6 Orthonormality of Characters

6.1 Review: Schur's Lemma

Last time, we presented Schur's Lemma.

Recall that $\mathrm{Hom}_G(\rho, \psi)$, which may also be written as $\mathrm{Hom}_G(V, W)$, denotes the space of homomorphisms (or in other words, linear maps) from V to W which are G-equivariant, meaning that

$$\mathrm{Hom}_G(\rho, \psi) = \{f : V \to W \mid f \text{ linear, and } f(\rho_g(v)) = \psi_g(f(v)) \text{ for all } g \in G, v \in V\}.$$

Theorem 6.1 (Schur's Lemma)
Suppose $\rho : G \to \mathrm{GL}(V)$ and $\psi : G \to \mathrm{GL}(W)$ are irreducible (and complex and finite-dimensional). Then

$$\dim(\mathrm{Hom}_G(\rho, \psi)) = \begin{cases} 0 & \text{if } \rho \not\cong \psi \\ 1 & \text{if } \rho \cong \psi. \end{cases}$$

In other words, the first statement means that if $\rho \not\cong \psi$, then the only G-equivariant homomorphism $V \to W$ is the zero map; the second statement means that if $\rho \cong \psi$, then the only G-equivariant homomorphisms $V \to W$ are the scalar maps — or in other words, $\mathrm{End}_G(\rho) = \mathbb{C} \cdot \mathrm{Id}$.

The first statement is true for representations over *any* field of coefficients (and the same proof we saw last time works in the general case). However, the second statement is *not* true for arbitrary fields of coefficients — in the proof, we used the fact that eigenvectors must always exist, which isn't true in general (for example, $\mathbb{R}$ is not algebraically closed, so it's possible that the characteristic polynomial does not have roots, and there are no real eigenvalues). In particular, there are examples of real representations for which $\mathrm{End}_G(\rho) \neq \mathbb{R}$:

Example 6.2
Consider the representation of $\mathbb{Z}/3\mathbb{Z}$ acting on $\mathbb{R}^2$, where $\overline{1}$ is mapped to the matrix

$$\begin{bmatrix} \cos 2\pi/3 & -\sin 2\pi/3 \\ \sin 2\pi/3 & \cos 2\pi/3 \end{bmatrix}$$

which corresponds to rotation by $2\pi/3$. In this case, $\mathrm{End}_{\mathbb{Z}/3\mathbb{Z}}(\mathbb{R}^2)$ is $\mathbb{C}$, not $\mathbb{R}$.

Proof. As usual, multiplication by any scalar is in $\mathrm{End}_{\mathbb{Z}_3}(\mathbb{R}^2)$; these elements correspond to $\mathbb{R}$ (in the case of $\mathbb{C}$, Schur's Lemma states that the *only* elements of $\mathrm{End}_G(\rho)$ are multiplication by scalars).

But there are actually other possible endomorphisms — rotation by $\pi/2$ about the origin is *also* a G-equivariant endomorphism, since any two rotations about the origin must commute. We can think of this element as i; then taking linear combinations, all elements $a + bi$ (for real a and b) are in $\mathrm{End}_{\mathbb{Z}/3\mathbb{Z}}(\mathbb{R}^2)$ (and it's possible to check that there's no others).

Composing two such endomorphisms corresponds to multiplying their corresponding complex numbers (since the endomorphism corresponding to z can be thought of as multiplication by z in the complex plane). So then this means $\mathrm{End}_{\mathbb{Z}/3\mathbb{Z}}(\mathbb{R}^2) = \mathbb{C}$. □

We won't discuss this, but there also exists a real irreducible representation ρ of some group G for which $\mathrm{End}_G(\rho) = \mathbb{H}$ (where $\mathbb{H}$ denotes the quaternions).

6.2 An Implication of Schur's Lemma

We've seen earlier that every representation $\rho : G \to \mathrm{GL}(V)$ can be written as the sum of irreducible representations, with certain coefficients. Using Schur's Lemma, we can describe what these coefficients are:

Corollary 6.3

Let $\rho : G \to \mathrm{GL}(V)$ be a representation. Let $\rho_1, \ldots, \rho_n$ be the list of all irreducible representations of G (up to isomorphism). Then

$$\rho \cong \bigoplus_{i=1}^{n} \rho_i^{d_i}$$

where $d_k = \dim \mathrm{Hom}_G(\rho_k, \rho)$ for all k.

