4 The Main Theorem

4.1 More on Maschke's Theorem

Last class, given a finite-dimensional complex representation $\rho : G \to \mathrm{GL}(V)$ of a finite group G, we found a G-invariant positive Hermitian form on V and used it to show that a G-invariant subspace W has an invariant complement, namely $W^\perp$ (its orthogonal complement with respect to the Hermitian form). We used this to deduce Maschke's Theorem — that every representation can be split as a direct sum of irreducible ones.

We'll now discuss a few features of this proof.

First, why did we use Hermitian forms specifically? A different choice of form which may also seem reasonable is the *symmetric bilinear form*, a form where we require $\langle w, v \rangle$ to equal $\langle v, w \rangle$ rather than its conjugate. (For example, $v \cdot w$ is the standard symmetric bilinear form, while $v \cdot \overline{w}$ is the standard Hermitian form.)

The reason is that in a symmetric bilinear form over $\mathbb{C}$, it's possible that $v \cdot v = 0$. For example, consider the representation of $\mathbb{Z}/3\mathbb{Z}$ acting on $\mathbb{C}^3$, where

$$\overline{1} \mapsto A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

We can see that $(1, \zeta, \zeta^2)^t$ is an eigenvector, since

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \zeta \\ \zeta^2 \end{bmatrix} = \begin{bmatrix} \zeta^2 \\ 1 \\ \zeta \end{bmatrix} = \zeta^2 \begin{bmatrix} 1 \\ \zeta \\ \zeta^2 \end{bmatrix}.$$

But if we tried to perform our construction, taking W to be the span of this eigenvector, we'd see that $W^\perp$ actually *contains* it, since $(1, \zeta, \zeta^2) \cdot (1, \zeta, \zeta^2) = 1 + \zeta^2 + \zeta^4 = 0$. This means $W^\perp$ isn't actually a complement of W, so this would break the construction. We require that our form is *Hermitian* (and positive) to avoid this issue, since in that case $W^\perp$ really is a complement of W.

Another useful takeaway from our proof was that we found an invariant Hermitian form by *averaging*. This trick of averaging over all $g \in G$ can produce many other invariant things.

> **Example 4.1**
> Given a representation $\psi : G \to \mathrm{GL}(V)$ and a vector $v \in V$, the vector
>
> $$\frac{1}{|G|} \sum_{g \in G} \psi_g v$$
>
> is G-invariant. This is because for any $h \in H$, we have
>
> $$\psi_h \frac{1}{|G|} \sum_{g \in G} \psi_g v = \frac{1}{|G|} \sum_{g \in G} \psi_{hg} v = \frac{1}{|G|} \sum_{g \in G} \psi_g v,$$
>
> since $g \mapsto hg$ is a bijection on G (as g runs over all of G, so does hg).

> **Note 4.2**
> We saw a similar trick in 18.701, when proving that every finite group of isometries of $\mathbb{R}^2$ (or more generally $\mathbb{R}^n$) has a fixed point — we can start with any point p, and consider the points gp in its orbit. Then the center of mass (or average) of all these points is a fixed point, for the same reason we saw here.
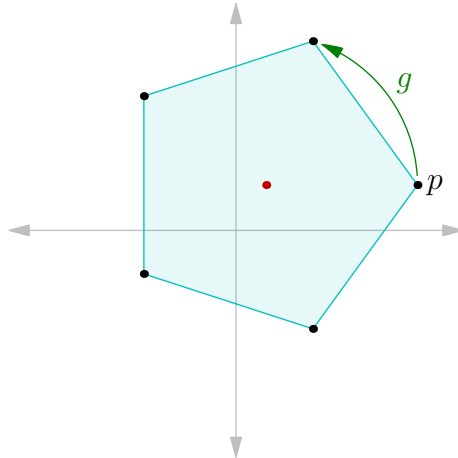
There's something to be careful of here, though — in Example 4.1, we don't actually know that this vector is nonzero — in fact, for many representations ψ it *must* be zero (often there's no nontrivial invariant vectors).

