RANDOM WALKS ON FINITE GROUPS

YUERONG ZHUANG

Abstract. This paper gives an overview of random walks and Fourier analysis on finite groups. First, random walks are discussed on $\mathbb{Z}^d$, focusing on the question of whether a walker returns to the origin walking randomly in finite time in d different dimensions. Then, we recall the Fourier transform on $\mathbb{R}$ and construct a corresponding theory for finite groups. Representation theory and Fourier analysis simplify the computations needed in the convolution operations that arise in multi-step random walks. Finally, we give an example of random transpositions in the symmetric group, showing that it takes $\frac{1}{2}k \log k$ steps to randomly mix k cards.

Contents

1. Introduction

Random walks are the process in which an object takes random paths in a space X from an origin. It is essential in various fields, such as recording the Brownian motion of particles or tracing stock prices [1]. Fourier transforms are also useful mathematical tools because they decompose complex functions into their fundamental building blocks and speed up information processing.

In this paper, Section 2 introduces some basic properties of random walks on integer lattices $\mathbb{Z}^d$ and studies the problem of returning to the origin. Section 3 moves to the Fourier transform on $\mathbb{R}$ first and then on finite groups, further including the idea of convolutions and representation theory. The card shuffling example in Section 4 demonstrates an application of random walks and Fourier analysis on the symmetric group S_n.

2. Random Walks in $\mathbb{Z}^d$

First, we formally define random walks.

Date: August 14th, 2020.

Definition 2.1. Let $\{X_k\}_{k=1}^{\infty}$ be a sequence of independent and identically distributed discrete random variables. For all $n \geq 1$, let $\Sigma_n = X_1 + X_2 + \cdots + X_n$. The sequence of partial sums $\{\Sigma_n\}_{n=1}^{\infty}$ is called a *random walk.*

We first recall classical results for random walks in $\mathbb{Z}^d$ and address the question of if a walker will go back to the origin in finite time. We now directly compute the probability of returning to the origin in $\mathbb{Z}$, following [1].

Lemma 2.2. *In $\mathbb{Z}$, the probability of a return to the origin at time t is*

$$\Pr(\Sigma_t = 0) = \begin{cases} \binom{2m}{m} 2^{-2m}, & t = 2m \\ 0, & t = 2m-1 \end{cases},$$

for $m \geq 1$.

Proof. In a 1-dimensional lattice, the walker moves 1 unit to either the positive or the negative direction: $\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$ for $1 \leq i \leq t$. When t is odd, the movements in the two directions can not cancel out. When $t = 2m$, the walker needs to choose *positive* in m out of $2m$ steps. There are $\binom{2m}{m}$ number of ways of choosing this, each with the probability of $\frac{1}{2}$. Further, the m *negative* steps are chosen with probability $\frac{1}{2}$ each, giving us the required probability. □

We can further specify a random walk's behavior by considering when it first arrives at the origin.

Lemma 2.3. *Let f_{2k} denote the event that the walker first returns to the origin at the $2k^{th}$ step. Then, for $m \geq 1$, we have*

$$\Pr(\Sigma_{2m} = 0) = \sum_{k=0}^{m} \Pr(f_{2k}) \Pr(\Sigma_{2m-2k} = 0).$$

Proof. Consider the collection of all the cases with $\Sigma_{2m} = 0$. We divide it into classes according to when the first return occurs. The probability of choosing the case that the first return happens at the $2k^{th}$ step is

$$\Pr(f_{2k}) 2^{2k} \Pr(\Sigma_{2m-2k} = 0).$$

The total probability of returning at the $2m^{th}$ step equals the sum of the probabilities in each class, so

$$\Pr(\Sigma_{2m} = 0) = \sum_{k=0}^{k=m} \Pr(f_{2k}) \Pr(\Sigma_{2m-2k} = 0),$$

which immediately proves the result. □

Lemma 2.4. *For $m \geq 1$,*

$$\Pr(f_{2m}) = \frac{\Pr(\Sigma_{2m} = 0)}{2m-1}.$$

For this proof, we need to introduce the concept of generating functions. Namely, given a countable set of values $A = \{a_1, a_2, \ldots\}$, we define the function

$$A(x) = P_A(x) := \sum_{i=1}^{\infty} a_i x^i$$

as the *generating function* corresponding to A.