Proof. From Maschke's Theorem, we know that we can write

$$\rho \cong \bigoplus_{i=1}^{n} \rho_i^{d_i}$$

for *some* coefficients d_i, so then

$$\mathrm{Hom}_G(\rho_k, \rho) = \mathrm{Hom}_G\left(\rho_k, \bigoplus \rho_i^{d_i}\right) = \bigoplus_{i=1}^{n} \mathrm{Hom}_G(\rho_k, \rho_i)^{d_i} = \mathbb{C}^{d_k},$$

using the fact that by Schur's Lemma, $\mathrm{Hom}_G(\rho_k, \rho_i)$ is 0 if $i \neq k$ and $\mathbb{C}$ if $i = k$. This means

$$\dim \mathrm{Hom}_G(\rho_k, \rho) = d_k$$

for each k, as desired. ☐

Student Question. *Why could we write*

$$\mathrm{Hom}_G\left(\rho_k, \bigoplus \rho_i^{d_i}\right) = \bigoplus_{i=1}^{n} \mathrm{Hom}_G(\rho_k, \rho_i)^{d_i}?$$

Answer. *It's enough to see that* $\mathrm{Hom}_G(U, V \oplus W) = \mathrm{Hom}_G(U, V) \oplus \mathrm{Hom}_G(U, W)$.

First, by looking at matrices, it's possible to see that $\mathrm{Hom}_{\mathbb{C}}(U, V \oplus W) = \mathrm{Hom}_{\mathbb{C}}(U, V) \oplus \mathrm{Hom}_{\mathbb{C}}(U, W)$ *(where this denotes* all *linear maps, not just the G-equivariant ones) — if* $\dim U = m$, $\dim V = n_1$, *and* $\dim W = n_2$, *then a linear map $U \to V \oplus W$ is a $(n_1 + n_2) \times m$ matrix, which we can think of as a pair of a $n_1 \times m$ matrix and a $n_2 \times m$ matrix:*

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}.$$

Then such a map is compatible with the G-action if and only if each component is.

It's possible to think of this without matrices, as well. By definition, $V \oplus W$ is the space of pairs (v, w) with $v \in V$ and $w \in W$. So in order to describe a linear map from U to $V \oplus W$, for an element $u \in U$, we need to specify the first *coordinate of its image (corresponding to a linear map $U \to V$) and the* second *coordinate of its image (corresponding to a linear map $U \to W$). Then since G essentially acts separately on V and W (by the definition of a direct sum of representations), the map $U \to V \oplus W$ is G-equivariant if and only if the two individual maps $U \to V$ and $U \to W$ are.*

Student Question. *What exactly does it mean to have a list of irreducible representations up to isomorphism — what happens if there are two isomorphic representations, but they act on different vector spaces?*

Answer. *We can think of $\rho_1, \ldots, \rho_n$ as an abstract list of representations, without thinking about the subspaces being acted on. For example, when writing down the character table of S_3, we saw that there are three irreducible representations; this means every irreducible representation is isomorphic to one of them. We're using "up to isomorphism" in the same sense here.*

More generally, when we write $\rho = \bigoplus_{i=1}^{n} \rho_i^{d_i}$, when $d_k > 0$, this really means that ρ_k is isomorphic to a sub-representation of ρ.

6.3 Matrices and a New Representation

We'll now rewrite the concept of G-equivariance in terms of matrices — this will give a useful construction of a new representation, which we can apply Schur's Lemma to in order to prove the orthonormality of irreducible characters.

Choose a basis for V and W. Then if $n = \dim V$ and $m = \dim W$, we can write a linear map $V \to W$ as a $m \times n$ matrix, where the map sends $v \in V$ to $Av \in W$.

We can also write our representations ρ and ψ as the matrix representations $R : G \to \mathrm{GL}_n(\mathbb{C})$ and $S : G \to \mathrm{GL}_m(\mathbb{C})$. Then by rewriting the definition of G-equivariance (that $f(\rho_g(v)) = \psi_g(f(v))$ for all v and g) in terms of these matrices, we have that a matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{C})$ corresponds to a linear map $f \in \mathrm{Hom}_G(\rho, \psi)$ if and only if

$$AR_g = S_g A \text{ for all } g \in G.$$

(Our initial definition was written in terms of v, but we can think of it instead as an equality of the *linear maps* themselves — that the linear maps $f \circ \rho_g$ and $\psi_g \circ f$ are the same — which corresponds to an equality of matrices.) We can rewrite this condition as

$$A = S_g A R_g^{-1} \text{ for all } g \in G.$$

The key point is that we can get *another* representation of G from this expression (starting with our original representations ρ and ψ). Note that the space $M = \mathrm{Mat}_{m\times n}(\mathbb{C})$ is *itself* a $\mathbb{C}$-vector space — we can forget everything we know about how to multiply matrices, and just imagine adding them and multiplying by scalars, which makes $\mathrm{Mat}_{m\times n}(\mathbb{C})$ a mn-dimensional vector space over $\mathbb{C}$.

