $n \in S$ is the orthogonal projection of f onto the the span of the ρ_n, $n \in S$. Hence, any finite sum of norms of the Fourier coefficients $\sum \|f_n\|^2$ in (*) is bounded by $\|f\|^2$, and the Fourier series converges. Moreover, the limit g has the property that $f - g$ is orthogonal to each ρ_n; in other words, g is the projection of f onto the span of all the ρ_n.

The more difficult part is to show that any $f \in L^2$ which is orthogonal to all ρ_n vanishes. One method is to derive it from a powerful theorem of Weierstraß', which says that the Fourier polynomials are dense in the space of continuous functions, in the sense of uniform convergence.

(The result holds for continuous functions on any compact subset of $\mathbb{R}^N$; here, $N = 2$.) Approximating a candidate f, orthogonal to all ρ_n, by a sequence of Fourier polynomials p_k leads to a contradiction, because

$$\|f - p_k\|^2 = \|f\|^2 + \|p_k\|^2 \geq \|f\|^2,$$

by orthogonality, and yet uniform convergence certainly implies convergence in mean-square.

We summarise the basic facts in the following

19.17 Theorem. *(i) Any continuous function on* $\mathrm{U}(1)$ *can be uniformly approximated by finite linear combinations of the* ρ_n.
(ii) Any square-integrable function $f \in L^2(\mathrm{U}(1))$ *has a series expansion* $f = \sum f_n \cdot \rho_n$*, with Fourier coefficients* f_n *given by (**).*

20 The group SU(2)

We move on to the group SU(2). We will describe its conjugacy classes, find the bi-invariant volume form, whose existence implies the unitarisability and complete reducibility of its continuous finite-dimensional representations. In the next lecture, we will list the irreducibles with their characters.

Giving away the plot, observe that SU(2) acts on the space P of polynomials in two variables z_1, z_2, by its natural linear action on the coordinates. We can decompose P into the homogeneous pieces P_n, $n = 0, 1, 2, \ldots$, preserved by the action of SU(2): thus, the space P_0 of constants is the trivial one-dimensional representation, the space P_1 of linear functions the standard 2-dimensional one, etc. Then, $P_0, P_1, P_2, \ldots$ is the complete list of irreducible representations of SU(2), up to isomorphism, and every continuous finite-dimensional representation is isomorphic to a direct sum of those.

(20.1) SU(2) *and the quaternions*

By definition, SU(2) is the group of complex 2×2 matrices preserving the complex inner product and with determinant 1; the group operation is matrix multiplication. Explicitly,

$$\mathrm{SU}(2) = \left\{ \begin{bmatrix} u & v \\ -\bar{v} & \bar{u} \end{bmatrix} : u, v \in \mathbb{C}, |u|^2 + |v|^2 = 1 \right\}.$$

Geometrically, this can be identified with the three-dimensional unit sphere in $\mathbb{C}^2$. It is useful to replace $\mathbb{C}^2$ with Hamilton's quaternions $\mathbb{H}$, generated over $\mathbb{R}$ by elements i, j satisfying the relations $\mathrm{i}^2 = \mathrm{j}^2 = -1$, $\mathrm{ij} = -\mathrm{ji}$. Thus, $\mathbb{H}$ is four-dimensional, spanned by $1, \mathrm{i}, \mathrm{j}$ and $\mathrm{k} := \mathrm{ij}$. The *conjugate* of a quaternion $q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}$ is $\bar{q} := a - b\mathrm{i} - c\mathrm{j} - d\mathrm{k}$, and the *quaternion norm* is

$$\|q\|^2 = q\bar{q} = \bar{q}q = a^2 + b^2 + c^2 + d^2.$$

The relation $\overline{q_1 q_2} = \bar{q}_2 \bar{q}_1$ shows that $\|\| : \mathbb{H} \to \mathbb{R}$ is multiplicative,

$$\|q_1 q_2\| = q_1 q_2 \cdot \overline{q_1 q_2} = q_1 q_2 \cdot \bar{q}_2 \bar{q}_1 = \|q_1\| \|q_2\|$$

in particular, the "unit quaternions" (the quaternions of unit norm) form a group under multiplication. Direct calculation establishes the following.

20.2 Proposition. *Sending* $\begin{bmatrix} u & v \\ -\bar{v} & u \end{bmatrix}$ *to* $q = u + v\mathrm{j}$ *gives an isomorphism of* SU(2) *with the multiplicative group of unit quaternions.* □

(20.3) Conjugacy classes

A theorem from linear algebra asserts that unitary matrices are diagonalisable in an orthonormal eigenbasis. The diagonalised matrix is then also unitary, so its diagonal elements are complex numbers of unit norm. These numbers, of course, are the eigenvalues of the matrix. They are only determined up to reordering. We therefore get a bijection of conjugacy classes in the unitary group with unordered N-tuples of complex numbers of unit norm:

$$\mathrm{U}(N)/\mathrm{U}(N) \leftrightarrow \mathrm{U}(1)^N/S_N,$$

where the symmetric group S_N acts by permuting the N factors of $U(1)^N$. The map from right to left is defined by the inclusion $\mathrm{U}(1)^N \subset \mathrm{U}(N)$ and is therefore continuous; a general theorem from topology ensures that it is in fact a homeomorphism.[7]

Clearly, restricting to $\mathrm{SU}(N)$ imposed the $\det = 1$ restriction on the right side $U(1)^N$. However, there is a potential problem in sight, because it might happen, in principle, that two matrices in $\mathrm{SU}(N)$ will be conjugate in $\mathrm{U}(N)$, but not in $\mathrm{SU}(N)$. For a general situation of a subgroup $H \subset G$ this warrants deserves careful consideration. In the case at hand, however, we are lucky. The scalar matrices $\mathrm{U}(1) \cdot \mathrm{Id} \subset \mathrm{U}(N)$ are central, that is, their conjugation action is trivial. Now, for any matrix $A \in \mathrm{U}(N)$ and any Nth root r of $\det A$, we have $r^{-1}A \in \mathrm{SU}(N)$, and conjugating by the latter has the same effect as conjugation by A. We then get a bijection

$$\mathrm{SU}(N)/S\mathrm{U}(N) \leftrightarrow \mathrm{S}\left(\mathrm{U}(1)^N\right)/S_N,$$

where $\mathrm{U}(1)^N$ is identified with the group of diagonal unitary $N \times N$ matrices and the S on the right, standing for "special", indicates its subgroup of determinant 1 matrices.

20.4 Proposition. *The normalised trace* $\frac{1}{2}\mathrm{Tr} : \mathrm{SU}(2) \to \mathbb{C}$ *gives a homomorphisms of the set of conjugacy classes in* $\mathrm{SU}(2)$ *with the interval* $[-1, 1]$.

The set of conjugacy classes is given the quotient topology.

Proof. As discussed, matrices are conjugate in $\mathrm{SU}(2)$ iff their eigenvalues agree up to order. The eigenvalues form a pair $\{z, z^{-1}\}$, with z on the unit circle, so they are the roots of

$$X^2 - (z + z^{-1})X + 1,$$

in which the middle coefficient ranges over $[-2, 2]$. $\square$

The quaternion picture provides a good geometric model of this map: the half-trace becomes the projection $q \mapsto a$ on the real axis $\mathbb{R} \subset \mathbb{H}$. The conjugacy classes are therefore the 2-*dimensional spheres of constant latitude on the unit sphere*, plus the two poles, the singletons $\{\pm 1\}$. The latter correspond to the central elements $\pm I_2 \in \mathrm{SU}(2)$.

