

Contents

1 Basic Definitions & Results	2
1.1 What is a representation?	2
1.2 Character theory	4
1.3 Fourier analysis on finite groups	9

Literature

References are roughly listed in ascending order of depth and complexity, i.e. for the quickest introduction to representation theory (at the cost of not getting an intuitive understanding why definitions are chosen that way, etc.) you might want to consult the first reference.

- [Ber16, Section 6]: A minimal introduction to representation theory up to the theorem establishing that characters span the space of class functions; interesting applications for mixing times are provided; it follows quite closely the following source:
- [Dia88]: Note that [Ber16, Section 6] is a condensed version of this. Here there are more explanations, a lot of examples (mainly for mixing times, however) and a section devoted to the representation theory of the symmetric group.
- [Sag13]: A very pedagogical (i.e. slow & detailed) approach. Consult for reference and if proofs found elsewhere are unclear or for really good understanding. Also seems to cover a lot more results specific to S_n .
- [FH13]: A more abstract and quite comprehensive book with a very systematic approach. Although it starts with the representation theory of finite groups and also deals with symmetric functions, large parts of the books are on Lie groups. Helpful if you want to read algebra papers on representation theory as they introduce the necessary language* that is (fortunately) left out in the other sources.

*E.g. they say homomorphism instead of writing $\rho(s)\rho(t) = \rho(st)$, they say the kernel of some map is a G -module, instead of saying that it is closed under action of G , they say “ G -module homomorphism from V to W ” instead of “linear map f from V to W s.t. $f \circ \rho_V(s) = \rho_W(s) \circ f$ for all $s \in G$ ”, etc.

- [OV96]: This article is nowhere as comprehensive as the above books, but it claims to have found a better approach to the representation theory of symmetric groups, making Young diagrams, etc. appear “more naturally”. I am dubious about that claim, however, since they need a lot of heavy machinery from abstract algebra, which kind of defeats the purpose of building the same theory bottom-up.*

Books I didn’t look into (yet):

- [JK81]: Cited in [AK13]. Did not look into it yet, but it seems to be cited frequently.
- [Ser77]: A more comprehensive version of [Dia88] (in fact Diaconis follows this one, but leaves out a lot). Seems to be a very solid introduction, also addresses induced representations and their characters, as well as Artin’s and Brauer’s theorems (which are for example not dealt with in [FH13] and several other long books).

Motivation

In this text we only deal with finite groups G . Note that finite groups can always be seen as subgroups of the symmetric group S_n . So we want to understand S_n .

1 Basic Definitions & Results

This section follows [Ber16] and [Dia88] very closely.

1.1 What is a representation?

Definition 1.1 (Representation). *A representation of G on a d_ρ -dimensional complex vector space V is a homomorphism $\rho : G \rightarrow GL(V)$, i.e.*

$$\forall s, t \in G : \quad \rho(s)\rho(t) = \rho(st).$$

We say

*The abstract algebra approach seems to be quite interesting, however, since their results seem to be frequently overlooked; see e.g. <https://arxiv.org/pdf/1202.4976.pdf>.

- ρ has dimension d_ρ ,
- if there exists a linear subspace $W \subseteq V$ that is stable under G (i.e. $\forall s \in G : \rho(s)W \subseteq W$), then ρ restricted to W gives a subrepresentation and
- if ρ admits no non-trivial subrepresentations (i.e. $W = V$ or $W = \{0\}$), then ρ is called irreducible.

Remark 1.2 (On G -modules). Every fixed representation ρ gives V the structure of a G -module. We use the definition from [Sag13]: Let V be a (complex) vector space and G be a group, then V is a G -module if there is a multiplication gv of elements of V by elements of G such that (1) $gv \in V$, (2) $g(\alpha v + \beta w) = \alpha(gv) + \beta(gw)$, (3) $(gh)v = g(hv)$ and (4) $id v = v$ for all $g, h \in G, v, w \in V, \alpha, \beta \in \mathbb{C}$. In fact we can turn any finite set S on which G acts (i.e. satisfies (1),(3),(4)) into a G -module by considering $\mathbb{C}S = \mathbb{C}[s_1, \dots, s_N]$, the vector space generated by S over $\mathbb{C}$, consisting of all formal linear combinations $c_1 s_1 + \dots + c_N s_N$.