Proof. We first define generating functions corresponding to the sets

$$S = \{\Pr(\Sigma_{2m} = 0) \,|\, m \in \mathbb{N}\}, \quad F = \{\Pr(f_{2m}) \,|\, m \in \mathbb{N}\} :$$
$$S(x) = \sum_{m=0}^{\infty} Pr(\Sigma_{2m} = 0)x^m, \quad F(x) = \sum_{m=0}^{\infty} \Pr(f_{2m})x^m.$$

By the construction of the functions, we know that they are well-defined with all coefficients in $[0, 1]$. Hence,

$$\begin{aligned} S(x) &= \sum_{m=0}^{\infty} \binom{2m}{m} 2^{-2m} x^m && \text{by Lemma 2.2} \\ &= \sum_{m=0}^{\infty} \binom{2m}{m} \left(\frac{x}{4}\right)^m \\ &= \frac{1}{\sqrt{1-x}} && \text{by Binomial Theorem.} \end{aligned}$$

Lemma 2.3 also gives the relationship between S and F:

$$S(x) = 1 + S(x)F(x).$$

Hence,

$$\begin{aligned} F(x) &= 1 - \frac{1}{S(x)} = 1 - \sqrt{1-x} \\ &= 1 - \sum_{m=1}^{\infty} \binom{\frac{1}{2}}{m} (-1)^{m-1} x^m && \text{by Binomial Theorem} \\ &= \sum_{m=1}^{\infty} \frac{\binom{2m}{m}}{(2m-1)2^{2m}} x^m \\ &= \sum_{m=1}^{\infty} \frac{\Pr(\Sigma_{2m} = 0)}{2m-1} x^m. \end{aligned}$$

□

We are now in a position to prove that a random walker has probability 1 of returning to the origin in finite time.

Theorem 2.5. *In $\mathbb{Z}$, the probability of a random walk returning to the origin in finite time is 1.*

Proof. Define a new generating function on $[0, 1]$ to be $F_M(x) = \sum_{m=0}^{M} \Pr(f_{2m})x^m$. Because we know $\sum_{m=0}^{\infty} \Pr(f_{2m}) \le 1$, the series

$$F(x) = \lim_{M \to \infty} F_M(x)$$

converges for any $x \in [0,1]$. The probability of eventually returning to the origin is:

$$\begin{aligned}
\sum_{m=0}^{\infty} \Pr(f_{2m}) &= \lim_{M\to\infty} \lim_{x\to 1} F_M(x) \\
&= \lim_{x\to 1} \lim_{M\to\infty} F_M(x) && \text{by the convergence of the series} \\
&= \lim_{x\to 1} F(x) && \text{by the definition of } F(x) \\
&= F(1) && \text{by the continuity of } F \text{ and the Weierstrass } M \text{ test} \\
&= 1.
\end{aligned}$$

□

Remark 2.6. If A is a countable set with $\sum_{a\in A} a = 1$, then the corresponding generating function $A(x)$ has a fixed point at $x = 1$: $A(1) = 1$.