(20.5) Characters as Laurent polynomials

The preceding discussion implies that the characters of continuous $\mathrm{SU}(2)$-representations are continuous functions on $[-1, 1]$. It will be more profitable, however, to parametrise that interval by the "latitude" $\phi \in [0, \pi]$, rather than the real part $\cos\phi$. We can in fact let ϕ range from $[0, 2\pi]$, provide we remember our functions are invariant under $\phi \leftrightarrow -\phi$ and periodic with period 2π. We call such functions of ϕ *Weyl-invariant*, or *even.* Functions which change sign under that same symmetry are *anti-invariant* or *odd.*

[7] We use the quotient topologies on both sides. The result I am alluding to asserts that a continuous bijection from a compact space to a Hausdorff space is a homeomorphism.

Notice that ϕ parametrises the subgroup of diagonal matrices, isomorphic to U(1),

$$\begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix} = \begin{bmatrix} z & 0 \\ 0 & z^{-1} \end{bmatrix} \in \mathrm{SU}(2),$$

and the transformation $\phi \leftrightarrow -\phi$ is induced by conjugation by the matrix $\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \in \mathrm{SU}(2)$. The character of an SU(2)-representation V will be the function of $z \in \mathrm{U}(1)$

$$\chi_V(z) = \mathrm{Tr}_V\left(\begin{bmatrix} z & 0 \\ 0 & z^{-1} \end{bmatrix}\right),$$

invariant under $z \leftrightarrow z^{-1}$. When convenient, we shall re-express it in terms of ϕ and abusively write $\chi_V(\phi)$.

A representation of SU(2) restricts to one of our U(1), and we know from last lecture that the characters of the latter are polynomials in z and z^{-1}. Such functions are called *Laurent polynomials.* We therefore obtain the

20.6 Proposition. *Characters of* SU(2)*-representations are even Laurent polynomials in z.* □

(20.7) The volume form and Weyl's Integration formula for SU(2)

For finite groups, *averaging over the group* was used to prove complete reducibility, to define the inner product of characters and prove the orthogonality theorems. For compact groups (such as U(1)), integration over the group must be used instead. In both cases, the essential property of the operation is its *invariance* under left and right translations on the group. To state this more precisely: the linear functional sending a continuous function φ on G to $\int_G \varphi(g)dg$ is invariant under left and right translations of φ. A secondary property (which is arranged by appropriate scaling) is that the average or integral of the constant function 1 is 1.

We thus need a volume form over SU(2) which is invariant under left and right translations. Our geometric model for SU(2) as the unit sphere in $\mathbb{R}^4$ allows us to find it directly, without appealing to Haar's (difficult) general result. Note, indeed, that the actions of SU(2) on $\mathbb{H}$ by left and right multiplication preserve the quaternion norm, hence the Euclidean distance. In particular, the usual Euclidean volume element on the unit sphere must be invariant under both left and right multiplication by SU(2) elements.

20.8 Remark. The right×left action of $\mathrm{SU}(2) \times \mathrm{SU}(2)$ on $\mathbb{R}^4 = \mathbb{H}$, whereby $\alpha \times \beta$ sends $q \in \mathbb{H}$ to $\alpha q \beta^{-1}$, gives a homomorphism $h : \mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$. (The action preserves orientation, because the group is connected.) Clearly, $(-\mathrm{Id}, -\mathrm{Id})$ acts as the identity, so h factors through the quotient $\mathrm{SU}(2) \times \mathrm{SU}(2)/\{\pm(\mathrm{Id}, \mathrm{Id})\}$, and indeed turns out to give an *isomorphism* of the latter with SO(4). We will discuss this in Lecture 22.

To understand the inner product of characters, we are interested in integrating *class functions* over SU(2). This must be expressible in terms of their restriction to U(1). This is made explicit by the following theorem, whose importance cannot be overestimated: as we shall see, it implies the character formulae for all irreducible representations of SU(2). Let dg be the Euclidean volume form on the unit sphere in $\mathbb{H}$, normalised to total volume 1, and let $\Delta(\phi) = e^{i\phi} - e^{-i\phi}$ be the *Weyl denominator.*

20.9 Theorem (Weyl Integration Formula). *For a continuous class function f on* SU(2)*, we have*

$$\int_{\mathrm{SU}(2)} f(g)dg = \frac{1}{2}\frac{1}{2\pi}\int_0^{2\pi} f(\phi) \cdot |\Delta(\phi)|^2 d\phi = \frac{1}{\pi}\int_0^{2\pi} f(\phi)\sin^2\phi\, d\phi.$$

Thus, the integral over SU(2) can be computed by restriction to the U(1) of diagonal matrices, after correcting the measure by the factor $\frac{1}{2}|\Delta(\phi)|^2$. We are abusively writing $f(\phi)$ for the value of f at $\begin{bmatrix} e^{\mathrm{i}\phi} & 0 \\ 0 & e^{-\mathrm{i}\phi} \end{bmatrix}$.

Proof. In the presentation of SU(2) as the unit sphere in $\mathbb{H}$, the function f being constant on the spheres of constant latitude. The volume of a spherical slice of width $d\phi$ is $C \cdot \sin^2 \phi \, d\phi$, with the constant C normalised by the "total volume 1" condition

$$\int_0^\pi C \cdot \sin^2 \phi \, d\phi = 1,$$

whence $C = 2/\pi$, in agreement with the theorem. $\square$

21 Irreducible characters of SU(2)

Let us check the irreducibility of some representations.

- The trivial 1-dimensional representation is obviously irreducible; its character has norm 1, due to our normalisation of the measure.
- The character of the standard representation on $\mathbb{C}^2$ is $e^{\mathrm{i}\phi} + e^{-\mathrm{i}\phi} = 2\cos\phi$. Its norm is

$$\frac{4}{\pi}\int_0^{2\pi} \cos^2\phi \cdot \sin^2\phi \, d\phi = \frac{1}{\pi}\int_0^{2\pi} (2\cos\phi\sin\phi)^2 \, d\phi = \frac{1}{\pi}\int_0^{2\pi} \sin^2 2\phi \, d\phi = 1,$$

  so $\mathbb{C}^2$ is irreducible. Of course, this can also be seen directly (no invariant lines).
- The character of the tensor square $\mathbb{C}^2 \otimes \mathbb{C}^2$ is $4\cos^2\phi$, so its square norm is

$$\begin{aligned} \frac{1}{\pi}\int_0^{2\pi} 16\cos^4\phi\sin^2\phi \, d\phi &= \frac{1}{\pi}\int_0^{2\pi} 4\cos^2\phi\sin^2 2\phi \, d\phi \\ &= \frac{1}{\pi}\int_0^{2\pi} (\sin 3\phi + \sin\phi)^2 \, d\phi \\ &= \frac{1}{\pi}\int_0^{2\pi} \left(\sin^2 3\phi + \sin^2\phi + 2\sin\phi\sin 3\phi\right) \, d\phi = 1 + 1 = 2, \end{aligned}$$

  so this is reducible. Indeed, we have $(\mathbb{C}^2)^{\otimes 2} = \mathrm{Sym}^2\mathbb{C}^2 \oplus \Lambda^2\mathbb{C}^2$; these must be irreducible. $\Lambda^2\mathbb{C}^2$ is the trivial 1-dim. representation, since its character is 1 (the product of the eigenvalues), and so $\mathrm{Sym}^2\mathbb{C}^2$ is a new, 3-dimensional irreducible representation, with character $2\cos 2\phi + 1 = z^2 + z^{-2} + 1$.