This justifies the occasional use of $g \cdot v$ or just gv instead of $\rho(g)(v)$. This will be important especially in later sections.

{ex:repsOfSn}

Example 1.1. Let $G = S_n$. There are some representations that will appear again and again, so we introduce them here:

1. The trivial representation is 1-dimensional. It assigns to each permutation the identity map, i.e.

$$\rho(\pi)x = x \quad \text{for } \pi \in S_n.$$

2. The alternating representation is also 1-dimensional. It is defined by

$$\rho(\pi)x = \text{sgn}(\pi)x \quad \text{for } \pi \in S_n.$$

Recall that sgn is a homomorphism on permutations to see that this is indeed well defined.

3. The permutation representation is an n -dimensional representation. Define its action on the basis vectors $\{e_i\}_{i=1}^n$ of $\mathbb{C}^n$:

$$\rho(\pi)e_i = e_{\pi(i)} \quad \text{for } \pi \in S_n.$$

This is not an irreducible representation: Choose $W = \text{span}\{(1, 1, \dots, 1)\}$ or its orthogonal complement $W' = \{x \in \mathbb{C}^n : \sum_i x_i = 0\}$ and note that both are invariant under permuting basis vectors.

Intuitively it should be enough to understand what is going on in invariant subspaces, since their sum constitutes the whole space. To make this precise we first need to clarify how a representation into $V \oplus W$ acts: Let ρ_V, ρ_W be representations into V and W , respectively. Then their sum $\sigma := \rho_V \oplus \rho_W$ is defined by its natural action

$$\sigma(s)(v + w) := \rho_V(s)(v) + \rho_W(s)(w) \quad \text{for } v \in V, w \in W.$$

Theorem 1.3 (Maschke's theorem). *Every representation ρ is the finite sum of irreducible representations.*

By induction on the dimension of ρ it suffices to prove the following claim:

Lemma 1.4. *Let $\rho : G \rightarrow GL(V)$ be a representation of G . Suppose $W \subset V$ is stable, then there is a^{*} complement W' (i.e. a subspace such that $W \cap W' = \{0\}$ and $W + W' = V$) such that W' is also stable.*

Proof. The idea is to define a scalar product $\langle \cdot, \cdot \rangle$ on V that is invariant under ρ in the sense that $\langle \rho(s)v, \rho(s)w \rangle = \langle v, w \rangle$ for all $v, w \in V, s \in G$. Given a subspace W of V we can now consider its orthogonal (w.r.t. $\langle \cdot, \cdot \rangle$) complement W' . Then we have

$$\langle w, \rho(s)w' \rangle = \langle \rho(s^{-1})w, w' \rangle = 0,$$

where the first equality was by invariance of $\langle \cdot, \cdot \rangle$ under ρ and the second equality by orthogonality and invariance of W under ρ . In particular this equation tells us that W' is stable too (since after applying $\rho(s)$ it is still orthogonal to W).

It remains to check that such an inner product exists: Take $(\cdot, \cdot)$ to be any inner product on V and define $\langle v, w \rangle := \sum_{s \in G} (\rho(s)v, \rho(s)w)$. The required invariance follows from the group properties. $\square$

In fact this proof also gave us the existence of a basis of V such that the matrix representation of $\rho(s)$ is unitary. We will assume such a choice of basis throughout the text without mentioning it anymore.

1.2 Character theory

Ultimately we want to use representations to be able to calculate certain functionals of G , i.e. $f : G \rightarrow \mathbb{C}$. Endowing these functionals with their natural vector space

^{*}This is not unique since we do not require it to be the orthogonal complement!

structure, it is natural to look for a “good” basis. Clearly one basis would be the set of delta functions centered on group elements $s \in G$. This is not a particularly useful basis since it does not make any use of the group structure at all. A more useful one turns out to be a certain subset of “characters” which span a slightly smaller, but more natural space of “class functions”:

