Notes on Probability and Representation Theory of Finite Groups

William Chuang

February 20, 2025

Contents

1 Probability on Finite Groups and Total Variation Distance

1.1 Basic Definitions

Let G be a finite set (often a finite group in our applications). A *probability distribution* (or *probability measure*) on G is a function

$$P\colon G \to [0,1]$$

such that $\sum_{g\in G} P(g) = 1$.

Definition 1.1 (Total Variation Distance)**.** *For two probability distributions P and Q on G, the* total variation distance *between them is defined as*

$$\|P - Q\|_{\mathrm{TV}} \;=\; \max_{A\subseteq G}\bigl|P(A) - Q(A)\bigr|.$$

Equivalently, it is well-known that

$$\|P - Q\|_{\mathrm{TV}} \;=\; \frac{1}{2}\sum_{g\in G}\bigl|P(g) - Q(g)\bigr| \;=\; \frac{1}{2}\|P - Q\|_1,$$

where $\|\cdot\|_1$ denotes the ℓ^1-norm.

Proof of equivalence. First, note that

$$\max_{A\subseteq G}\big|P(A)-Q(A)\big| \;=\; \max_{A\subseteq G}\Bigg|\sum_{g\in A}\big(P(g)-Q(g)\big)\Bigg|.$$

One can choose A to be the set of points g for which $P(g)-Q(g)\geq 0$. Then

$$\max_{A\subseteq G}|P(A)-Q(A)| \;=\; \frac{1}{2}\sum_{g\in G}|P(g)-Q(g)|.$$

Hence the two definitions match. □

1.2 Bounding Lemmas and Remarks

A common task is to bound $\|P-Q\|_{\mathrm{TV}}$ for certain special cases. For instance, if P and Q are obtained by running a random walk on a group G for k steps, one often seeks an upper bound on $\|P-Q\|_{\mathrm{TV}}$ in terms of k and properties of G or of the step distribution.

Lemma 1.2 (Upper Bound Lemma, Diaconis-style)**.** *Suppose P and Q are probability measures on G. In many scenarios, one has an upper bound on $\|P-Q\|_{\mathrm{TV}}$ by exploiting symmetry or Fourier techniques (discussed below). In particular, if P and Q arise from repeated convolution of an initial measure, character bounds can give rates of convergence to the uniform distribution.*

Remark 1.3. *The idea is that for a finite group G, one can write the difference $P-Q$ in terms of the irreducible characters of G. Each step of a random walk (a convolution by some driving measure) dampens all but the trivial character. Estimating that damping gives explicit upper bounds on $\|P-Q\|_{\mathrm{TV}}$.*

1.3 Example: A Special Case on a Cyclic Group

Let $G=\mathbb{Z}/n\mathbb{Z}$ be the cyclic group of order n. If μ is a probability measure on G with some support that generates the whole group (e.g., $\mu(1)=p$, $\mu(0)=1-p$, etc.), then repeated convolution μ^{*k} tends to the uniform distribution $u=(1/n,\ldots,1/n)$ as $k\to\infty$. The total variation distance $\|\mu^{*k}-u\|_{\mathrm{TV}}$ can often be bounded using discrete Fourier analysis, leading to explicit mixing times.

2 Fourier Analysis on Finite Groups

We now review some basics of the Fourier transform on finite groups, which is a key tool in bounding total variation distances of random walks and in many other contexts.

2.1 Group Algebras and Irreducible Representations

Let G be a finite group of order $|G|$. Consider the complex vector space $\mathbb{C}[G]$, whose elements are formal linear combinations of elements of G. Often, we identify $\mathbb{C}[G]$ with the space of complex-valued functions on G, denoted $L^2(G)$ (with dimension $|G|$). The inner product on $L^2(G)$ is given by

$$\langle f,h\rangle \;=\; \frac{1}{|G|}\sum_{g\in G} f(g)\,\overline{h(g)}.$$

A *representation* of G on a complex vector space V is a group homomorphism $\rho\colon G \to GL(V)$. A representation is called *irreducible* if V has no nontrivial proper subrepresentation. Every finite group has only finitely many irreducible representations up to isomorphism, say

$$\rho_1, \rho_2, \ldots, \rho_r,$$

with dimensions $d_1, d_2, \ldots, d_r$, respectively. We have the fundamental fact (the *orthogonality relations*) that

$$\sum_{g \in G} \chi_i(g)\,\overline{\chi_j(g)} \;=\; |G|\,\delta_{ij},$$

where $\chi_i(g) = \mathrm{trace}(\rho_i(g))$ is the *character* of the representation ρ_i.

2.2 Fourier Transform on a Finite Group

Definition 2.1 (Fourier Transform). *For $f \in L^2(G)$, its* Fourier transform *is the tuple (of matrices) given by*

$$\widehat{f}(\rho_i) \;=\; \sum_{g \in G} f(g)\,\rho_i(g), \quad \textit{for each } i = 1, 2, \ldots, r.$$

Each $\widehat{f}(\rho_i)$ is a $d_i \times d_i$ matrix. Collectively, the family $\{\widehat{f}(\rho_i)\}$ encodes the frequencies of f along each irreducible representation.