So far we have found the irreducible representations $\mathbb{C} = \mathrm{Sym}^0\mathbb{C}^2$, $\mathbb{C}^2 = \mathrm{Sym}^1\mathbb{C}^2$, $\mathrm{Sym}^2\mathbb{C}^2$, with characters 1, $z + z^{-1}$, $z^2 + 1 + z^{-2}$. This surely lends credibility to the following.

21.1 Theorem. *The character χ_n of $\mathrm{Sym}^n\mathbb{C}^2$ is $z^n + z^{n-2} + \ldots + z^{2-n} + z^{-n}$. Its norm is 1, and so each $\mathrm{Sym}^n\mathbb{C}^2$ is an irreducible representation of SU(2). Moreover, this is the complete list of irreducibles, as $n \geq 0$.*

21.2 Remark. Note that $\mathrm{Sym}^n\mathbb{C}^2$ has dimension $n + 1$.

Proof. The standard basis vectors $\mathbf{e}_1, \mathbf{e}_2$ scale under $g = \begin{bmatrix} z & 0 \\ 0 & z^{-1} \end{bmatrix}$ by the factors $z^{\pm 1}$. Now, the basis vectors of $\mathrm{Sym}^n\mathbb{C}^2$ arise by symmetrising the vectors $\mathbf{e}_1^{\otimes n}, \mathbf{e}_1^{\otimes(n-1)} \otimes \mathbf{e}_2, \ldots, \mathbf{e}_2^{\otimes n}$ in the nth tensor power $(\mathbb{C}^2)^{\otimes n}$. The vectors appearing in the symmetrisation of $\mathbf{e}_1^{\otimes p} \otimes \mathbf{e}_2^{\otimes(n-p)}$ are tensor products containing p factors of $\mathbf{e}_1$ and $n-p$ factors of $\mathbf{e}_2$, in some order; each of these is an eigenvectors for g, with eigenvalue $z^p \cdot z^{n-p}$. Hence, the eigenvalues of g on Sym^n are $\{z^n, z^{n-2}, \ldots, z^{-n}\}$ and the character formula follows.

We have

$$\chi_n(z) = \frac{z^{n+1} - z^{-n-1}}{z - z^{-1}},$$

whence $\chi_n(z)|^2|\Delta(z)|^2 = |z^{n+1} - z^{-n-1}|^2 = 4\sin^2[(n+1)\phi]$. Thus,

$$\|\chi_n\|^2 = \frac{1}{4\pi}\int_0^{2\pi} 4\sin^2[(n+1)\phi]\, d\phi = \frac{1}{\pi} \cdot \pi = 1,$$

proving irreducibility.

To see completeness, observe that the functions $\Delta(z)\chi_n(z)$ span the vector space of odd Laurent polynomials (§20.5). For any character χ, the function $\chi(z)\Delta(z)$ is an odd Laurent polynomial. If χ was a new irreducible character, $\chi(z)\Delta(z)$ would have to be orthogonal to all anti-invariant Laurent polynomials, in the standard inner product on the circle. (Combined with the weight $|\Delta|^2$, this is the correct SU(2) inner product up to scale, as per the integration formula). But this is impossible, as the orthogonal complement of the odd polynomials is spanned by the even ones. $\square$

(21.3) Another look at the irreducibles

The group SU(2) acts linearly on the vector space P of polynomials in z_1, z_2 by transforming the variables in the standard way,

$$\begin{bmatrix} u & v \\ -\bar{v} & \bar{u} \end{bmatrix} \cdot \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} uz_1 + vz_2 \\ -\bar{v}z_1 + \bar{u}z_2 \end{bmatrix}.$$

This preserves the decomposition of P as the direct sum of the spaces P_n of *homogeneous* polynomials of degree n.

21.4 Proposition. *As* SU(2)*-representations,* $\mathrm{Sym}^n\mathbb{C}^2 \cong P_n$*, the space of homogeneous polynomials of degree n in two variables.*

Proof. An isomorphism from P_n to $\mathrm{Sym}^n\mathbb{C}^2$ is defined by sending $z_1^p z_2^{n-p}$ to the symmetrisation of $\mathbf{e}_1^{\otimes p} \otimes \mathbf{e}_2^{\otimes(n-p)}$. By linearity, this will send $(uz_1 + vz_2)^p \cdot (-\bar{v}z_1 + \bar{u}z_2)^{n-p}$ to the symmetrisation of $(u\mathbf{e}_1 + v\mathbf{e}_2)^{\otimes p} \otimes (-\bar{v}\mathbf{e}_1 + \bar{u}\mathbf{e}_2)^{\otimes(n-p)}$. This shows that our isomorphism commutes with the SU(2)-action. $\square$

21.5 Remark. Equality of the representations up to isomorphism can also be seen by checking the character of a diagonal matrix g, for which the $z_1^p z_2^{n-p}$ are eigenvectors.

(21.6) Tensor product of representations

Knowledge of the characters allows us to find the decomposition of tensor products. Obviously, tensoring with $\mathbb{C}$ does not change a representation. The next example is

$$\chi_1 \cdot \chi_n = (z + z^{-1})\frac{z^{n+1} - z^{-n-1}}{z - z^{-1}} = \frac{z^{n+2} + z^n - z^{-n} - z^{-2-n}}{z - z^{-1}} = \chi_{n+1} + \chi_{n-1},$$

provided $n \geq 1$; if $n = 0$ of course the answer is χ_1. More generally, we have

21.7 Theorem. *If* $p \geq q$, $\chi_p \cdot \chi_q = \chi_{p+q} + \chi_{p+q-2} + \ldots + \chi_{p-q}$. *Hence,*

$$\mathrm{Sym}^p\mathbb{C}^2 \otimes \mathrm{Sym}^q\mathbb{C}^2 \cong \bigoplus_{k=0}^{q} \mathrm{Sym}^{p+q-2k}\,\mathbb{C}^2.$$

Proof.

$$\begin{aligned}\chi_p(z)\cdot\chi_q(z) &= \frac{z^{p+1} - z^{-p-1}}{z - z^{-1}} \cdot \left(z^q + z^{q-2} + \ldots + z^{-q}\right) \\ &= \sum_{k=0}^{q} \frac{z^{p+q+1-2k} - z^{2k-p-q-1}}{z - z^{-1}} \\ &= \sum_{k=0}^{q} \chi_{p+q-2k}\ ;\end{aligned}$$

the condition $p \geq q$ ensures that there are no cancellations in the sum. □

21.8 Remark. Let us define formally $\chi_{-n} = -\chi_{n-2}$, a "virtual character". Thus, $\chi_0 = 1$, the trivial character, $\chi_{-1} = 0$ and $\chi_{-2} = -1$. The decomposition formula becomes

$$\chi_p \cdot \chi_q = \chi_{p+q} + \chi_{p+q-2} + \ldots + \chi_{p-q},$$

regardless whether $p \geq q$ or not.

(21.9) Multiplicities

The multiplicity of the irreducible representation Sym^n in a representation with character $\chi(z)$ can be computed as the inner product $\langle \chi | \chi_n \rangle$, with the integration (20.9). The following recipe, which exploits the simple form of the functions $\Delta(z)\chi_n(z) = z^{n+1} - z^{-n-1}$, is sometimes helpful.