Now we may continue the discussion in higher dimensions. Let f^d_{2m} be the event that the walker first returns to the origin on $\mathbb{Z}^d$ at the $(2m)$th step, and let Σ^d_i be the walker's position on $\mathbb{Z}^d$ at the ith step. We can define similar generating functions on $[0,1]$ as

$$S^{(d)}(x) = \sum_{m=0}^{\infty} \Pr(\Sigma^d_{2m} = 0)x^m, \quad F^{(d)}(x) = \sum_{m=0}^{\infty} \Pr(f^d_{2m})x^m.$$

Previous results from 1D lattices generalize well:

$$\Pr(\Sigma^d_{2m} = 0) = \sum_{k=0}^{m} \Pr(f^d_{2k})\Pr(\Sigma^d_{2m-2k} = 0), \quad F^{(d)}(x) = \frac{S^{(d)}(x) - 1}{S^{(d)}(x)}.$$

Let r^d_i be the probability that the walker returns to the origin at least once in the first i steps on $\mathbb{Z}^d$. When the walker ultimately returns to the origin, we have

$$\begin{aligned}
r^d_\infty &= \lim_{i\to\infty} r^d_i \\
&= \lim_{i\to\infty} \sum_{m=0}^{\lceil i/2 \rceil} \Pr(f^d_{2m}) \\
&= \lim_{i\to\infty} \lim_{x\to 1} \sum_{m=0}^{\lceil i/2 \rceil} \Pr(f^d_{2m})x^m \\
&= \lim_{x\to 1} \lim_{i\to\infty} \sum_{m=0}^{\lceil i/2 \rceil} \Pr(f^d_{2m})x^m \\
&= \lim_{x\to 1} F^{(d)}(x) \\
&= \lim_{x\to 1} \frac{S^{(d)}(x) - 1}{S^{(d)}(x)} \\
&= 1 - \lim_{x\to 1} \frac{1}{S^{(d)}(x)}. \qquad (2.7)
\end{aligned}$$

Lemma 2.8.

$$\lim_{x\to 1} S^{(d)}(x) = \sum_{m=0}^{\infty} \Pr(\Sigma_{2m}^{d} = 0).$$

Proof. First, we use the convergence of the series and get

$$\begin{aligned}\sum_{m=0}^{M} \Pr(\Sigma_{2m}^{d} = 0) &= \lim_{x\to 1} \sum_{m=0}^{M} \Pr(\Sigma_{2m}^{d} = 0)x^{m} \\ &\leq \lim_{x\to 1} \sum_{m=0}^{\infty} \Pr(\Sigma_{2m}^{d} = 0)x^{m} \\ &= \lim_{x\to 1} S^{(d)}(x).\end{aligned}$$

To prove the reverse direction, we use the fact that the coefficients in $S^{(d)}$ are positive. The series is monotonic:

$$\lim_{x\to 1} S^{(d)}(x) \leq \sum_{m=0}^{\infty} \Pr(\Sigma_{2m}^{d} = 0).$$

□

We can now give results for returning to the origin in higher dimensions.

Theorem 2.9. *In $\mathbb{Z}^2$, the probability of a random walk returning to the origin in finite time is 1.*

Proof. First, by Stirling's formula,

$$n! \to \sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}, \quad \text{for } n \to \infty.$$

So, since we have

$$\binom{2m}{m} \to \frac{2^{2m}}{\sqrt{\pi m}},$$

we know from Lemma 2.2 that

$$\Pr(\Sigma_{2m}^{2} = 0) = \left(\frac{1}{2^{2m}}\binom{2m}{m}\right)^{2} = \frac{1}{4^{2m}}\binom{2m}{m}^{2} \to \frac{1}{\pi m}.$$

This means that the series $\sum_{m=0}^{\infty} \Pr(\Sigma_{2m}^{2} = 0)$ diverges to infinity. Then, from previous results,

$$\begin{aligned}r_{\infty}^{2} &= 1 - \lim_{x\to 1} \frac{1}{S^{2}(x)} && \text{by 2.7} \\ &= 1 - \frac{1}{\sum_{m=0}^{\infty} \Pr(\Sigma_{2m}^{2} = 0)} && \text{by 2.8} \\ &= 1.\end{aligned}$$

□

Theorem 2.10. *In $\mathbb{Z}^3$, the probability of a random walk returning to the origin is less than 1.*