Definition 1.5 (Character). *Let ρ be a representation of G into V . Then the character associated to ρ is the function $\chi_\rho : G \rightarrow \mathbb{C}$ defined by*

$$\chi_\rho(s) := \text{tr}(\rho(s)) \quad \text{for } s \in G.$$

Example 1.2. *The character χ_ρ associated to the permutation representation ρ of S_n is the function giving the number of fixed points of $s \in S_n$.*

We state now a few useful identities:

Lemma 1.6. *We have (1) $\chi_\rho(e) = d_\rho$, (2) $\chi_\rho(t^{-1}st) = \chi_\rho(s)$, (3) $\chi_\rho(s^{-1}) = \overline{\chi_\rho(s)}$.*

Proof. The proofs of 1. and 2. are trivial. To see 3., consider the basis where the representations are unitary: Here we have $\rho(s^{-1}) = \rho(s)^{-1} = \rho(s)^*$. $\square$

Note that 2. states that characters are constant on conjugacy classes, i.e. any character evaluated at s gives the same value as evaluated at gsg^{-1} . So we cannot hope for characters to span the space of all functionals on G . The next best thing we can hope for is that they span the space of all functions that are constant on conjugacy classes.

Definition 1.7 (Class functions). *Functions $f : G \rightarrow \mathbb{C}$ that are constant on conjugacy classes (i.e. $f(s) = f(gsg^{-1})$ for all $s, g \in G$) are called class functions. We denote the space of class functions by $\mathbb{C}_{\text{class}}(G)$ and turn it into a Hilbert space by endowing it with the natural inner product*

$$(f|g) := \frac{1}{|G|} \sum_{s \in G} f(s)\overline{g(s)}.$$

In order to prove statements like Theorem 1.10 we first need Schur’s lemma:

{lem:Schur}

Lemma 1.8 (Schur’s lemma). *Let ρ_V, ρ_W be two irreducible representations into V, W , respectively. Let $f : V \rightarrow W$ be linear such that*

$$f \circ \rho_V(s) = \rho_W(s) \circ f \quad \text{for all } s \in G.$$

Then:

- (a) If ρ_V, ρ_W are not equivalent, then $f = 0$.
- (b) If $V = W$ and $\rho_V = \rho_W$, then $f = \lambda I$ for some $\lambda \in \mathbb{C}$.

Another useful way to state Schur's Lemma is the following:

{cor:Schur}

Corollary 1.9 (Another version of Schur's lemma). *Same setting as in Lemma 1.8.*

Let $h : V \rightarrow W$ be linear and define $\tilde{h} : V \rightarrow W$ to be

$$\tilde{h} := \frac{1}{|G|} \sum_{s \in G} \rho_W(s^{-1}) h \rho_V(s).$$

Then:

- (a) If ρ_V, ρ_W are not equivalent, then $\tilde{h} = 0$.
- (b) If $V = W$ and $\rho_V = \rho_W$, then $\tilde{h} = \lambda I$ with $\lambda = \text{tr}(h)/d_\rho$.

This follows from Lemma 1.8 by noting that for all $t \in G$ we have $\rho_W(t^{-1})\tilde{h}\rho_V(t) = \sum_{s \in G} \rho_W((st)^{-1})h\rho_V(st) = \tilde{h}$.

Proof of Lemma 1.8. The kernel of f is invariant: $f(v) = 0 \Rightarrow \rho_W(s)(f(v)) = 0$, but $\rho_W(s)(f(v)) = f(\rho_V(s)(v))$ by assumption; thus $f(\rho_V(s)(v)) = 0$ for all $s \in G, v \in \text{Ker } f$.

The image of f is invariant: $w = f(v)$, then $\rho_W(s)(w) = f(\rho_V(s)(v))$ is in the image too.