Lemma 6.4
There is a representation C of G acting on $\mathrm{Mat}_{m\times n}(\mathbb{C})$, where for each $g \in G$, g is sent to the matrix

$$C_g : A \mapsto S_g A R_g^{-1}.$$

In other words, if we think of $\mathrm{Mat}_{m\times n}(\mathbb{C})$ as a $\mathbb{C}$-vector space, then for each $g \in G$, the map $A \mapsto S_g A R_g^{-1}$ is a linear operator on this vector space. So we're sending g to the $mn \times mn$ matrix corresponding to that linear operator (which we denote by C_g).

Proof. It suffices to check that $C_{gh} = C_g C_h$ for all g and h. But for any $A \in \mathrm{Mat}_{m\times n}(\mathbb{C})$, we have

$$C_{gh}(A) = S_{gh} A R_{gh}^{-1} = S_g S_h A R_h^{-1} R_g^{-1} = C_g(C_h(A)).$$

So then $C_{gh} = C_g C_h$, as desired. □

In this new representation, $\mathrm{Hom}_G(\rho, \psi)$ is exactly the space of G-invariant vectors (note that here "vectors" means matrices of dimension $m \times n$, since the vector space our representation is acting on is actually the space of such matrices).

We can also describe this construction without using matrices — let $M = \mathrm{Hom}_{\mathbb{C}}(V, W)$ be the space of *all* linear maps $V \to W$ (which we thought of as $\mathrm{Mat}_{m\times n}(\mathbb{C})$ when writing down the construction in matrices). Then M is a vector space over $\mathbb{C}$. So we can define a representation γ acting on M, where for each $g \in G$, we send g to the linear map $\gamma_g : M \to M$ which sends $E \mapsto \psi_g E \rho_g^{-1}$ for all $E \in M$. As before, $\mathrm{Hom}_G(\rho, \psi)$ is exactly the space of G-invariant vectors in γ.

Student Question. *Why is $E \mapsto \psi_g E \rho_g^{-1}$ a linear map?*

Answer. *Thinking in terms of matrices, we want to see that $A \mapsto S_g A R_g^{-1}$ is a linear map. But this follows from the distributive property of matrix multiplication — for example, we have*

$$S_g(A + B)R_g^{-1} = S_g A R_g^{-1} + S_g B R_g^{-1}.$$

Proposition 6.5
For the representation γ described above, we have

$$\chi_\gamma = \chi_\psi \overline{\chi_\rho}.$$

The proposition quickly reduces to a statement about matrices:

Lemma 6.6
Let A and B be $n \times n$ and $m \times m$ matrices. Consider the linear map $A \otimes B$ from $\mathrm{Mat}_{m\times n}(\mathbb{C})$ to itself defined as

$$A \otimes B : E \mapsto BEA.$$

Then we have

$$\mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A) \cdot \mathrm{Tr}(B).$$

Proof. The space of matrices has a basis consisting of the matrices E_{ij} which have a 1 in the ith row and jth column, and 0's everywhere else (for all $1 \le i \le m$ and $1 \le j \le n$).

But by straightforward computation, we can see that $BE_{ij}A$ has $b_{ii}a_{jj}$ in its ith row and jth column — this means $BE_{ij}A$ is $b_{ii}a_{jj}E_{ij}$, plus some entries corresponding to the other basis elements. So if we write out the matrix corresponding to $A \otimes B$ (in the basis formed by the E_{ij}), the diagonal entry corresponding to E_{ij} is $b_{ii}a_{jj}$. The trace of $A \otimes B$ is then the sum of the diagonal entries of this matrix, which is

$$\mathrm{Tr}(A \otimes B) = \sum_{i=1}^{m}\sum_{j=1}^{n} b_{ii}a_{jj} = \sum_{i=1}^{m} b_{ii} \sum_{j=1}^{n} a_{jj} = \mathrm{Tr}(B) \cdot \mathrm{Tr}(A). \qquad \square$$

Proof of Proposition 6.5. Using the above lemma, for all $g \in G$ we have

$$\chi_\gamma(g) = \mathrm{Tr}(\rho_{g^{-1}} \otimes \psi_g) = \mathrm{Tr}(\rho_{g^{-1}}) \cdot \mathrm{Tr}(\psi_g) = \chi_\rho(g^{-1}) \cdot \chi_\psi(g).$$