Proof. As in the proof of Theorem 2.9 , we have:

$$\Pr(\Sigma_{2m}^3 = 0) = \frac{1}{2^{2m}} \binom{2m}{m} \sum_{j,k \geq 0}^{j+k \leq m} \frac{1}{3^m} \frac{m!}{j!k!(m-j-k)!}.$$

Letting

$$K := \max_{j,k} \left\{ \frac{m!}{j!k!(m-j-k)!} \right\},$$

we have that

$$\Pr(\Sigma_{2m}^3 = 0) \leq \frac{1}{2^{2m}} \binom{2m}{m} \sum_{j,k \geq 0}^{j+k \leq m} \frac{K}{3^m} \frac{m!}{j!k!(m-j-k)!}.$$

By Stirling's formula, we have $K \to \frac{C}{m}$ for some constant C. We also know that

$$\sum_{j,k \geq 0}^{j+k \leq m} \frac{1}{3^m} \frac{m!}{j!k!(m-j-k)!} = 1.$$

Thus, by the previous inequality, $\sum_{m=0}^{\infty} \Pr(\Sigma_{2m}^3 = 0)$ can be compared to a $p-$series and thus converges to a finite number. Now,

$$r_\infty^3 = 1 - \frac{1}{\sum_{m=0}^{\infty} \Pr(\Sigma_{2m}^3 = 0)} < 1.$$

□

Corollary 2.11. *The probability of returning to the origin is less then 1 for $\mathbb{Z}^d$ with $d > 3$.*

Proof. We claim that the probability of returning to the origin in $\mathbb{Z}^{d+1}$ is less than that in $\mathbb{Z}^d$ for $d \geq 2$. The statement is true when $d = 2$, and induction is needed for higher dimensions. We now think of Σ_m^d as a vector of $d+1$ components, with the last component being 0. Σ_m^{d+1} can also be thought of as a vector with $d+1$ components. Thus, all the events that the walker returns to the origin in $d+1$ dimensions are included in the events where they return to the origin in d dimensions, which proves the claim. As the probability of returning to the origin in $\mathbb{Z}^3$ is less than 1, the claim proves the corollary. □

3. Fourier Analysis on Finite Groups

When calculating the probability of reaching a certain point n after k steps, we use the idea of convolution because although each step is independent, the final position depends on every random choice. We now define convolutions for functions on finite groups.

Definition 3.1. Let G be a group, and $P, Q : G \to \mathbb{C}$ be arbitrary functions. The convolution $\star$ is defined as

$$P \star Q(g) := \sum_{h \in G} P(h) Q(gh^{-1}).$$

In general, we can repeatedly convolve a function with itself:

$$P^{\star n}(g) = P \star P^{\star(n-1)}(g) = \sum_{h \in G} P(h) P^{\star(n-1)}(gh^{-1}).$$

Remark 3.2. The definition above requires the convolution to be associative. We will assume this property here and in the following proofs.

The convolution operation makes intuitive sense; each term in the convolution operation represents a potential movement taken in a random walk. For example, if we move towards g in 2 steps, then the first movement goes to h, and the second movement cancels h by h^{-1} and then approaches g. However, because of the summation, the practical calculation is hard. We need to use some nicer methods to simplify this process, and this is where the Fourier transform helps. We recall the definitions from classical function spaces to give intuitions for the analogs we will build for finite groups.

Definition 3.3. The L^2 norm of a function $f : X \to \mathbb{C}$ is defined as

$$||f||_2 := \left(\int_X |f|^2\right)^{1/2},$$

assuming the above integral is finite.

Definition 3.4. The *inner product* $\langle \cdot, \cdot \rangle_2$ on $L^2(X)$ is a map from $L^2(X) \times L^2(X)$ to $\mathbb{C}$ given by

$$\langle f, g \rangle_2 := \int_X f\overline{g}.$$

We denote two functions $f, g : X \to \mathbb{C}$ as *equal almost everywhere* if $||f - g||_2 = 0$, written as $f \sim g$. We can now let $L^2(X)$ equal the set of equivalence classes under $\sim$, and use the inner product described above to give $L^2(X)$ a *Hilbert space* structure. Let's now take a closer look at the functions

$$e_k(x) = e^{2\pi ikx}$$

on the interval $[0, 1]$.