By irreducibility the kernel and the image are either the whole space or $\{0\}$. For (a) suppose $f \neq 0$. Then $\text{Ker } f = \{0\}$ and $\text{Im } f = W$ so f is an isomorphism. For (b) suppose $f \neq 0$ and $V = W$. (If $f = 0$ then we are done.) Then f has some non-zero eigenvalue λ . Define $\tilde{f} := f - \lambda I$. This map $\tilde{f}$ satisfies $\tilde{f} \circ \rho_V = \rho_W \circ \tilde{f}$ ($\rho_V = \rho_W$ by assumption) and has a non-trivial kernel. So by (a) we have $\tilde{f} = 0$ and hence $f \equiv \lambda I$. □

{thm:charsOrthogon}

Theorem 1.10. *The characters associated to irreducible representations are orthonormal functions with respect to $(\cdot | \cdot)$.*

Proof. Fix any two irreducible representations ρ, ρ' and denote their characters by χ, χ' , respectively. Let $r_{ij}(t), r'_{ij}(t)$ denote the entries of the matrix $\rho(t), \rho'(t)$, respectively.

Assume $\rho \neq \rho'$. We need to show

$$(\chi, \chi') \stackrel{(\text{def})}{=} \sum_{ij} (r_{ii} | r'_{jj}) \stackrel{(\text{def})}{=} \sum_{ij} \sum_{t \in G} r_{ii}(t) r'_{jj}(t) \stackrel{!}{=} 0.$$

Denote by $\tilde{x}_{ij}$ the entries of the matrix given by $\tilde{h}$ for some choice of h (it will be fixed later). By Corollary 1.9 we have that $\tilde{h} = 0$, so in particular $\tilde{x}_{ij} = 0$. But using the definition of $\tilde{h}$ we can also write

$$\tilde{x}_{ij} = \frac{1}{|G|} \sum_{t \in G} \sum_{kl} r'_{ik}(t^{-1}) x_{kl} r_{lj}(t),$$

where x_{kl} are the entries of the matrix induced by h . Choosing h such that $x_{kl} = \delta_{i=k, j=l}$ we get

$$0 = \sum_{t \in G} r'_{ii}(t^{-1}) r_{jj}(t).$$

It follows that $\sum_{ij} \sum_{t \in G} r_{ii}(t) r'_{jj}(t) = 0$.

Similar calculations for the case where $\chi = \chi'$ finish the proof. $\square$

We will also need the following consequences:

Theorem 1.11. *Let ρ be a (not necessarily irreducible) representation into V and ψ an irreducible representation into W . The number of times W arises in the decomposition of V is equal to $(\chi_\rho | \chi_\psi)$.*

Proof. We “split V into atoms”, i.e. write $V = W_1 \oplus \cdots \oplus W_m$ with W_i minimal invariant subspaces in the sense that ρ_i is an irreducible representation into W_i . Denote its character by χ_i . Then we have $\chi_\rho = \sum_i \chi_i$ and hence $(\chi_\rho | \chi_\psi) = \sum_{i=1}^m (\chi_i | \chi_\psi)$. Note that χ_i and χ_ψ are characters of irreducible representations, so by Theorem 1.10 we have that $(\chi_i | \chi_\psi)$ is equal to 0 or 1 according to whether $W = W_i$ or not. The result follows. $\square$

Corollary 1.12. *Two representations are equivalent if and only if they have the same character.*

Corollary 1.13. *Let ρ be a representation. Then $(\chi_\rho | \chi_\rho) \in \mathbb{N}$ and equal to 1 if and only if ρ is irreducible.*

This can be used to show that the $(n - 1)$ dimensional representation of S_n introduced in Example 1.1 (3) is irreducible. Since we have found two induced representations of the permutation representation ρ , showing $(\chi_\rho|\chi_\rho) = 2$ implies that both of these induced representations are irreducible. See [Ber16] for a simple proof of $(\chi_\rho|\chi_\rho) = 2$.

We define the *regular representation* ρ of G on $V := \{f : G \rightarrow \mathbb{C}\}$ by its action the basis elements of V as follows:

$$\rho(s)(\delta_t) := \delta_{st}.$$

{cor:multiplicity}

Corollary 1.14. *Every irreducible representation ψ is contained in the regular representation ρ with multiplicity equal to its dimension. In particular there are only a finite number of irreducible representations.*

Proof. Clearly $\chi_\rho(e) = |G|$ (since dimension of V is $|G|$ and $\rho(e)$ is the identity) and if $s \neq e$ then

$$\chi_\rho(s) \stackrel{(\text{def})}{=} \text{tr}(\rho(s)) = \sum_{t \in G} \langle \delta_t, \underbrace{\rho(s)\delta_t}_{\neq \delta_t} \rangle = 0. \quad (1.1) \quad \{\text{eq:regRepChar0}\}$$