21.10 Proposition. $\langle \chi | \chi_n \rangle$ *is the coefficient of* z^{n+1} *in* $\Delta(z)\chi(z)$. □

22 Some SU(2)-related groups

We now discuss two groups closely related to SU(2) and describe their irreducible representations and characters. These are SO(3), SO(4) and U(2). We start with the following

22.1 Proposition. $\mathrm{SO}(3) \cong \mathrm{SU}(2)/\{\pm\mathrm{Id}\}$, $\mathrm{SO}(4) \cong \mathrm{SU}(2) \times \mathrm{SU}(2)/\{\pm(\mathrm{Id},\mathrm{Id})\}$ *and* $\mathrm{U}(2) \cong \mathrm{U}(1) \times \mathrm{SU}(2)/\{\pm(\mathrm{Id},\mathrm{Id})\}$.

Proof. Recall the left×right multiplication action of SU(2), viewed as the space of unit norm quaternions, on $\mathbb{H} \cong \mathbb{R}^4$. This gives a homomorphism from $\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$. Note that $\alpha \times \beta$ send $1 \in \mathbb{H}$ to $\alpha\beta^{-1}$, so $\alpha \times \beta$ fixes 1 iff $\alpha = \beta$. The pair $\alpha \times \alpha$ fixes every other quaternion iff α is central in SU(2), that is, $\alpha = \pm\mathrm{Id}$. So the kernel of our homomorphism is $\{\pm(\mathrm{Id},\mathrm{Id})\}$.

We see surjectivity and construct the isomorphism $\mathrm{SU}(2) \to \mathrm{SO}(3)$ at the same time. Restricting our left×right action to the diagonal copy of SU(2) leads to the conjugation action of SU(2) on the space of *pure quaternions*, spanned by $\mathrm{i}, \mathrm{j}, \mathrm{k}$. I claim that this generates the full rotation group SO(3): indeed, rotations in the $\langle \mathrm{i}, \mathrm{j} \rangle$-plane are implemented by elements $a + b\mathrm{k}$, and similarly with any permutation of $\mathrm{i}, \mathrm{j}, \mathrm{k}$, and these rotations generate SO(3). This constructs a surjective homomorphism from SU(2) to SO(3), but we already know that the kernel is $\{\pm\mathrm{Id}\}$. So we have the assertion about SO(3).

Returning to SO(4), we can take any orthonormal, pure quaternion frame to any other one by a conjugation. We can also take $1 \in \mathbb{H}$ to any other unit vector by a left multiplication.

Combining these, it follows that we can take any orthonormal 4-frame to any other one by a suitable conjugation, followed by a left multiplication.

Finally, the assertion about $\mathrm{U}(2)$ is clear, as both $\mathrm{U}(1)$ and $\mathrm{SU}(2)$ sit naturally in $\mathrm{U}(2)$ (the former as the scalar matrices), and intersect at $\{\pm\mathrm{Id}\}$. $\square$

The group isomorphisms in the proposition are in fact homeomorphisms; that means, the inverse map is continuous (using the quotient topology on the quotient groups). It is not difficult (although a bit painful) to prove this directly from the construction, but this follows more easily from the following topological fact which is of independent interest.

22.2 Proposition. *Any continuous bijection from a Hausdorff space to a compact space is a homeomorphism.* $\square$

(22.3) Representations

It follows that continuous representations of the three groups in Proposition 22.1 are the same as continuous representations of $\mathrm{SU}(2), \mathrm{SU}(2) \times \mathrm{SU}(2)$ and $\mathrm{SU}(2) \times \mathrm{U}(1)$, respectively, which send $-\mathrm{Id}, -(\mathrm{Id}, \mathrm{Id})$ and $-(\mathrm{Id}, \mathrm{Id})$ to the identity matrix.

22.4 Corollary. *The complete list of irreducible representations of* $\mathrm{SO}(3)$ *is* $\mathrm{Sym}^{2n}\mathbb{C}^2$*, as* $n \geq 0$.

This formulation is slightly abusive, as $\mathbb{C}^2$ itself is not a representation of $\mathrm{SO}(3)$ but only of $\mathrm{SU}(2)$ ($-\mathrm{Id}$ acts as $-\mathrm{Id}$). But the sign problem is fixed on all even symmetric powers. For example, $\mathrm{Sym}^2\mathbb{C}^2$ is the standard 3-dimensional representation of $\mathrm{SO}(3)$. (There is little choice about it, as it is the only 3-dimensional representation on the list).

The image of our copy $\{\mathrm{diag}[z, z^{-1}]\} \subset \mathrm{SU}(2)$ of $\mathrm{U}(1)$, in our homomorphism, is the family of matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix},$$

with (caution!) $\varphi = 2\phi$. These cover all the conjugacy classes in $\mathrm{SO}(3)$, and the irreducible characters of $\mathrm{SO}(3)$ restrict to this as

$$1 + 2\cos\varphi + 2\cos(2\varphi) + \cdots + 2\cos(n\varphi).$$

(22.5) Representations of a product

Handling the other two groups, $\mathrm{SO}(4)$ and $\mathrm{U}(2)$, requires the following

22.6 Lemma. *For the product $G \times H$ of two compact groups G and H, the complete list of irreducible representations consists of the tensor products $V \otimes W$, as V and W range over the irreducibles of G and H, independently.*

Proof using character theory. From the properties of the tensor product of matrices, it follows that the character of $V \otimes W$ at the element $g \times h$ is

$$\chi_{V\otimes W}(g \times h) = \chi_V(g) \cdot \chi_W(h).$$

Now, a conjugacy class in $G \times H$ is a Cartesian product of conjugacy classes in G and H, and character theory ensures that the χ_V and χ_W form Hilbert space bases of the L^2 class functions on the two groups. It follows that the $\chi_V(g) \cdot \chi_W(h)$ form a Hilbert space basis of the class functions on $G \times H$, so this is a complete list of irreducible characters. $\square$

Incidentally, the proof establishes the *completeness* of characters for the product $G \times H$, if it was known on the factors. So, given our knowledge of $\mathrm{U}(1)$ and $\mathrm{SU}(2)$, it does give a complete argument. However, we also indicate another proof which does not rely on character theory, but uses complete reducibility instead. The advantage is that we only need to know complete reducibility under *one* of the factors G, H.

Proof without characters. Let U be an irreducible representation of $G \times H$, and decompose it into isotypical components U_i under the action of H. Because the action of G commutes with H, it must be block-diagonal in this decomposition, and irreducibility implies that we have a single block. Let then W be the unique irreducible H-type appearing in U, and let $V := \mathrm{Hom}^H(W,U)$. This inherits and action of G from U, and we have an isomorphism $U \cong W \otimes \mathrm{Hom}^H(W,U) = V \otimes W$ (see Example Sheet 2), with G and H acting on the two factors. Moreover, any proper G-subrepresentation of V would lead to a proper $G \times H$-subrepresentation of U, and this implies irreducibility of V. $\square$

22.7 Corollary. *The complete list of irreducible representations of* $\mathrm{SO}(4)$ *is* $\mathrm{Sym}^m\mathbb{C}^2 \otimes \mathrm{Sym}^n\mathbb{C}^2$, *with the two* $\mathrm{SU}(2)$ *factors acting on the two* Sym *factors. Here,* $m, n \geq 0$ *and* $m = n \mod 2$.