Theorem 3.5. *The set $E = \{e_k | k \in \mathbb{Z}\}$ of functions defined above is an orthogonal topological basis of $L^2([0, 1])$.*

Proof. See [2]. □

Remark 3.6. E being an orthogonal topological basis means that it satisfies two properties. First, it satisfies orthogonality: $\langle e_m, e_n \rangle_2 = 0$ for $m \neq n$, and $\langle e_n, e_n \rangle = 1$. Second, it is a basis: any function from $L^2([0, 1])$ can be written as a weighted sum of elements of E.

Now, we can define the Fourier transform on $[0, 1]$ for any f as a function $\widehat{f} : \mathbb{Z} \to \mathbb{C}$ such that:

$$\widehat{f}(k) = \int_0^1 f(x)e^{-2\pi ikx}dx.$$

With some effort, we can generalize to Fourier transforms on $\mathbb{R}$, which are defined as

$$\widehat{f}(y) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ixy}dx, \quad f(x) = \int_{-\infty}^{\infty} \widehat{f}(y)e^{2\pi ixy}dy.$$

Now, we can expand Fourier analysis to finite groups. The following example introduces an important function on abelian groups.

Example 3.7. Define a delta function to be:

$$\delta_g(x) = \begin{cases} 1 & x = g \\ 0 & \text{otherwise} \end{cases}.$$

For every function $f \in L^2(G)$, we have

$$f = \sum_{g \in G} f(g)\delta_g,$$

so the set $\{\delta_g | g \in G\}$ spans $L^2(G)$. Also, for $g \neq h$, $\langle \delta_g, \delta_h \rangle = 0$, meaning the delta functions are orthogonal. As the set is orthogonal with the order $|G|$, it forms a basis of $L^2(G)$ as a complex vector space.

Besides the special case of abelian group, Fourier transforms can be defined for non-abelian groups as well, but the proofs require some extra representation theory.

Definition 3.8. Let V be a finite-dimensional vector space over $\mathbb{C}$. A *representation* (ρ, V) of a group G is a group homomorphism $\rho : G \to GL(V)$. The *dimension* of the representation is denoted by $d_\rho = \dim(V)$.

Definition 3.9. Let (ρ, V) be a group representation. A *subrepresentation* (π, W) is a representation of G where $W \subseteq V$ such that for all $w \in W$ and $g \in G$, $\rho(g)w \in W$, and ρ agrees with π when restricted to W: $\rho(g)|_W = \pi(g)$.

Definition 3.10. Let (ρ, V) be a group representation. It is *irreducible* if its only subrepresentations are (ρ, V) and $(\pi, \{0\})$.

Definition 3.11. Let (ρ, V) and (π, W) be two representations of G. The *direct sum* of the representations is a map $\phi : G \to GL(V \oplus W)$ such that:

$$\phi(g)(v, w) = (\rho(g)(v), \pi(g)(w))$$

for $(v, w) \in V \oplus W$.

To illustrate the above definitions, we may take a look at the non-abelian group S_3.

Example 3.12. The trivial representation is a function sending all group elements to the identity.

$$1 : G \to GL(\mathbb{C}) = \mathbb{C}^\times,$$
$$1(g) = 1.$$

Example 3.13. There exists a map $L : S_3 \to GL(\mathbb{C}^3)$ which represents every element of S_3 as a permutation matrix. The map is called a permutation representation, and one example of it can be

$$L\left((1,3)\right) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

We can check that this representation is reducible. Specifically, the subspace spanned by $v = (1, 1, 1)$ is preserved under the action of S_3. It thus forms an irreducible subrepresentation isomorphic to the trivial representation.