Using this we immediately get

$$(\chi_\rho|\chi_\psi) \stackrel{(\text{def})}{=} \frac{1}{|G|} \sum_{s \in G} \chi_\psi(s) \bar{\chi}_\rho(s) = \chi_\psi(e) = d_\psi.$$

□

{cor:dimFormula}

Corollary 1.15. *Taking the sum over all irreducible representations ρ , we have*

$$\sum_{\rho} d_{\rho}^2 = |G|$$

and if $s \neq e$ we also have $\sum_{\rho} d_{\rho} \chi_{\rho}(s) = 0$.

Proof. This follows from the fact that the character associated to the regular distribution (call it χ) can be written as

$$\chi(s) = \sum_{\rho} d_{\rho} \chi_{\rho}(s),$$

by Corollary 1.14. Taking $s = e$ gives the first formula, while taking $s \neq e$ follows from (1.1). □

While Theorem 1.10 is not quite enough to see that the characters of irreducible representations form an orthonormal *basis* for $\mathbb{C}_{\text{class}}$ (maybe more elements are needed to span $\mathbb{C}_{\text{class}}$), this last step can be shown using a generalisation of Fourier analysis – more precisely a Fourier inversion formula. But first we need to introduce what we mean by Fourier analysis on groups and derive basic properties.

1.3 Fourier analysis on finite groups

Definition 1.16. For $f : G \rightarrow \mathbb{C}$ we define its Fourier transform, evaluated at a representation ρ to be

$$\hat{f}(\rho) := \sum_{s \in G} f(s) \rho(s).$$

Note that $\hat{f}(\rho)$ is a matrix (or equivalently a linear map from V to itself). It is a good exercise to take $G = \mathbb{Z}/n\mathbb{Z}$ and see how the usual Fourier transform is recovered in that case. (Cf. [Ber16, p.52])

In particular, this analogy suggests to think of class functions as (generalisations of) periodic functions and of characters associated to irreducible representations as (generalisations of) eigenfunctions of the Laplacian. In particular one might hope that the characters associated to irreducible representations are a orthonormal basis of the class functions. This turns out to be true:

{thm:charBasis}

Theorem 1.17. The characters of irreducible representations form an orthonormal basis of $\mathbb{C}_{\text{class}}$.

{cor:numOfIrreps}

Corollary 1.18. The number of irreducible representations is the dimension of $\mathbb{C}_{\text{class}}$; since there is “one degree of freedom” on each conjugacy class, this number is just the number of conjugacy classes of G .

Remark 1.19. From Corollary 1.18 it is not hard to see that the following statements are equivalent:

- G is Abelian.
- All irreducible representations of G have dimension 1.

In Section ?? we will explore the converse implication: If there are no low dimensional representations, then G is very “non-Abelian” and hence the SRW on its Cayley graph mixes fast.

However, before we can prove this fact, we need two more generalisations of classical results:

{thm:FourierInv}

Theorem 1.20 (Fourier inversion theorem). *For all $s \in G$ we have*

$$f(s) = \frac{1}{|G|} \sum_{\rho} d_{\rho} \operatorname{tr} \left(\rho(s^{-1}) \hat{f}(\rho) \right),$$

where the sum is over all irreducible representations.

This can be seen to reduce to the classical Fourier inversion theorem for $G = \mathbb{Z}/n\mathbb{Z}$; see [Ber16, p.52].

