7 Proof of the Main Theorem

7.1 Review: Orthonormality of Characters

Last time, we proved the orthonormality of characters of irreducible representations — if we have a finite group G and $\rho_1, \ldots, \rho_n$ is the full list of irreducible representations of G up to isomorphism, then if we use χ_i to denote χ_{ρ_i}, we have

$$\langle \chi_i, \chi_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

In order to prove this, we interpreted $\langle \chi_i, \chi_j \rangle$ as the trace of an averaging operator, acting on the space of linear maps $\text{Mat}_{m\times n}(\mathbb{C}) = \text{Hom}_{\mathbb{C}}(V_i, V_j)$, where V_i and V_j are the vector spaces that ρ_i and ρ_j act on.

From orthonormality, we immediately get a few corollaries:

Corollary 7.1
Any representation $\rho : G \to \text{GL}(V)$ can be split as a sum of irreducibles as

$$\rho \cong \bigoplus \rho_i^{n_i},$$

where for each k,

$$n_k = \langle \chi_\rho, \chi_k \rangle.$$

Recall that last class, we saw a different formula for the n_i, which was quite abstract — it involved the dimensions of $\text{Hom}_G(\rho_k, \rho)$. In contrast, this formula is quite concrete — it's easy to calculate the pairings $\langle \chi_\rho, \chi_k \rangle$.

Proof. We know ρ can be written in this form for some coefficients n_i, by Maschke's Theorem. But then

$$\chi_\rho = \sum n_i \chi_i,$$

so by linearity we have

$$\langle \chi_\rho, \chi_k \rangle = \sum n_i \langle \chi_i, \chi_k \rangle = n_k,$$

since orthonormality implies that $\langle \chi_i, \chi_k \rangle$ is 0 for $i \neq k$ and 1 for $i = k$. □

Using this, we can get a few concrete results:

Example 7.2
The dimension of the space of invariant vectors in ρ is

$$\frac{1}{|G|} \sum_{g \in G} \chi_\rho(g).$$

Proof. This dimension is the multiplicity of the trivial representation χ_1 when we decompose ρ into a sum of irreducibles, which is n_1. This is because ρ acts trivially on the space of invariant vectors, by definition, and so each basis vector corresponds to one copy of the trivial representation. But the character of the trivial representation is $\chi_1(g) = 1$ for all g, so using Corollary 7.1, we have

$$n_1 = \langle \chi_\rho, \chi_1 \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_\rho(g).$$ □

Corollary 7.3
If $\rho = \bigoplus \rho_i^{n_i}$ as before, then

$$\langle \chi_\rho, \chi_\rho \rangle = \sum n_i^2.$$

In particular, ρ is irreducible if and only if $\langle \chi_\rho, \chi_\rho \rangle = 1$.

Proof. Using linearity, we can expand

$$\langle \chi_\rho, \chi_\rho \rangle = \sum_i \sum_j n_i n_j \langle \chi_i, \chi_j \rangle = \sum n_i^2,$$

since again by orthonormality, $\langle \chi_i, \chi_j \rangle$ is 0 when $i \neq j$ and 1 when $i = j$. (Intuitively, if we have any pairing and we take a basis of vectors which are orthonormal under that pairing — here, the χ_i — then it becomes the usual pairing on $\mathbb{C}^n$.)

The second statement is then clear because the n_i are nonnegative integers, so the only way for their sum of squares to be 1 is if one is 1 and the rest are 0. □

7.2 The Regular Representation

We've already proved part of the main theorem — we've shown that the characters χ_i are orthonormal, which means they are linearly independent. But to show that they form a basis for the space of class functions, we also need to show that they span that space. To do this — and to prove the sum of squares formula stated earlier as well — we'll introduce the regular representation.

If G acts on a finite set X of n elements, then we can form a matrix representation of G of dimension n, where G acts by permutation matrices — we index the basis vectors by elements of X, and for each $g \in G$, we map g to the permutation matrix that describes how g acts on X.

Example 7.4
As we've seen earlier, S_n acts on the set $\{1, 2, \dots, n\}$. This gives a n-dimensional representation of S_n — the permutation representation, where each permutation is mapped to its corresponding permutation matrix.

