Introduction to Representation Theory, Character Theory, and Applications to Random Walks

William Chuang

February 20, 2025

Contents

1 Introduction

These notes introduce some fundamental ideas in the representation theory of finite groups and illustrate how character theory can be used to study random walks on groups. We will cover:

- The definition of a representation and basic facts about group algebras.
- Projection operators onto invariant subspaces.
- The Fourier transform (or discrete Fourier analysis) on a finite group.

- Character theory, with a focus on the symmetric group S_n.
- Applications of representation theory to random walks, mixing times, and bounding convergence to the uniform distribution.

Our aim is to present this material in a way that is accessible to an undergraduate student with some background in linear algebra and group theory.

2 Representations of Finite Groups

2.1 Basic Definitions

Definition 2.1 (Representation)**.** Let G be a finite group. A *representation* of G over a field $\mathbb{C}$ is a group homomorphism

$$\pi : G \to \mathrm{GL}(V),$$

where V is a finite-dimensional vector space over $\mathbb{C}$, and $\mathrm{GL}(V)$ denotes the group of all invertible linear transformations on V. The dimension of V is called the *dimension* of the representation.

Definition 2.2 (Subrepresentation)**.** If $\pi : G \to \mathrm{GL}(V)$ is a representation and $W \subseteq V$ is a subspace such that $\pi(g)(W) \subseteq W$ for all $g \in G$, then W is called a *G-invariant subspace* of V, and the restriction of π to W is called a *subrepresentation*.

Definition 2.3 (Irreducible Representation)**.** A nonzero representation $\pi : G \to \mathrm{GL}(V)$ is *irreducible* if the only G-invariant subspaces of V are $\{0\}$ and V itself. Equivalently, π is irreducible if it has no nontrivial subrepresentations.

Theorem 2.4 (Maschke's Theorem)**.** *If G is a finite group and $\mathrm{char}(\mathbb{C}) = 0$ (or more generally if $|G|$ is invertible in the field), then every finite-dimensional representation of G over $\mathbb{C}$ is completely reducible. That is, any representation can be decomposed as a direct sum of irreducible representations.*

2.2 Group Algebra and the Regular Representation

Definition 2.5 (Group Algebra)**.** Let G be a finite group. The *group algebra* $\mathbb{C}[G]$ is the vector space over $\mathbb{C}$ with basis $\{g \mid g \in G\}$ and multiplication linearly extended from the group multiplication. An element of $\mathbb{C}[G]$ looks like

$$\sum_{g \in G} \alpha_g \, g, \quad \alpha_g \in \mathbb{C}.$$

Definition 2.6 (Regular Representation)**.** The group G acts on $\mathbb{C}[G]$ by left multiplication:

$$h \cdot \left(\sum_{g \in G} \alpha_g g \right) = \sum_{g \in G} \alpha_g \, (hg).$$

This action defines a representation, called the *left regular representation*.

A key fact from Maschke's Theorem is that the regular representation decomposes as

$$\mathbb{C}[G] \cong \bigoplus_{\rho \in \widehat{G}} (\rho \otimes \mathbb{C}^{\dim(\rho)}),$$

where $\widehat{G}$ is the set of (inequivalent) irreducible representations of G, and each ρ appears with multiplicity $\dim(\rho)$.

3 Projection Operators and Invariant Subspaces

A crucial tool in representation theory is the projection operator onto the subspace of G-invariant vectors.

Definition 3.1 (Invariant Subspace)**.** If $\pi : G \to \mathrm{GL}(V)$ is a representation, the subspace of *invariant vectors* is

$$V^G \;=\; \{v \in V \mid \pi(g)(v) = v \text{ for all } g \in G\}.$$

Proposition 3.2 (Projection Onto Invariants)**.** *Let $\pi : G \to \mathrm{GL}(V)$ be a representation of a finite group G. Define the operator*

$$P \;=\; \frac{1}{|G|}\sum_{g\in G}\pi(g).$$

Then P is a projection onto V^G. In other words,

$$P^2 = P, \quad \mathrm{Im}(P) = V^G.$$

Proof. First note that P commutes with every $\pi(h)$:

$$\pi(h)P \;=\; \frac{1}{|G|}\sum_{g\in G}\pi(h)\pi(g) \;=\; \frac{1}{|G|}\sum_{g\in G}\pi(hg) \;=\; \frac{1}{|G|}\sum_{g'\in G}\pi(g') \;=\; P,$$

where we substituted $g' = hg$. Hence P is in the commutant of $\{\pi(g)\}_{g\in G}$.