In fact, we can describe our construction *directly* in terms of the averaging trick as described in Example 4.1. We can think of the space of Hermitian forms as a *real* vector space; then G has a real representation acting on this space, sending $\langle v, w \rangle$ to $\langle gv, gw \rangle$. In our construction, we started with some positive form, and averaged it over all g to get an invariant "vector" (meaning an invariant Hermitian form). Here we don't have the issue of the invariant vector possibly being zero, because when we add two positive Hermitian forms, our resulting Hermitian form is again positive.

Student Question. *How do we describe the space of Hermitian forms as a real vector space?*

Answer. *Every Hermitian form can be described as $\langle v, w \rangle = vA\overline{w}$ for some matrix A with $A^t = \overline{A}$. We can think of all entries of this matrix in terms of their real and complex parts. Then the n entries on the diagonal must all be real, we get to choose both the real and complex parts of the $n(n-1)/2$ entries below the diagonal, and this immediately determines the $n(n-1)/2$ entries above the diagonal (which must be the conjugate of their reflection). So we get to choose*

$$n + 2 \cdot \frac{n(n-1)}{2} = n^2$$

real numbers, which means the space of Hermitian forms has dimension n^2.

On a different note, the fact that every representation has an invariant positive Hermitian form (as shown in our proof) is equivalent to stating that every representation of a finite group is conjugate to a *unitary representation*:

> **Definition 4.3**
> A unitary representation is a homomorphism $\rho : G \to \mathrm{U}_n$, where $\mathrm{U}_n \subset \mathrm{GL}_n$ is the set of unitary matrices.[a]
> _____
> [a]Matrices A for which $A^t = \overline{A}$

Equivalently, we can define unitary representations without referring to matrices — a linear operator is unitary if it preserves the standard Hermitian form $\langle v, w \rangle = v \cdot \overline{w}$, so a representation is unitary if and only if $gv \cdot g\overline{w} = v \cdot \overline{w}$ for all $g \in G$ and $v, w \in \mathbb{C}^n$.

To see why these two ideas are equivalent, note that given a positive Hermitian form, we can choose an orthonormal basis with respect to that form. In that basis, the form will just be the standard Hermitian form $\langle v, w \rangle = v \cdot \overline{w}$. So by changing the basis, we have produced a unitary representation.

Finally, we'll describe the proof of Maschke's Theorem (also known as *complete reducibility*) more explicitly.

> **Theorem 4.4** (Maschke's Theorem)
> Every complex representation of a finite group is isomorphic to a direct sum of irreducible representations.

Proof. We use induction on the dimension. For the base case, any one-dimensional representation is already irreducible.

Now suppose $\rho : G \to \mathrm{GL}(V)$ is a representation. If ρ is already irreducible, we're done. Otherwise, we can pick an invariant subspace W, which is neither 0 nor V. Then let $\langle -, - \rangle$ be an invariant positive Hermitian form, and decompose $V = W \oplus W^{\perp}$. Since both W and $W^{\perp}$ are G-invariant, then we get the subrepresentations $\psi : G \to \mathrm{GL}(W)$ and $\eta : G \to \mathrm{GL}(W^{\perp})$ in ρ, and we can decompose $\rho \cong \psi \oplus \eta$. But η and ψ have smaller dimension than ρ, so both are a direct sum of irreducible representations (by the inductive hypothesis), and therefore ρ is a direct sum of irreducible representations as well. $\qquad \square$

4.2 More on Characters

Let's continue our discussion from earlier about characters. We'll first state a few basic properties. (We will assume that all our representations are of *finite* groups, unless stated otherwise.)