Given a group G, there can be many interesting sets X that it acts upon. But there's one set that we automatically always have; namely, G itself. Every group G acts on itself by left multiplication, where an element $g \in G$ sends $h \mapsto gh$. We can use this to form a representation of G of dimension $|G|$:

Definition 7.5
Let V be a vector space with basis $\{v_h\}$ indexed by elements $h \in G$. Then the **regular representation** of G is the representation $\rho : G \to \mathrm{GL}(V)$ such that for all $g \in G$, ρ_g is the linear operator on V sending $v_h \mapsto v_{gh}$ for all $h \in G$.

So in the regular representation, each $g \in G$ is sent to the permutation matrix of how left multiplication by g permutes the elements of G.[‖]

Example 7.6
In the group $\mathbb{Z}/3\mathbb{Z}$, operating (in this case the group operation is addition) on the left by $\overline{1}$ corresponds to the permutation $\overline{0} \mapsto \overline{1}$, $\overline{1} \mapsto \overline{2}$, $\overline{2} \mapsto \overline{0}$. So in its regular representation, $\overline{1}$ acts by the permutation matrix

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

‖ For left multiplication, $\rho_{g_1}\rho_{g_2} = \rho_{g_1 g_2}$, so ρ is a homomorphism. Using right multiplication rather than left multiplication, which would send $\rho'_g : v_h \longmapsto v_{hg}$, would be a anti-homomorphism rather than a homomorphism. That is, $\rho'_{g_1}\rho'_{g_2}(v_h) = v_{hg_2g_1} = \rho'_{g_2g_1}(v_h)$.

Example 7.7
In S_3, left multiplication by (12) swaps (1) and (12), swaps (23) and (123), and (13) and (132). So in the regular representation, (12) acts by the block diagonal matrix

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

where the rows and columns are indexed by elements of S_3 in the order (1), (12), (23), (123), (13), (132).

It's clear from the definition that the regular representation has exactly one invariant vector up to scaling, the sum of all basis elements. This is because if $\sum a_h v_h$ is an invariant vector, then we must have

$$\sum a_h v_h = \rho_g \left(\sum a_h v_h\right) = \sum a_h v_{gh}$$

for all $g \in G$, which implies that $a_h = a_{gh}$ for all g and h, and therefore all a_h are equal. In particular, the regular representation is not irreducible unless G is trivial (the regular representation has an invariant vector, so it must have the trivial representation in its decomposition).

Note 7.8
It's also possible to think of elements of V as $\mathbb{C}$-valued functions on G — instead of thinking of them as linear combinations of abstract basis vectors, we can think of $\sum a_h v_h$ as the function mapping $h \mapsto a_h$ for all $h \in G$.

We'll now use ρ to denote the regular representation.

Guiding Question
What can we say about the character of ρ and its decomposition into irreducibles?

This turns out to have a simple answer.

Proposition 7.9
The character of the regular representation is

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

Proof. The trace of a permutation matrix is its number of fixed points (a permutation matrix consists only of 0's and 1's, and 1's on its diagonal correspond to fixed points).

It's clear that the permutation corresponding to 1 fixes all elements of G (since multiplication by 1 doesn't change any element), so $\chi_\rho(1) = |G|$. Meanwhile, for all $g \in G$ other than 1, the permutation corresponding to g has no fixed points — if h were a fixed point, then we would have $h = gh$, which implies $g = 1$. So then $\chi_\rho(g) = 0$ for all $g \neq 1$. ∎

Using this, we can decompose ρ into a sum of irreducibles pretty easily, using Corollary 7.1 — since our character has such a simple form, it's not hard to compute its pairing with anything.

Proposition 7.10
The regular representation decomposes into irreducibles as

$$\rho \cong \bigoplus \rho_i^{d_i},$$

where d_i denotes the dimension of ρ_i.

Proof. By Corollary 7.1, we know $\rho = \bigoplus \rho_i^{n_i}$, where

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

(all terms involving $g \neq 1$ disappear, since $\chi_\rho(g) = 0$). But $\rho_i(1)$ is the identity matrix of dimension d_i, which has trace d_i — so $\chi_i(1) = d_i$, which means $n_i = d_i$ as well. □

Example 7.11
In an abelian group, all dimensions of irreducible representations are 1, so $\rho \cong \rho_1 \oplus \cdots \oplus \rho_n$.

From this, we can immediately deduce the sum of squares formula stated in the main theorem.