Next,

$$P^2 = \frac{1}{|G|^2}\sum_{g,h\in G}\pi(g)\pi(h) = \frac{1}{|G|^2}\sum_{g,h\in G}\pi(gh) = \frac{1}{|G|}\sum_{t\in G}\pi(t) = P,$$

so P is idempotent.

To see that $\mathrm{Im}(P) = V^G$, observe that for $v \in V^G$, we have $\pi(g)(v) = v$ for all g, thus

$$P(v) = \frac{1}{|G|}\sum_{g\in G}\pi(g)(v) = \frac{1}{|G|}\sum_{g\in G}v = v.$$

Hence $v \in \mathrm{Im}(P)$, so $V^G \subseteq \mathrm{Im}(P)$. Conversely, if $v \in \mathrm{Im}(P)$, then $v = P(w)$ for some w, and v is G-invariant because P commutes with all $\pi(g)$. Thus $\mathrm{Im}(P) \subseteq V^G$. $\square$

4 Fourier Transform on a Finite Group

For an *abelian* group G, the classical discrete Fourier transform diagonalizes convolution operators in $\mathbb{C}[G]$. More generally, if G is non-abelian, we can still decompose the group algebra into irreducible blocks. This is sometimes called the *Fourier transform on G*.

Proposition 4.1. *Let G be a finite group, and let $\{\rho^i\}_{i=1}^r$ be the irreducible representations of G, each of dimension d_i. Then we have a decomposition*

$$\mathbb{C}[G] \;\cong\; \bigoplus_{i=1}^{r}\mathrm{Mat}_{d_i}(\mathbb{C}),$$

where an element of $\mathbb{C}[G]$ acts by left multiplication on each block, and the isomorphism is given by collecting matrix coefficients of the irreps. Under this isomorphism, a convolution operator (by a function $f \in \mathbb{C}[G]$) becomes block diagonal in the irreducible basis, by Schur's Lemma.

Remark 4.2. For abelian groups, all irreps have dimension 1, so the decomposition is simply a direct sum of 1-dimensional spaces, which recovers the classical discrete Fourier transform.

5 Character Theory and S_n

5.1 Characters

Definition 5.1 (Character)**.** If $\pi : G \to \mathrm{GL}(V)$ is a representation, its *character* χ_π is the function $\chi_\pi : G \to \mathbb{C}$ given by

$$\chi_\pi(g) = \mathrm{Trace}\big(\pi(g)\big).$$

Definition 5.2 (Irreducible Characters)**.** If π is an irreducible representation, then χ_π is called an *irreducible character*. Characters of irreducible representations are fundamental in the study of representations.

Key properties of characters include:

- Characters are class functions: $\chi_\pi(hgh^{-1}) = \chi_\pi(g)$ for all $g, h \in G$.
- Orthogonality relations: For irreducible characters χ_i, χ_j of G,

$$\frac{1}{|G|} \sum_{g \in G} \chi_i(g)\overline{\chi_j(g)} = \delta_{ij}.$$

5.2 The Symmetric Group S_n

The irreducible representations of the symmetric group S_n are classified by *partitions* of n. If $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ is a partition of n, there is a corresponding irreducible representation V^λ often constructed via the *Young diagram* and the *Specht modules* approach. The character χ^λ of this irreducible representation can be computed by combinatorial methods (e.g. the *Murnaghan–Nakayama rule*).

Example 5.3 (Character Table for S_3)**.** The group S_3 has three conjugacy classes:

$$C_1 = \{e\}, \quad C_2 = \{\text{transpositions}\}, \quad C_3 = \{\text{3-cycles}\}.$$

The irreps are:

(1) The trivial representation χ_{triv}, (2) The sign representation χ_{sgn}, (3) The 2-dimensional standard representation χ_{std}.