(22.8) A closer look at $\mathrm{U}(2)$

Listing the representations of $\mathrm{U}(2)$ requires a comment. As $\mathrm{U}(2) = \mathrm{U}(1) \times \mathrm{SU}(2)/\{\pm(\mathrm{Id},\mathrm{Id})\}$, in principle we should tensor together pairs of representations of $\mathrm{U}(1)$ and $\mathrm{SU}(2)$ of matching parity, so that $-(\mathrm{Id},\mathrm{Id})$ acts as $+\mathrm{Id}$ in the product. However, observe that $\mathrm{U}(2)$ acts naturally on $\mathbb{C}^2$, extending the $\mathrm{SU}(2)$ action, so the action of $\mathrm{SU}(2)$ on $\mathrm{Sym}^n\,\mathbb{C}^2$ also extends to all of $\mathrm{U}(2)$. Thereunder, the factor $\mathrm{U}(1)$ of scalar matrices does *not* act trivially, but rather by the nth power of the natural representation (since it acts naturally on $\mathbb{C}^2$).

In addition, $\mathrm{U}(2)$ has a 1-dimensional representation $\det : \mathrm{U}(2) \to \mathrm{U}(1)$. Denote by $\det^{\otimes m}$ its mth tensor power; it restricts to the trivial representation on $\mathrm{SU}(2)$ and to the $2m$th power of the natural representation on the scalar subgroup $\mathrm{U}(1)$. The classification of representations of $\mathrm{U}(2)$ in terms of their restrictions to the subgroups $\mathrm{U}(1)$ and $\mathrm{SU}(2)$ coverts then into the

22.9 Proposition. *The complete list of irreducible representations of* $\mathrm{U}(2)$ *is* $\det^{\otimes m} \otimes \mathrm{Sym}^n\,\mathbb{C}^2$, *with* $m, n \in \mathbb{Z}$, $n \geq 0$. $\square$

(22.10) Characters of $\mathrm{U}(2)$

We saw earlier that the conjugacy classes of $\mathrm{U}(2)$ were labelled by unordered pairs $\{z_1, z_2\}$ of eigenvalues. These are double-covered by the subgroup of diagonal matrices $\mathrm{diag}[z_1, z_2]$. The trace of this matrix on $\mathrm{Sym}^n\,\mathbb{C}^2$ is $z_1^n + z_1^{n-1}z_2 + \cdots + z_2^n$, and so the character of $\det^{\otimes m} \otimes \mathrm{Sym}^n\,\mathbb{C}^2$ is the symmetric[8] Laurent polynomial

$$z_1^{m+n}z_2^m + z_1^{m+n-1}z_2^{m+1} + \cdots + z_1^m z_2^{m+n}. \tag{22.11}$$

Moreover, these are all the irreducible characters of $\mathrm{U}(2)$. Clearly, they span the space of symmetric Laurent polynomials. General theory tells us that they should form a Hilbert space basis of the space of class functions, with respect to the $\mathrm{U}(2)$ inner product; but we can check this directly from the

22.12 Proposition (Weyl integration formula for $\mathrm{U}(2)$**).** *For a class function* f *on* $\mathrm{U}(2)$, *we have*

$$\int_{\mathrm{U}(2)} f(u)du = \frac{1}{2}\frac{1}{4\pi^2}\iint_{0\times 0}^{2\pi\times 2\pi} f(\phi_1,\phi_2)\cdot|e^{\mathrm{i}\phi_1} - e^{\mathrm{i}\phi_2}|^2 d\phi_1 d\phi_2.$$

Here, $f(\phi_1,\phi_2)$ abusively denotes $f(\mathrm{diag}[e^{\mathrm{i}\phi_1}, e^{\mathrm{i}\phi_2}])$. We call $\Delta(z_1,z_2) := z_1 - z_2 = e^{\mathrm{i}\phi_1} - e^{\mathrm{i}\phi_2}$ the *Weyl denominator* for $\mathrm{U}(2)$. The proposition is proved by lifting f to the double cover $\mathrm{U}(1) \times \mathrm{SU}(2)$ of $\mathrm{U}(2)$ and applying the integration formula for $\mathrm{SU}(2)$; we omit the details. Armed with the integration formula, we could have discovered the character formulae (22.11) *a priori*, before constructing the representations. Of course, proving that all formulae are actually realised as characters requires either a direct construction (as given here) or a general existence argument, as secured by the Peter-Weyl theorem.

[8] Under the switch $z_1 \leftrightarrow z_2$.

23 The unitary group*

The representation theory of the general unitary groups $\mathrm{U}(N)$ is very similar to that of $\mathrm{U}(2)$, with the obvious change that the characters are now *symmetric Laurent polynomials in N variables*, with integer coefficients; the variables represent the eigenvalues of a unitary matrix. These are polynomials in the z_i and z_i^{-1} which are invariant under permutation of the variables. Any such polynomial can be expressed as $(z_1 z_2 \cdots z_N)^n \cdot f(z_1, \ldots, z_N)$, where f is a genuine symmetric polynomial with integer coefficients. The fact that all $\mathrm{U}(N)$ characters are of this form follows from the classification of conjugacy classes in Lecture 20, and the representation theory of the subgroup $\mathrm{U}(1)^N$: representations of the latter are completely reducible, with the irreducible being tensor products of N one-dimensional representations of the factors (Lemma 22.6).

(23.1) Symmetry vs. anti-symmetry

Let us introduce some notation. For an N-tuple $\boldsymbol{\lambda} = \lambda_1, \ldots, \lambda_N$ of integers, denote by $\mathbf{z}^{\boldsymbol{\lambda}}$ the (Laurent) monomial $z_1^{\lambda_1} z_2^{\lambda_2} \cdots z_N^{\lambda_N}$. Given $\sigma \in S_N$, the symmetric group on N letters, $\mathbf{z}^{\sigma(\boldsymbol{\lambda})}$ will denote the monomial for the permuted N-tuple $\sigma(\boldsymbol{\lambda})$. There is a distinguished N-tuple $\boldsymbol{\delta} = (N-1, N-2, \ldots, 0)$. The *anti-symmetric* Laurent polynomials

$$a_{\boldsymbol{\lambda}}(\mathbf{z}) := \sum\nolimits_{\sigma \in S_N} \varepsilon(\sigma) \cdot \mathbf{z}^{\sigma(\boldsymbol{\lambda})},$$

where $\varepsilon(\sigma)$ denotes the signature of σ, are distinguished by the following simple observation:

23.2 Proposition. *As $\boldsymbol{\lambda}$ ranges over the decreasingly ordered N-tuples, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$, the $a_{\boldsymbol{\lambda}+\boldsymbol{\delta}}$ form a $\mathbb{Z}$-basis of the integral anti-symmetric Laurent polynomials.* □

The polynomial $a_{\boldsymbol{\delta}}(\mathbf{z})$ is special, as the following identities show:

$$a_{\boldsymbol{\delta}}(\mathbf{z}) = \det[z_p^{N-q}] = \prod_{p<q} (z_p - z_q).$$

The last product is also denoted $\Delta(\mathbf{z})$ and called the *Weyl denominator*. The first identity is simply the big formula for the determinant of the matrix with entries z_p^{N-q}; the second is called the *Vandermonde formula* and can be proved in several ways. One can do induction on N, combined with some clever row operations. Or, one notices that the determinant is a homogeneous polynomial in the z_p, of degree $N(N-1)/2$, and vanishes whenever $z_p = z_q$ for some $p \neq q$; this implies that the product on the right divides the determinant, so we have equality up to a constant factor, and we need only match the coefficient for a single term in the formula, for instance $\mathbf{z}^{\boldsymbol{\delta}}$.