Proposition 7.12
If the irreducible representations of G have dimensions $d_1, \ldots, d_n$, then we have

$$|G| = d_1^2 + \cdots + d_n^2.$$

Proof. We can compare the dimensions on the two sides of Proposition 7.10. On the left-hand side, we have $\dim(\rho) = |G|$. Meanwhile, on the right-hand side, for each i we have d_i copies of ρ_i, which itself has dimension d_i — this contributes d_i^2 to the dimension (since when we take the direct sum of two representations, we add their dimension). So then

$$|G| = \dim(\rho) = \dim\left(\bigoplus \rho_i^{d_i}\right) = \sum d_i^2,$$

which gives the desired equality. □

7.3 Span of Irreducible Characters

Finally, we'll again use the regular representation to prove that the characters of irreducible representations span the space of class functions. We know that these characters are linearly independent (since they're orthonormal), so it will then follow that they form a *basis* for the space of class functions (the first statement of our main theorem).

In order to prove this, we'll show that any class function can be written in the following form:

Proposition 7.13
For any class function f, we have

$$f = \sum \langle f, \chi_i \rangle \chi_i.$$

Note that we already know this statement is true when f is the character of any representation. More generally, if we can write $f = \sum n_i \chi_i$, then by orthonormality we know that $n_i = \langle f, \chi_i \rangle$ for each i. So Proposition 7.13 essentially tells us that this formula really works for *all* class functions. Note that the coefficients $\langle f, \chi_i \rangle$ are all in $\mathbb{C}$, so Proposition 7.13 implies that the χ_i span the space of class functions — this proposition would follow immediately if we already knew that the χ_i span the space of class functions, but we can actually use it to prove that statement instead.

Proposition 7.13 is equivalent to the following statement:

Proposition 7.14
If f is a class function and $\langle f, \chi_k \rangle = 0$ for all k, then f is the zero function.

First, to make this equivalence between Proposition 7.13 and Proposition 7.14 more explicit, we'll write out the details of how the second implies the first. (For the other direction, the first implies the second because if the χ_i do span the space of class functions and f has zero pairing with each one, then f also has zero pairing with *itself*. But the space of class functions under $\langle -, - \rangle$ is a Hermitian space, meaning that $\langle f, f \rangle > 0$ unless $f = 0$.)

Proof of Proposition 7.13. Given any class function, we can define $f^* = \sum \langle f, \chi_i \rangle \chi_i$. But then $f - f^*$ has zero pairing with each χ_k, since

$$\begin{aligned}\langle f - f^*, \chi_k \rangle &= \langle f, \chi_k \rangle - \left\langle \sum \langle f, \chi_i \rangle \chi_i, \chi_k \right\rangle \\ &= \langle f, \chi_k \rangle - \sum \langle f, \chi_i \rangle \langle \chi_i, \chi_k \rangle \\ &= \langle f, \chi_k \rangle - \langle f, \chi_k \rangle \\ &= 0\end{aligned}$$

by orthonormality. (Intuitively, since f^* is a linear combination of the χ_i, the pairing of f^* with each χ_k is the coefficient of χ_k, which we *constructed* to be the same as the pairing of f with χ_k.) But using Proposition 7.14, the only class function that has zero pairing with every χ_k is the zero function, so we must have $f - f^* = 0$, and therefore $f = f^*$. □

In order to prove Proposition 7.14, we'll first introduce a bit of convenient notation:

Definition 7.15
For a function $f : G \to \mathbb{C}$ (not necessarily a class function) and a representation $\rho : G \to \mathrm{GL}(V)$, define

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

Note that $\rho(f)$ is in $\mathrm{End}(V)$ (or equivalently, in $\mathrm{Mat}_{n \times n}(\mathbb{C})$ if we write down a basis and use matrix representations), since each ρ_g is in $\mathrm{End}(V)$ and the coefficients $f(g)$ are scalars. We can think of this definition as extending the definition of ρ to *linear combinations* of group elements (in the obvious way), instead of just the group elements themselves — in this interpretation, the function f stores the coefficient of each group element in the linear combination.

We can make a few observations about this definition:

Lemma 7.16
For any representation ρ and function f, we have $\mathrm{Tr}\, \rho(f) = \langle \chi_\rho, \overline{f} \rangle$.