Theorem 2.2 (Plancherel's Theorem for Finite Groups). *The map $f \mapsto \{\widehat{f}(\rho_i)\}$ is an isometric isomorphism from $L^2(G)$ onto the direct sum of the matrix spaces corresponding to the irreducible representations of G. Concretely,*

$$\|f\|^2_{L^2(G)} \;=\; \frac{1}{|G|}\sum_{g \in G} |f(g)|^2 \;=\; \frac{1}{|G|}\sum_{i=1}^{r} d_i\,\|\widehat{f}(\rho_i)\|^2_{HS},$$

where $\|\cdot\|_{HS}$ is the Hilbert–Schmidt norm on matrices.

Sketch of Proof. See, e.g., Serre's *Linear Representations of Finite Groups* or any standard text on representation theory of finite groups. The proof follows from the orthogonality relations of characters and the fact that $L^2(G)$ decomposes into the direct sum of all irreducible representations, each occurring with multiplicity equal to its dimension. □

2.3 Fast Fourier Transform Techniques

For an *abelian* finite group G, all irreducible representations have dimension 1, so the Fourier transform reduces to taking discrete characters. In particular, for $G \cong \mathbb{Z}/n\mathbb{Z}$, the Fourier transform is exactly the *discrete Fourier transform* (DFT) of length n. Algorithms like the Fast Fourier Transform (FFT) compute this in $O(n \log n)$ time rather than the naive $O(n^2)$.

For certain nonabelian groups (e.g., some metabelian groups, S_n, etc.), there are analogs of "fast" transforms but they may be more involved. The idea is to exploit the group structure and the known block decomposition of the group algebra.

3 Character Theory of Some Specific Groups

3.1 Example: The Symmetric Group S_4

The group S_4 (the permutations of 4 elements) has 5 conjugacy classes, typically labeled by cycle type:

Cycle type	(1)(2)(3)(4)	(12)	(12)(34)	(123)	(1234)
Class name	$1A$	$2A$	2^2	$3A$	$4A$
Size of class	1	6	3	8	6

Correspondingly, there are 5 irreducible representations of S_4: the trivial representation, the sign representation, the standard 3-dimensional representation, and two others. One can list their characters in a 5×5 table, known as the *character table* of S_4. (Sometimes the notation for classes differs, e.g. $1A, 2A, 2B, 3A, 4A$, etc., but the concept is the same.)

3.2 Example: The Group $\mathrm{SL}_2(\mathbb{Z}_3)$

The group $\mathrm{SL}_2(\mathbb{Z}_3)$ consists of all 2×2 matrices with entries in the finite field $\mathbb{Z}_3$ and determinant 1. It is a nonabelian group of order 24. One can study its irreps either by direct construction or by exploiting known isomorphisms (e.g. $\mathrm{SL}_2(\mathbb{Z}_3)$ is isomorphic to the binary tetrahedral group, though that may be more advanced).

A classical fact is that $\mathrm{SL}_2(\mathbb{Z}_3)$ is *not* a direct product of smaller groups. One can see this from its character table or from the fact that it is a perfect group of small order, etc.

Remark 3.1. *Sometimes* $\mathrm{SL}_2(\mathbb{Z}_3)$ *is related to* A_4 *(the alternating group on 4 elements) via a double cover or a projective representation, but these details go beyond a simple example. The main point is that it has interesting representations of dimensions 1, 2, 3, etc., and they can be understood by group-theoretic and character-theoretic methods.*

4 Connections and Concluding Remarks

4.1 Mixing of Random Walks

A major application of Fourier analysis on finite groups is bounding the convergence of a random walk to its stationary distribution. In the case of a group of order $|G|$, if we convolve an initial distribution with a probability measure μ on G (assuming μ is a *generating* measure or has some spectral gap), one often shows that:

$$\|\mu^{*k} - u\|_{\mathrm{TV}} \;\leq\; \max_{\rho \neq \text{trivial}} \big\|\rho(\mu)\big\|^k,$$

where $\|\rho(\mu)\|$ is an operator norm (or something analogous) that measures how far the representation ρ is from annihilating μ. Since the trivial representation always has eigenvalue 1, all other irreps typically have eigenvalues strictly less than 1 in absolute value (under suitable assumptions), so this distance decays exponentially in k.

4.2 Summary

These notes touched on:

- **Total variation distance** and its basic properties.
- **Fourier transform on finite groups**, including:
  - Irreducible representations and characters,
  - Plancherel's theorem,
  - Fast Fourier Transform for abelian (and some nonabelian) groups.
- **Examples** like cyclic groups, S_4, and $\mathrm{SL}_2(\mathbb{Z}_3)$.

In more advanced treatments, one uses these tools to derive mixing rates for random walks, build fast algorithms for group-theoretic problems, and study the representation theory of more complicated groups.

References

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

[2] P. Diaconis, *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics, 1988.

[3] W. Fulton and J. Harris, *Representation Theory: A First Course*, Springer, 1991.