Proposition 4.5

If $\rho : G \to \mathrm{GL}(V)$ is a complex representation, then

(a) $\chi_{\rho}(g)$ is a sum of roots of unity;

(b) $\chi_{\rho}(g^{-1}) = \overline{\chi_{\rho}(g)}$;

(c) $\overline{\chi_{\rho}}$ is the character of another representation of the same dimension, denoted ρ^{*} and called the *dual representation*.

Proof. These properties all come from the definition of the character as the trace of a matrix.

For (a), since G is a finite group, each $g \in G$ and therefore $\rho_g \in \mathrm{GL}(V)$ has finite order, so the eigenvalues of ρ_g, which we denote by $\lambda_i(g)$, are all roots of unity. Then $\mathrm{Tr}(\rho_g) = \sum \lambda_i(g)$ is a sum of roots of unity (as the trace is always the sum of eigenvalues with multiplicity).

For (b), the eigenvalues of $\rho_{g^{-1}}$ are the inverses of the eigenvalues of ρ_g (since the two matrices are inverses), and since the eigenvalues of ρ_g are all roots of unity, their inverses are equal to their conjugates. So we have

$$\mathrm{Tr}(\rho_{g^{-1}}) = \sum (\lambda_i(g))^{-1} = \sum \overline{\lambda_i(g)} = \overline{\mathrm{Tr}(\rho_g)}.$$

Finally, for (c), let V^{*} be the *dual space* of V, consisting of linear maps $f : V \to \mathbb{C}$. For convenience, we'll denote $f(v)$ by $\langle f, v \rangle$ (to emphasize the fact that we can think of it as a pairing between a vector v and a covector f). Then the dual representation ρ^{*} is given by

$$\langle \rho_g^{*}(f), v \rangle = \langle f, \rho_{g^{-1}}(v) \rangle.$$

(This defines $\rho_g^{*}(f)$ for each $f \in V$, by describing where it takes each vector.)

In other words, we make G act on V^{*} such that for every $v \in V$ and $f \in V^{*}$, we have

$$\langle f, v \rangle = \langle \rho_{g^*}(f), \rho_g(v) \rangle.$$

That is, the dual representation is defined such that operating on both the vector v (with the representation ρ) and the covector f (with the dual representation ρ^{*}) does not affect the pairing $\langle f, v \rangle$ given by the dual space. (Our original definition then follows from plugging in $\rho_{g^{-1}}(v)$ in place of v, to make the definition more explicit.)

This may seem somewhat abstract, but we can make it more concrete by describing it in terms of matrices. Fix a basis of V, so then from ρ we get a matrix representation $R : G \to \mathrm{GL}_n(\mathbb{C})$. Then the dual representation in terms of matrices is given by

$$R_g^{*} = R_{g^{-1}}^{t}.$$

This is a valid representation since $(AB)^t = B^t A^t$ and $(AB)^{-1} = B^{-1} A^{-1}$, so $((AB)^t)^{-1} = (B^t A^t)^{-1} = (A^t)^{-1}(B^t)^{-1}$ (intuitively, each of taking the inverse and transposing means we need to swap the two matrices, so doing both means we need to swap twice and get back our original order). Since $\mathrm{Tr}(A^t) = \mathrm{Tr}(A)$, we have

$$\chi_{R^*}(g) = \chi_R(g^{-1}) = \overline{\chi_R(g)}. \qquad \square$$

Student Question. *Why is the transpose important — if there was no transpose, would we still get a representation?*

Answer. *We'd get what's called an* anti-representation *instead — we'd have a map with the property that $\rho_{gh} = \rho_h \rho_g$. (This is a representation if G is abelian.) It's possible to get an anti-representation in two ways — by inverting the elements, or by taking their transposes — and doing* both *gives us back a valid representation.*

Student Question. *Why are the two definitions of the dual representation (the abstract one and the one given in terms of matrices) equivalent, and how do we get the formula for the character from the abstract definition?*