Proof. This follows directly from the linearity of trace — plugging in the definitions and using linearity, we have

$$\langle \chi_\rho, \overline{f} \rangle = \sum_{g \in G} (\chi_\rho(g) f(g)) = \sum_{g \in G} (f(g) \,\mathrm{Tr}\, \rho_g) = \mathrm{Tr} \left(\sum_{g \in G} f(g) \rho_g \right) = \mathrm{Tr}\, \rho(f).$$

(The reason we have $\overline{f}$ and not f here is because we conjugate the second element of the pairing.) □

Lemma 7.17
If f is a class function, then $\rho(f) \in \mathrm{End}_G(\rho)$.

Proof. Recall that class functions are functions which are invariant under conjugation, meaning that for each conjugacy class, they take the same value on all its elements.

But this means $\rho(f)$ must be invariant under conjugation as well — more explicitly, for any $g \in G$, we have

$$\rho_g\rho(f)\rho_g^{-1} = \sum_{h\in G} f(h)\rho_g\rho_h\rho_g^{-1} = \sum_{h\in G} f(h)\rho_{ghg^{-1}} = \sum_{h\in G} f(h)\rho_h = \rho(f),$$

since the new sum just permutes elements within each conjugacy class, and $f(ghg^{-1}) = f(h)$ for all h. We can rearrange this to

$$\rho_g\rho(f) = \rho(f)\rho_g$$

for all $g \in G$, so $\rho(f)$ is indeed G-equivariant. □

Using these observations, we can now prove our claim, that the only class function f which has zero pairing with every χ_i is the zero function.

Proof of Proposition 7.14. For each i, we are given that $\langle \chi_i, f\rangle = 0$, so by Lemma 7.16 we also have

$$\operatorname{Tr}\rho_i(\overline{f}) = 0.$$

But we also know that $\rho_i(\overline{f}) \in \operatorname{End}_G(\rho_i)$. And by Schur's Lemma (since ρ_i is irreducible), the only elements of $\operatorname{End}_G(\rho_i)$ are scalar matrices! So then $\rho_i(\overline{f})$ is a scalar matrix with trace 0; this means it's the zero matrix.

So now we know that $\rho_i(\overline{f})$ is the zero matrix for all irreducible representations. But by Maschke's Theorem, every representation is the direct sum of irreducible representations — so this means $\rho(\overline{f})$ is the zero matrix for *any* representation ρ. (If we write ρ as a direct sum of irreducibles ρ_i, then $\rho(\overline{f})$ is the corresponding direct sum of the matrices $\rho_i(\overline{f})$, and a direct sum of zero matrices is also the zero matrix.)

Now we'd like to use this to conclude that f itself is 0. To do so, we can take ρ to be the regular representation. It turns out that we can essentially just read off the function by looking at its action in the regular representation — in particular, f acts on the basis vector v_1 as

$$\rho(f)(v_1) = \sum_g f(g)\rho_g(v_1) = \sum_g f(g)v_g.$$

Since $\rho(f)$ is the zero map, then the right-hand side must be zero as well; so $f(g) = 0$ for all $g \in G$. □

This concludes our proof of the Main Theorem — we have proven that the irreducible characters χ_i are orthonormal and form a basis for the space of class functions, and the dimensions of the irreducible representations d_i satisfy $\sum d_i^2 = |G|$. The only remaining statement which we have not proven is that d_i divides $|G|$ for each i. We will not discuss this proof in class, but it is posted on Canvas; in these notes, a writeup of this proof is included in the appendix.

7.4 Generalizations to Compact Groups

We've worked with representations of *finite* groups, but much of this generalizes to *compact* subgroups of $\mathrm{GL}_n(\mathbb{C})$, such as $\mathrm{U}(n)$ and $\mathrm{O}(n)$. In this case, we consider *continuous* representations.

Example 7.18

Consider $\mathrm{U}(1) = \{z \in \mathbb{C}^\times \mid |z| = 1\}$. The irreducible representations are all one-dimensional (as in the finite case, this has to do with the fact that the group is abelian), and are indexed by integers, where $\rho_n : z \mapsto z^n$.

All functions $f : \mathrm{U}(1) \to \mathbb{C}$ are class functions, and we can think of such a function as a function $f : \mathbb{R} \to \mathbb{C}$ which is 2π-periodic (by thinking of $\mathrm{U}(1)$ as the unit circle). Then if we try to decompose such a function f in the same way as Proposition 7.13, we'll get its **Fourier series** — under some reasonable conditions on f (in particular, here we require it to be continuous), the expression $\sum \chi\langle f, \chi\rangle$ ends up being the Fourier series of f.

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