Their character table is:

	e	(12)	(123)
χ_{triv}	1	1	1
χ_{sgn}	1	-1	1
χ_{std}	2	0	-1

For larger n, one organizes the irreps of S_n via partitions $\lambda \vdash n$. The character values can be computed using various combinatorial formulas (e.g. the Frobenius formula, the Murnaghan–Nakayama rule, etc.).

6 Random Walks on Groups and Mixing Times

6.1 Random Walk Setup

A *random walk* on a finite group G is determined by a probability measure p on G. Starting at the identity, one chooses a group element g with probability $p(g)$ and moves there; repeating this process yields a Markov chain whose states are elements of G. We often study the distribution

$$p^{(k)} = p * p * \cdots * p \quad (k \text{ times}),$$

where $*$ denotes the *convolution* in $\mathbb{C}[G]$. We want to know how quickly $p^{(k)}$ converges to the uniform distribution $u(g) = \frac{1}{|G|}$.

6.2 Representation-Theoretic Approach to Mixing

One powerful method to analyze mixing is to expand p in terms of irreducible characters. For each irreducible representation ρ^i with character χ_i, one can write

$$p(g) = \frac{1}{|G|} \sum_i d_i \langle \chi_i, p \rangle \chi_i(g),$$

where $d_i = \dim(\rho^i)$ and $\langle \chi_i, p \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_i(g) \overline{p(g)}$. Then

$$p^{(k)}(g) = (p * p^{(k-1)})(g)$$

and repeated convolution has a spectral interpretation: each irreducible subrepresentation may contract at a rate given by the eigenvalue associated with that representation.

Remark 6.1. In many random walk analyses, the second-largest eigenvalue in magnitude of the transition operator (acting on an appropriate subspace) determines the mixing rate. Representation theory provides a systematic way to identify these eigenvalues for many symmetric measures p.

6.3 Example: Random Transposition Shuffle on S_n

Consider the random walk on S_n given by picking a random transposition $(i\ j)$ (with uniform probability among all $\binom{n}{2}$ transpositions) and multiplying the current permutation by that transposition. This is a classic random walk used to model card shuffling.

Theorem 6.2 (Mixing Time of Random Transposition Shuffle)**.** *For the random transposition shuffle on S_n, the total variation distance from the uniform distribution u satisfies*

$$\left\| p^{(k)} - u \right\|_{\mathrm{TV}} \leq C\, e^{-c\,k/n}$$

for some positive constants C and c. More refined analysis shows that the cutoff *(sharp transition to near-uniform) occurs around $\frac{1}{2} n \log n$ steps.*

Sketch of proof. One uses the representation theory of S_n to show that all non-trivial irreducible representations (except the trivial one) have characters that lead to eigenvalues strictly less than 1. In fact, the second-largest eigenvalue in magnitude is on the order of $1 - O(1/n)$, giving an exponential convergence rate in k/n. A more delicate analysis locates the exact cutoff around $\frac{1}{2} n \log n$. □

Cutoff Phenomenon for the Random Transposition Shuffle on S_n

We analyze the mixing time of the random transposition shuffle on S_n, which consists of repeatedly selecting a random transposition and multiplying the current permutation by it. Our goal is to show that this Markov chain exhibits a cutoff at time $\frac{1}{2}n\log n$.

Proof. We analyze the mixing time using spectral decomposition via representation theory.

Step 1: Transition Matrix and Group Algebra The transition matrix p satisfies:

$$p(\sigma,\tau) = \begin{cases} \frac{2}{n(n-1)}, & \text{if } \tau = \sigma(i\ j) \text{ for some transposition } (i\ j), \\ 1 - \frac{2}{n(n-1)}, & \text{if } \tau = \sigma, \\ 0, & \text{otherwise.} \end{cases}$$

Since transpositions generate S_n, the Markov chain defined by p is irreducible and aperiodic.

Step 2: Spectral Decomposition via Representations Consider the group algebra $\mathbb{C}[S_n]$, which decomposes as:

$$\mathbb{C}[S_n] \cong \bigoplus_{\lambda \vdash n} V^\lambda \otimes V^\lambda.$$

The transition operator p acts as a convolution operator on $\mathbb{C}[S_n]$, and its eigenvalues correspond to irreducible representations.