Answer. *One way to think about this is to first think in terms of matrices — in that setting, it's clear that $\chi_{R^*} = \overline{\chi_R}$. Now in the abstract setting, we can pick a basis for V. This gives a basis for V^* as well — given a basis $\{v_1, \ldots, v_n\}$ for V, we take f_i to be the function which is 1 on v_i and 0 on each of the other basis vectors. Then our abstract definition is equivalent to taking the inverse transpose of the corresponding matrices.*

Student Question. *If $\rho : G \to \mathrm{GL}(V)$, does there exist a representation on the* same *vector space V with character $\overline{\chi_\rho}$?*

Answer. *Technically, yes. The dual space V^* is isomorphic to V — they have the same dimension, so fixing a basis for each gives an isomorphism between them. But they're not isomorphic in a canonical way.*

4.3 The Main Theorem

To understand the main theorem, we need to understand the space of *class functions*.

> **Definition 4.6**
> A **class function** is a function $f : G \to \mathbb{C}$ which is fixed on each conjugacy class of G. That is, for a class function f, if g and h are conjugate, then $f(g) = f(h)$.

The space of class functions is a vector space over $\mathbb{C}$, with addition and scalar multiplication defined as usual for functions.

We'll now state the main theorem in our story about representations, which gives surprisingly detailed information about the characters of irreducible representations.

> **Theorem 4.7** (Main Theorem)
> Let G be a finite group. Then:
>
> (a) The characters of irreducible representations form a basis in the space of class functions on G.
>
> (b) This basis is *orthonormal* with respect to the Hermitian form on the space of class functions given by
> $$\langle f_1, f_2 \rangle = \frac{1}{|G|} \sum_{g \in G} f_1(g) \overline{f_2(g)}.$$
>
> (c) If $d_1, \ldots, d_m$ are the dimensions of the irreducible representations of G, then
> $$d_1^2 + d_2^2 + \cdots + d_n^2 = |G|,$$
> and each d_i divides $|G|$.

In particular, (a) also implies that the characters of different irreducible representations are *distinct*, since they must form a basis.

For (c), note that although this isn't written in terms of characters, we can interpret it as a statement about characters as well, since $\dim(\rho) = \chi_\rho(1)$.

We'll prove these properties in later classes; first we'll look at a few important implications.

> **Corollary 4.8**
> The character of a representation uniquely determines the representation, up to isomorphism.

Proof. By Maschke's Theorem, we know that any representation ρ can be decomposed as a sum of irreducibles — if we use $\rho_1, \ldots, \rho_n$ to denote the irreducible representations, then by grouping together isomorphic summands,

we can write

$$\rho = \bigoplus_i \rho_i^{n_i}$$

for some integers n_i (the notation ψ^k, or $\psi^{\oplus k}$, denotes a direct sum of k copies of the representation ψ). But then we have

$$\chi_\rho = \sum n_i \chi_{\rho_i}.$$

So the coefficients n_i when decomposing ρ as a sum of irreducibles are the same as the coefficients when decomposing χ_ρ as a sum of χ_{ρ_i}. But since the χ_{ρ_i} form a *basis* of the space of class functions, there's a *unique* way to write χ_ρ (which is a class function) as a linear combination of these χ_{ρ_i}! So the n_i are uniquely determined from the character of ρ, and therefore so is ρ itself. $\qquad \square$

> **Corollary 4.9**
> The number of irreducible representations of G is the number of conjugacy classes on G.

Proof. The number of irreducible representations of G is the dimension of the space of class functions (since their characters form a basis for this space). But this dimension is just the number of conjugacy classes, since to specify a class function $f : G \to \mathbb{C}$, we need to specify its value on each conjugacy class. $\qquad \square$

This theorem gives a lot of concrete information that we can figure out about irreducible representations by just looking at the group; we'll see some examples of this next class.

Resource: Algebra II Student Notes
Spring 2022
Instructor: Roman Bezrukavnikov
Notes taken by Sanjana Das and Jakin Ng