23.3 Remark. There is a determinant formula (although not a product expansion) for every $a_{\boldsymbol{\lambda}}$:

$$a_{\boldsymbol{\lambda}}(\mathbf{z}) = \det[z_p^{\lambda_q}].$$

23.4 Proposition. *Multiplication by Δ establishes a bijection between symmetric and anti-symmetric Laurent polynomials with integer coefficients.*

Proof. The obvious part (which is the only one we need) is that $\Delta \cdot f$ is anti-symmetric and integral, if f is symmetric integral. In the other direction, anti-symmetry of a g implies its vanishing whenever $z_p = z_q$ for some $p \neq q$, which (by repeated use of Bézout's theorem, that $g(a) = 0 \Rightarrow (X - a)|g(X))$ implies that $\Delta(\mathbf{z})$ divides g, with integral quotient. □

23.5 Remark. The proposition and its proof apply equally well to the genuine polynomials, as opposed to Laurent polynomials.

(23.6) The irreducible characters of $\mathrm{U}(N)$

23.7 Definition. The *Schur function* $s_{\boldsymbol{\lambda}}$ is defined as

$$s_{\boldsymbol{\lambda}}(\mathbf{z}) := \frac{a_{\boldsymbol{\lambda}+\boldsymbol{\delta}}(\mathbf{z})}{a_{\boldsymbol{\delta}}(\mathbf{z})} = \frac{a_{\boldsymbol{\lambda}+\boldsymbol{\delta}}(\mathbf{z})}{\Delta(\mathbf{z})}.$$

By the previous proposition, the Schur functions form an *integral basis* of the space of symmetric Laurent polynomials with integer coefficients. The *Schur polynomials* are the genuine polynomials among the $s_{\boldsymbol{\lambda}}$'s; they correspond to the $\boldsymbol{\lambda}$'s where each $\lambda_k > 0$, and they form a basis of the space of integral symmetric polynomials. Our main theorem is now

23.8 Theorem. *The Schur functions are precisely the irreducible characters of* $\mathrm{U}(N)$*.*

Following the model of SU(2) in the earlier sections and the general orthogonality theory of characters, the theorem breaks up into two statements.

23.9 Proposition. *The Schur functions are orthonormal in the inner product defined by integration over* $\mathrm{U}(N)$*, with respect to the normalised invariant measure.*

23.10 Proposition. *Every* $s_{\boldsymbol{\lambda}}$ *is the character of some representation of* $\mathrm{U}(N)$*.*

We shall not prove Proposition 23.9 here; as in the case of SU(2), it reduces to the following integration formula, whose proof, however, is now more difficult, due to the absence of a concrete geometric model for $\mathrm{U}(N)$ with its invariant measure.

23.11 Theorem (Weyl Integration for $\mathrm{U}(N)$). *For a class function f on* $\mathrm{U}(N)$*, we have*

$$\int_{\mathrm{U}(N)} f(u)du = \frac{1}{N!}\frac{1}{(2\pi)^N}\int_0^{2\pi}\cdots\int_0^{2\pi} f(\phi_1,\ldots,\phi_N)\cdot|\Delta(\mathbf{z})|^2 d\phi_1\cdots d\phi_N.$$

We have used the standard convention $f(\phi_1,\ldots,\phi_N) = f(\mathrm{diag}[z_1,\ldots,z_N])$ and $z_p = e^{\mathrm{i}\phi_p}$.

(23.12) Construction of representations

Much like in the case of SU(2), the irreps of $\mathrm{U}(N)$ can be found in terms of the tensor powers of its standard representation on $\mathbb{C}^N$. Formulating this precisely requires a comment. A representation of $\mathrm{U}(N)$ will be called *polynomial* iff its character is a genuine (as opposed to Laurent) polynomial in the z_p. For example, the determinant representation $u \mapsto \det u$ is polynomial, as its character is $z_1\cdots z_N$, but its dual representation has character $(z_1\cdots z_N)^{-1}$ and is not polynomial.

23.13 Remark. A symmetric polynomial in the z_p is expressible in terms of the elementary symmetric polynomials. If the z_p are the eigenvalues of a matrix u, these are the coefficients of the characteristic polynomial of u. As such, they have polynomial expressions in terms of the entries of u. Thus, in any polynomial representation ρ, the trace of $\rho(u)$ is a polynomial in the entries of $u \in \mathrm{U}(N)$. It turns out in fact that all *entries* of the matrix $\rho(u)$ are polynomially expressible in terms of those of u, without involving the entries of u^{-1}.

Now, the character of $(\mathbb{C}^N)^{\otimes d}$ is the polynomial $(z_1 + \cdots + z_N)^d$, and it follows from our discussion of Schur functions and polynomials that any representation of $\mathrm{U}(N)$ appearing in the irreducible decomposition of $(\mathbb{C}^N)^{\otimes d}$ is polynomial. The key step in proving the existence of representation now consists in showing that *every Schur polynomial appears in the irreducible decomposition of* $(\mathbb{C}^N)^{\otimes d}$*, for some* d. This implies that all Schur polynomials are characters of representations. However, any Schur function converts into a Schur polynomial upon multiplication by a large power of det; so this does imply that all Schur *functions* are characters of $\mathrm{U}(N)$-representations (a polynomial representation tensored with a large inverse power of det).

Note that the Schur function $s_{\boldsymbol{\lambda}}$ is homogeneous of degree $|\boldsymbol{\lambda}| = \lambda_1 + \cdots + \lambda_N$; so if it appears in any $(\mathbb{C}^N)^{\otimes d}$, it must do so for $d = |\boldsymbol{\lambda}|$. Checking its presence requires us only to show that

$$\langle s_{\boldsymbol{\lambda}} | (z_1 + \cdots + z_N)^{|\boldsymbol{\lambda}|} \rangle > 0$$

in the inner product (23.11). While there is a reasonably simple argument for this, it turns out that we can (and will) do much more with little extra effort.

(23.14) Schur-Weyl duality

To state the key theorem, observe that the space $(\mathbb{C}^N)^{\otimes d}$ carries an action of the symmetric group S_d on d letters, which permutes the factors in every vector $\mathbf{v}_1 \otimes \mathbf{v}_2 \otimes \cdots \otimes \mathbf{v}_d$. Clearly, this commutes with the action of $\mathrm{U}(N)$, in which a matrix $u \in \mathrm{U}(N)$ transforms all tensor factors simultaneously. This way, $(\mathbb{C}^N)^{\otimes d}$ becomes a representation of the product group $\mathrm{U}(N) \times S_d$.

23.15 Theorem (Schur-Weyl duality). *Under the action of* $\mathrm{U}(N) \times S_d$, $(\mathbb{C}^N)^{\otimes d}$ *decomposes as a direct sum of products*

$$\bigoplus_{\boldsymbol{\lambda}} V_{\boldsymbol{\lambda}} \otimes S^{\boldsymbol{\lambda}}$$

labelled by partitions of d with no more than N parts, in which $V_{\boldsymbol{\lambda}}$ is the irreducible representation of $\mathrm{U}(N)$ *with character $s_{\boldsymbol{\lambda}}$ and $S^{\boldsymbol{\lambda}}$ is an irreducible representation of S_d.*

The representations $S^{\boldsymbol{\lambda}}$, for various $\boldsymbol{\lambda}$'s, depend only on $\boldsymbol{\lambda}$ (and not on N), are pairwise non-isomorphic. For $N \geq d$ they exhaust all irreducible representations of S_d.