We saw earlier that $\chi_\rho(g^{-1}) = \overline{\chi_\rho(g)}$ — this follows from the fact that $\chi_\rho(g^{-1})$ is the sum of the eigenvalues of $\rho_{g^{-1}} = \rho_g^{-1}$, which are the inverses of the eigenvalues of ρ_g, and since all these eigenvalues have magnitude 1, their inverses are also their conjugates. So

$$\chi_\gamma(g) = \overline{\chi_\rho(g)} \cdot \chi_\psi(g)$$

for all $g \in G$, as desired. $\square$

6.4 Orthonormality of Characters

We have now developed the tools that we can use to prove one part of the main theorem stated earlier, the orthonormality of irreducible characters.

Proposition 6.7
The characters of the irreducible representations of G are orthonormal.

Proof. Let ρ and ψ be irreducible representations, acting on the spaces V and W. Then our Hermitian form is

$$\langle \chi_\psi, \chi_\rho \rangle = \frac{1}{|G|} \sum_{g\in G} \overline{\chi_\rho(g)} \chi_\psi(g),$$

so we want to check this expression is 0 if $\rho \ncong \psi$ and 1 if $\rho \cong \psi$.

But we saw a representation γ, acting on the space $M = \mathrm{Hom}_{\mathbb{C}}(V, W)$, whose character is exactly the expression inside the sum! So we can rewrite our sum as

$$\langle \chi_\psi, \chi_\rho \rangle = \frac{1}{|G|} \sum_{g\in G} \chi_\gamma(g) = \mathrm{Tr}\left(\frac{1}{|G|} \sum_{g\in G} \gamma_g\right).$$

(Here we used the fact that $\operatorname{Tr}(A+B) = \operatorname{Tr} A + \operatorname{Tr} B$ — our original sum calculates the traces of each γ_g *first*, and then averages them, but we can instead average the γ_g and *then* compute the trace).

Now let's consider what the linear operator on M given by $\sum_{g \in G} \gamma_g / |G|$ looks like; denote this operator by f.

Since f is an averaging operator (we're averaging over all $g \in G$), it must always output a G-invariant vector — similar to the averaging trick we saw earlier, for any $v \in M$ and any fixed $h \in G$, we have

$$\gamma_h(f(v)) = \gamma_h \cdot \frac{1}{|G|} \sum_{g \in G} \gamma_g v = \frac{1}{|G|} \sum_{g \in G} \gamma_h \gamma_g v = \frac{1}{|G|} \sum_{g \in G} \gamma_{hg} v = \frac{1}{|G|} \sum_{g \in G} \gamma_g v = f(v),$$

since if we fix h and let g range over all elements in G, then hg also ranges over all elements in G.

But we know what the G-invariant vectors are! Recall that vectors in the space M that γ acts on are actually homomorphisms $V \to W$, and by the way γ was defined, the vectors in M which are G-invariant are exactly the homomorphisms $V \to W$ which are G-equivariant — which are described by Schur's Lemma.

If $\rho \not\cong \psi$, then by Schur's Lemma, then $\operatorname{Hom}_G(\rho, \psi)$ only contains the zero map, so the only G-invariant vector in M is the zero vector. So since our operator f sends every vector $v \in M$ to *some* G-invariant vector, it must actually send every vector v to 0. So f is the zero operator, and

$$\langle \chi_\psi, \chi_\rho \rangle = \operatorname{Tr}(f) = 0.$$

Now suppose $\rho \cong \psi$. Then by Schur's Lemma, $\operatorname{Hom}_G(\rho, \psi)$ only contains scalar maps; so there's only one G-invariant vector in M up to scaling (the identity matrix and its scalar multiples). Let u be such a (nonzero) invariant vector.

This means f must send every vector $v \in M$ to some scalar multiple of u (since f sends every v to some G-invariant vector, and the only G-invariant vectors are multiples of u). On the other hand, since we're averaging and u is already G-invariant, f must send u to itself — we have

$$\frac{1}{|G|} \sum_{g \in G} \gamma_g u = \frac{1}{|G|} \sum_{g \in G} u = u.$$

Now choose a basis for M whose first element is u. Then in this basis, the matrix for f is of the form

$$\begin{bmatrix} 1 & * & * & \cdots & * \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

which has trace 1. So in this case,

$$\langle \chi_\psi, \chi_\rho \rangle = \operatorname{Tr}(f) = 1.$$

□

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