Proof. Both sides are linear, so it suffices to prove it for $f = \delta_t$ with $t \in G$ arbitrary. In that case $\hat{f}(\rho) = \rho(t)$, so the r.h.s. equals

$$\frac{1}{|G|} \sum_{\rho} d_{\rho} \chi_{\rho}(s^{-1}t),$$

which is nonzero only if $s = t$, in which case it is equal to 1 by Corollary 1.15. $\square$

Lemma 1.21. *Let $f \in \mathbb{C}_{\text{class}}$ and let ρ be an irreducible representation. Then $\hat{f}(\rho) = \lambda I$, where $\lambda = \frac{|G|}{d_{\rho}}(f|\bar{\chi}_{\rho})$.*

Proof. Fix any $s \in G$. We would like to apply Schur's lemma 1.8 to $\hat{f}$ so consider the following:

$$\begin{aligned} \rho(s) \hat{f}(\rho) \rho(s)^{-1} &= \sum_{t \in G} f(t) \rho(s) \rho(t) \rho(s^{-1}) \\ &= \sum_{t \in G} f(t) \rho(sts^{-1}) \\ &= \sum_{t \in G} f(sts^{-1}) \rho(sts^{-1}) = \hat{f}(\rho), \end{aligned}$$

since f is a class function. Now by Schur's lemma, $\hat{f}(\rho) = \lambda I$ for some $\lambda \in \mathbb{C}$. Taking traces we get

$$\begin{aligned} \lambda d_{\rho} &= \operatorname{tr} \left(\hat{f}(\rho) \right) \\ &= \sum_{s \in G} f(s) \operatorname{tr} (\rho(s)) \\ &= \sum_{s \in G} f(s) \chi_{\rho}(s) = |G|(f|\bar{\chi}_{\rho}). \end{aligned}$$

$\square$

We are now ready to prove our main result:

Proof of Theorem 1.17. We already know that characters are class functions themselves and that they are orthonormal. It remains to show that if $(f|\chi_\rho) = 0$ for all irreducible representations ρ , then $f = 0$. By the above lemma, however, $(f|\chi_\rho) = 0$ for all ρ implies that $\hat{f} = 0$ and by the Fourier inversion theorem 1.20 this implies that $f = 0$. $\square$

For the sake of completeness we note that it is easy to prove that for arbitrary functions $f, g : G \rightarrow \mathbb{C}$ and the convolution $(f * g)(s) := \sum_{t \in G} f(st^{-1})g(t)$ we have

$$\widehat{f * g}(\rho) = \hat{f}(\rho)\hat{g}(\rho),$$

as well as a generalisation of the Plancherel formula:

Theorem 1.22 (Plancherel formula). *Let $f, g : G \rightarrow \mathbb{C}$ be two functions. Then*

$$\sum_{s \in G} f(s)g(s^{-1}) = \frac{1}{|G|} \sum_{\rho} d_{\rho} \operatorname{tr} \left(\hat{f}(\rho)\hat{g}(\rho) \right).$$

Proof. By linearity it suffices to consider $f = \delta_t$ for which the proof reduces to showing $g(t^{-1}) = \frac{1}{|G|} \sum_{\rho} d_{\rho} \operatorname{tr} (\rho(t)\hat{g}(\rho))$, which is precisely the Fourier inversion theorem 1.20. $\square$

This theorem is very important in the analysis of the convergence of group-valued random walks to their equilibrium since the n -step probability is just an n -fold convolution of the transition kernel.

References

- [AK13] Gil Alon and Gady Kozma. The probability of long cycles in interchange processes. *Duke Mathematical Journal*, 162(9):1567–1585, 2013.
- [Ber16] Nathanaël Berestycki. Lecture notes: Mixing times of markov chains. <http://www.statslab.cam.ac.uk/~beresty/teach/Mixing/mixing3.pdf>, 2016. [Online; accessed 20-March-2019].
- [Dia88] Persi Diaconis. Group representations in probability and statistics. *Lecture notes-monograph series*, 11:i–192, 1988.

- [FH13] William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013.
- [JK81] Gordon James and Adalbert Kerber. *The Representation Theory of the Symmetric Group*, volume 16. Encyclopedia Math. Appl, 1981.
- [OV96] Andrei Okounkov and Anatoly Vershik. A new approach to representation theory of symmetric groups. *Selecta Mathematica, New Series*, 2(4):581–605, 1996.
- [Sag13] Bruce E Sagan. *The symmetric group: representations, combinatorial algorithms, and symmetric functions*, volume 203. Springer Science & Business Media, 2013.
- [Ser77] Jean-Pierre Serre. *Linear representations of finite groups*. Springer, 1977.