Recall that a *partition of d with n parts* is a sequence $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n > 0$ of integers summing to d. Once $N \geq d$, all partitions of d are represented in the sum. As the number of conjugacy classes in S_d equals the number of partitions, the last sentence in the theorem is obvious. Nothing else in the theorem is obvious; it will be proved in the next section.

24 The symmetric group and Schur-Weyl duality*

(24.1) Conjugacy classed in the symmetric group

The duality theorem relates the representations of the unitary groups $\mathrm{U}(N)$ to those of the symmetric groups S_d, so we must say a word about the latter. Recall that every permutation has a unique decomposition as a product of disjoint cycles, and that two permutations are conjugate if and only if they have the same *cycle type*, which is the collection of all cycle lengths, with multiplicities. Ordering the lengths decreasingly leads to a partition $\boldsymbol{\mu} = \mu_1 \geq \ldots \geq \mu_n > 0$ of d. Assume that the numbers $1, 2, \ldots, d$ occur $m_1, m_2, \ldots, m_d$ times, respectively in $\boldsymbol{\mu}$; that is, m_k is the number of k-cycles in our permutation. We shall also refer to the cycle type by the notation $(\mathbf{m})$. Thus, the conjugacy classes in S_d are labelled by $(\mathbf{m})$'s such that $\sum_k k \cdot m_k = d$. It follows that the number of irreducible characters of S_d is the number of partitions of d. What is also true but unexpected is that there is a natural way to assign irreducible representations to partitions, as we shall explain.

24.2 Proposition. *The conjugacy class of cycle type* $(\mathbf{m})$ *has order*

$$\frac{d!}{\prod m_k! \cdot \prod k^{m_k}}.$$

Proof. By the orbit-stabiliser theorem, applied to the conjugation action of S_d on itself, it suffices to show that the centraliser of a permutation of cycle type $(\mathbf{m})$ is $\prod m_k! \cdot \prod k^{m_k}$. But an element of the centraliser must permute the k-cycles, for all k, while preserving the cyclic order of the elements within. We get a factor of $m_k!$ from the permutations of the m_k k-cycles, and k^{m_k} from independent cyclic permutations within each k-cycle. $\square$

(24.3) The irreducible characters of S_d

Choose $N > 0$ and denote by p_μ or $p_{(\mathbf{m})}$ the product of *power sums*

$$p_\mu(\mathbf{z}) = \prod_k (z_1^{\mu_k} + z_2^{\mu_k} + \cdots + z_N^{\mu_k}) = \prod_{k>0} (z_1^k + \cdots + z_N^k)^{m_k};$$

note that if we are to allow a number of trailing zeroes at the end of $\boldsymbol{\mu}$, they must be ignored in the first product. For any partition $\boldsymbol{\lambda}$ of d with no more than N parts, we define a function on conjugacy classes of S_d by

$$\omega_{\boldsymbol{\lambda}}(\mathbf{m}) = \text{ coefficient of } \mathbf{z}^{\boldsymbol{\lambda}+\boldsymbol{\delta}} \text{ in } p_{(\mathbf{m})}(\mathbf{z}) \cdot \Delta(\mathbf{z}).$$

It is important to observe that $\omega_{\boldsymbol{\lambda}}(\mathbf{m})$ does *not* depend on N, provided the latter is larger than the number of parts of $\boldsymbol{\lambda}$; indeed, an extra variable z_{N+1} can be set to zero, $p_{(\mathbf{m})}$ is then unchanged, while both $\Delta(\mathbf{z})$ and $\mathbf{z}^{\boldsymbol{\lambda}+\boldsymbol{\delta}}$ acquire a factor of $z_1 z_2 \cdots z_N$, leading to the same coefficient in the definition.

Anti-symmetry of $p_{(\mathbf{m})}(\mathbf{z})\cdot\Delta(\mathbf{z})$ and the definition of the Schur functions lead to the following

24.4 Proposition. $p_{(\mathbf{m})}(\mathbf{z}) = \sum_{\boldsymbol{\lambda}} \omega_{\boldsymbol{\lambda}}(\mathbf{m}) \cdot s_{\boldsymbol{\lambda}}(\mathbf{z})$*, the sum running over the partitions of d with no more than N parts.*

Indeed, this is equivalent to the identity

$$p_{(\mathbf{m})}(\mathbf{z}) \cdot \Delta(\mathbf{z}) = \sum_{\boldsymbol{\lambda}} \omega_{\boldsymbol{\lambda}}(\mathbf{m}) \cdot a_{\boldsymbol{\lambda}+\boldsymbol{\delta}}(\mathbf{z}),$$

which just restates the definition of the $\omega_{\boldsymbol{\lambda}}$.

24.5 Theorem (Frobenius character formula). *The $\omega_{\boldsymbol{\lambda}}$ are the irreducible characters of the symmetric group S_d.*

We will prove that the $\omega_{\boldsymbol{\lambda}}$ are orthonormal in the inner product of class functions on S_d. We will then show that they are characters of representations of S_d, which will imply their irreducibility. The second part will be proved in conjunction with Schur-Weyl duality.

Proof of orthonormality. We must show that for any partitions $\boldsymbol{\lambda}, \boldsymbol{\nu}$,

$$\sum_{(\mathbf{m})} \frac{\omega_{\boldsymbol{\lambda}}(\mathbf{m}) \cdot \omega_{\boldsymbol{\nu}}(\mathbf{m})}{\prod m_k! \cdot \prod k^{m_k}} = \delta_{\lambda\nu}. \tag{24.6}$$

We will prove these identities simultaneously for all d, via the identity

$$\sum_{(\mathbf{m});\boldsymbol{\lambda},\boldsymbol{\nu}} \frac{\omega_{\boldsymbol{\lambda}}(\mathbf{m}) \cdot \omega_{\boldsymbol{\nu}}(\mathbf{m})}{\prod m_k! \cdot \prod k^{m_k}} s_{\boldsymbol{\lambda}}(\mathbf{z}) s_{\boldsymbol{\nu}}(\mathbf{w}) \equiv \sum_{\boldsymbol{\lambda}} s_{\boldsymbol{\lambda}}(\mathbf{z}) s_{\boldsymbol{\lambda}}(\mathbf{w}) \tag{24.7}$$

for variables $\mathbf{z}, \mathbf{w}$, with $(\mathbf{m})$ ranging over all cycle types of all symmetric groups S_d and $\boldsymbol{\lambda}, \boldsymbol{\nu}$ ranging over all partitions with no more than N parts of all integers d. Independence of the Schur polynomials implies (24.6).