Step 3: Eigenvalues of p The eigenvalues of p are indexed by partitions $\lambda \vdash n$. The trivial representation $\lambda = (n)$ corresponds to the uniform stationary distribution with eigenvalue $\lambda_1 = 1$. The second-largest eigenvalue corresponds to the standard representation $\lambda = (n-1,1)$, given by:

$$\lambda_2 = 1 - \frac{2}{n-1}.$$

More generally, the eigenvalues associated with the irreducible representation $\lambda = (\lambda_1, \lambda_2, \dots)$ are:

$$\lambda_\lambda = 1 - \frac{\sum_i \lambda_i(\lambda_i - 1)}{n(n-1)}.$$

Since $\lambda_2 \approx 1 - O(1/n)$, it determines the mixing rate.

Step 4: Bounding the Total Variation Distance The total variation distance satisfies:

$$\left\| p^k - u \right\|_{\text{TV}} \le \sum_{\lambda \neq (n)} d_\lambda e^{-k(1-\lambda_\lambda)},$$

where d_λ is the dimension of the irreducible representation indexed by λ. The leading term is determined by the second-largest eigenvalue:

$$\left\| p^k - u \right\|_{\text{TV}} \approx e^{-ck/n}.$$

Setting this to be small (e.g., $O(1)$), we obtain the cutoff time:

$$k \approx \frac{1}{2} n \log n.$$

□

We have shown that the random transposition shuffle exhibits a cutoff phenomenon at $\frac{1}{2}n \log n$. The proof relies on decomposing the transition operator using representation theory and character theory, identifying the second-largest eigenvalue, and bounding the total variation distance using eigenvalue decay.

7 Bounding Norms via Markov and Chebyshev Inequalities

To make rigorous statements about convergence, one often uses inequalities relating the random walk's deviation from the mean or from uniform. Let f be a function on G with expectation

$$\mathbb{E}_p[f] = \sum_{g \in G} p(g)\, f(g).$$

Then

$$\mathrm{Var}_p(f) = \mathbb{E}_p[f^2] - (\mathbb{E}_p[f])^2.$$

Classical inequalities such as *Markov's inequality* and *Chebyshev's inequality* can give bounds on probabilities of large deviations, which in turn relate to the total variation distance from uniform if f is chosen suitably (e.g. an indicator function of a certain set).

- **Markov's inequality:** If $X \geq 0$ is a random variable, then

$$\Pr[X \geq a] \;\leq\; \frac{\mathbb{E}[X]}{a}.$$

- **Chebyshev's inequality:** If X is a random variable with mean μ and variance σ^2, then

$$\Pr[|X - \mu| \geq t] \;\leq\; \frac{\sigma^2}{t^2}.$$

These bounds are often used in conjunction with spectral estimates (e.g. bounding the variance of certain class functions under the distribution $p^{(k)}$) to show that the random walk is close to uniform after a certain number of steps.

8 Other Examples: Hypercube Random Walk

Another classical example is the random walk on the n-dimensional hypercube $G = (\mathbb{Z}/2\mathbb{Z})^n$, where each step flips one coordinate chosen uniformly at random. The irreducible representations of $(\mathbb{Z}/2\mathbb{Z})^n$ are all 1-dimensional (since the group is abelian), and one can compute eigenvalues easily. It follows that the walk has a mixing time on the order of $n \log n$ (in fact, for the simple random walk that flips each bit with probability $1/2$, the cutoff is around $n \log n$). Representation theory (in this abelian setting, classical Fourier analysis on $\mathbb{Z}/2\mathbb{Z}$) underlies this computation.

9 Concluding Remarks

We have seen how representation theory and character theory provide powerful tools for:

- Decomposing the group algebra into irreducible blocks,
- Constructing projection operators onto invariant subspaces,

- Diagonalizing or block-diagonalizing convolution operators, and
- Analyzing the convergence rates of random walks on groups.

These methods are broadly applicable, from understanding the structure of finite groups themselves to analyzing algorithms (e.g. card shuffling, random sampling) in combinatorics and computer science.

References

[1] J.-P. Serre, *Linear Representations of Finite Groups*, Springer, 1977.

[2] P. Diaconis, *Group Representations in Probability and Statistics*, IMS, 1988.

[3] P. Diaconis and M. Shahshahani, *Generating a random permutation with random transpositions*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 57(2), 159–179, 1981.

[4] B. E. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*, 2nd ed., Springer, 2001.