From Proposition 24.4, the left side in (24.7) is

$$\sum_{(\mathbf{m})} \frac{p_{(\mathbf{m})}(\mathbf{z}) \cdot p_{(\mathbf{m})}(\mathbf{w})}{\prod m_k! \cdot \prod k^{m_k}},$$

which is also

$$\sum_{(\mathbf{m})}\prod_{k>0}\frac{(z_1^k+\cdots+z_N^k)^{m_k}(w_1^k+\cdots+w_N^k)^{m_k}}{m_k!\cdot k^{m_k}}$$

$$=\sum_{(\mathbf{m})}\prod_{k>0}\frac{(\sum_{p,q}(z_pw_q)^k/k)^{m_k}}{m_k!}=\prod_{k>0}\exp\left\{\sum\nolimits_{p,q}(z_pw_q)^k/k\right\}$$

$$=\exp\left\{\sum\nolimits_{p,q;k}(z_pw_q)^k/k\right\}=\exp\left\{-\sum\nolimits_{p,q}\log(1-z_pw_q)\right\}$$

$$=\prod_{p,q}(1-z_pw_q)^{-1}.$$

We are thus reduced to proving the identity

$$\sum_{\boldsymbol{\lambda}}s_{\boldsymbol{\lambda}}(\mathbf{z})s_{\boldsymbol{\lambda}}(\mathbf{w})=\prod_{p,q}(1-z_pw_q)^{-1}. \tag{24.8}$$

There are (at least) two ways to proceed. We can multiply both sides by $\Delta(\mathbf{z})\Delta(\mathbf{w})$ and show that each side agrees with the $N\times N$ *Cauchy determinant* $\det[(1-z_pw_q)^{-1}]$. The equality

$$\det[(1-z_pw_q)^{-1}]=\Delta(\mathbf{z})\Delta(\mathbf{w})\prod_{p,q}(1-z_pw_q)^{-1}$$

can be proved by clever row operations and induction on N. (See Appendix A of Fulton-Harris, *Representation Theory*, for help if needed.) On the other hand, if we expand each entry in the matrix in a geometric series, multi-linearity of the determinant gives

$$\begin{aligned}\det[(1-z_pw_q)^{-1}]&=\sum_{l_1,\dots,l_N\geq 0}\det[(z_pw_q)^{l_q}]\\&=\sum_{l_1,\dots,l_N\geq 0}\det[z_p^{l_q}]\cdot\prod_q w_q^{l_q}\\&=\sum_{l_1,\dots,l_N\geq 0}a_{\mathbf{l}}(\mathbf{z})\mathbf{w}^{\mathbf{l}}=\sum_{l_1>l_2>\dots>l_N\geq 0}a_{\mathbf{l}}(\mathbf{z})a_{\mathbf{l}}(\mathbf{w}),\end{aligned}$$

the last identity from the anti-symmetry in $\mathbf{l}$ of the $a_{\mathbf{l}}$. The result is the left side of (24.8) multiplied by $\Delta(\mathbf{z})\Delta(\mathbf{w})$, as desired.

Another proof of (24.8) can be given by exploiting the orthogonality properties of Schur functions and a judicious use of Cauchy's integration formula. Indeed, the left-hand side $\Sigma(\mathbf{z},\mathbf{w})$ of the equation has the property that

$$\frac{1}{N!}\oint\cdots\oint\Sigma(\mathbf{z},\mathbf{w}^{-1})f(\mathbf{w})\Delta(\mathbf{w})\Delta(\mathbf{w}^{-1})\frac{dw_1}{2\pi\mathrm{i}w_1}\cdots\frac{dw_N}{2\pi\mathrm{i}w_N}=f(\mathbf{z})$$

for any symmetric polynomial $f(\mathbf{w})$, while any homogeneous symmetric Laurent polynomial containing negative powers of the w_q integrates to zero. (The integrals are over the unit circles.) An N-fold application of Cauchy's formula shows that the same is true for $\prod_{p,q}(1-z_pw_q^{-1})^{-1}$, the function obtained from the right-hand side of (24.8) after substituting $w_q\leftrightarrow w_q^{-1}$, subject to the convergence condition $|z_p|<|w_q|$ for all p,q. I will not give the detailed calculation, as the care needed in discussing convergence makes this argument more difficult, in the end, than the one with the Cauchy determinant, but just observe that the algebraic manipulation

$$\frac{\Delta(\mathbf{w})\Delta(\mathbf{w}^{-1})}{w_1\cdots w_N\prod_{p,q}(1-z_pw_q^{-1})}=\frac{\prod_{p\neq q}(w_q-w_p)}{\prod_{p,q}(w_q-z_p)}$$

and the desired integral is rewritten as

$$\frac{1}{N!}\oint\cdots\oint\frac{\prod_{p\neq q}(w_q-w_p)}{\prod_{p,q}(w_q-z_p)}f(\mathbf{w})\frac{dw_1}{2\pi\mathrm{i}}\cdots\frac{dw_N}{2\pi\mathrm{i}}=f(\mathbf{z}),$$

for any symmetric polynomial f. For $N=1$ one recognises the Cauchy formula; in general, we have simple poles at the z_p, to which values the w_q are to be set, in all possible permutations. (The numerator ensures that setting different w's to the same z value gives no contribution). Note that, if f is not symmetric, the output is the symmetrisation of f. □

(24.9) Proof of the key theorems

We now prove the following, with a single calculation:

1. The $\omega_{\boldsymbol{\lambda}}$ are characters of representations of S_d.
2. Every Schur polynomial is the character of some representation of $\mathrm{U}(N)$.
3. The Schur-Weyl duality theorem.

Note that (1) and (2) automatically entail the irreducibility of the corresponding representations. Specifically, we will verify the following

24.10 Proposition. *The character of* $(\mathbb{C}^N)^{\otimes d}$*, as a representation of* $S_d\times\mathrm{U}(N)$*, is*

$$\sum\nolimits_{\boldsymbol{\lambda}}\omega_{\boldsymbol{\lambda}}(\mathbf{m})\cdot s_{\boldsymbol{\lambda}}(\mathbf{z}),$$

with the sum ranging over the partitions of d *with no more than* N *parts.*

The character is the trace of the action of $M:=\sigma\times\mathrm{diag}[z_1,\ldots,z_N]$, for any permutation σ of cycle type $(\mathbf{m})$. By Lemma 22.6, $(\mathbb{C}^N)^{\otimes d}$ breaks up as a sum $\bigoplus W_i\otimes V_i$ of tensor products of irreducible representations of S_d and $\mathrm{U}(N)$, respectively. Collecting together the irreducible representations of $\mathrm{U}(N)$ expresses the character of $(\mathbb{C}^N)^{\otimes d}$ as

$$\sum\nolimits_{\boldsymbol{\lambda}}\chi_{\boldsymbol{\lambda}}(\mathbf{m})\cdot s_{\boldsymbol{\lambda}}(\mathbf{z}),$$

for certain characters $\chi_{\boldsymbol{\lambda}}$ of S_d. Comparing this with the proposition implies (1)–(3) in our list.

Proof. Let $\mathbf{e}_1,\ldots,\mathbf{e}_N$ be the standard basis of $\mathbb{C}^N$. A basis for $(\mathbb{C}^N)^{\otimes d}$ is $\mathbf{e}_{i_1}\otimes\mathbf{e}_{i_2}\otimes\cdots\otimes\mathbf{e}_{i_d}$, for all possible indexing functions $\mathbf{i}:\{1,\ldots,d\}\to\{1,\ldots,N\}$. Under M, this basis vector gets transformed to $z_{i_1}\cdots z_{i_d}\cdot\mathbf{e}_{i_{\sigma(1)}}\otimes\mathbf{e}_{i_{\sigma(2)}}\otimes\cdots\otimes\mathbf{e}_{i_{\sigma(d)}}$. This is off-diagonal unless $i_k=i_{\sigma(k)}$ for each k, that is, unless $\mathbf{i}$ is constant within each cycle of σ. So the only contribution to the trace comes from such $\mathbf{i}$.

The diagonal entry for such a $\mathbf{i}$ is a product of factors of z_i^k for each k-cycle, with an appropriate i. Summing this over all these $\mathbf{i}$ leads to

$$\prod_{k>0}(z_1^k+\cdots+z_N^k)^{m_k}=p_{(\mathbf{m})}(\mathbf{z}),$$

and our proposition now follows from Prop. 24.4. □