Consider now the vector space $W = \mathbb{C}^3/\langle v \rangle = \mathrm{Span}_{\mathbb{C}}\{(1, -1, 0), (0, 1, -1)\}$. Using the new basis of $\mathbb{C}^3$

$$B = \{v, (1, -1, 0), (0, 1, -1)\},$$

we can rewrite the permutation representation in a block-diagonal matrix. The upper block is a trivial representation, and the bottom block is another representation called the standard representation.

$$\rho((1,3)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \in GL(\mathbb{C}v \oplus W) \cong GL(\mathbb{C}^3).$$

In this example, we see that reducible representations can be built from smaller representations. The following theorem shows that it is true for any representation of any group.

Theorem 3.14 (Maschke's Theorem)**.** *Let (ρ, V) be a representation of a finite group G. Then, we can write (ρ, V) as a direct sum of a finite number of irreducible representations $(\pi_1, W_1), \ldots, (\pi_n, W_n)$.*

Proof. Let G and (ρ, V) be as given. The proof is by induction on the dimension of the representation. When $d_\rho = 1$, (ρ, V) must be irreducible. Assume that the theorem is true for all $d_\rho < k$. Then, when $d_\rho = k$, the proof is done if the representation itself is irreducible. If it is not, there exists some subrepresentation (π, W) with $W \subset V$ and $W \neq \{0\}$. Since the dimension of (π, W) is less than k by assumption, (π, W) can be written as a direct sum of irreducible representations.

Consider now the quotient space V/W. Let the cosets of the space be $\{v+W\}$ and let $g \cdot (v + W) = \rho(g)(v) + W$ for all $g \in G$. Because (π, W) is a subrepresentation of (ρ, V), $\rho(g)$ preserves W. Hence, from ρ, we can define a new representation $(\rho', V/W)$ with a dimension lower than k.

Therefore, we have $V = W \oplus V/W$, with $\dim(W)$ and $\dim(V/W)$ both less than $\dim(V)$, so we can apply the inductive hypothesis to finish the proof. □

Now we see how representations can interact.

Definition 3.15. Let (ρ, V) and (π, W) be two representations of G. A map $L : V \to W$ *intertwines* ρ and π if

$$L\rho(g) = \pi(g)L.$$

Two representations are *equivalent* if there exists an intertwining isomorphism L.

Lemma 3.16. *Let $L : V \to W$ intertwine $(\rho, V), (\pi, W)$. The restriction of ρ to the kernel of L is a subrepresentation of ρ. Similarly, the restriction of π to the image of L gives a subrepresentation of π.*

Proof. See [2]. □

Lemma 3.17 (Schur's Lemma)**.** *Let $(\rho, V), (\pi, W)$ be irreducible representations, and let L be as stated in the previous lemma. Then, L is either the 0 map or an isomorphism.*

Proof. Suppose $L \neq 0$, so $\mathrm{Ker}(L) \neq V$. By the previous lemma, $\mathrm{Ker}(L)$ is a subrepresentation of ρ. Hence, $\mathrm{Ker}(L) = \{0\}$ and L is injective. Similarly, $\mathrm{Im}(L) = W$. L is also surjective and is an isomorphism. □

Lemma 3.18. *Let (ρ, V) be an irreducible representation and $L : V \to V$ intertwine ρ with itself. Then, L is a scalar product of the identity linear transformation*

$$L = xI.$$

Proof. Let x be an eigenvalue of L, and let W be the corresponding eigenspace. We have

$$L(\rho(g)w) = \rho(g)xw = x\rho(g)w, w \in W$$

because L intertwines ρ with itself. Also, since $L(w) = xw$,

$$\rho(g)L(w) = \rho(g)xw = x\rho(g)w.$$

W has a subrepresentation of ρ because W is closed under the action of G. Since ρ is irreducible, $W = \{0\}$ or $W = V$, so $L = xI$. □

Our next goal is to define the Fourier transforms on any finite groups. To do so, we can show that the matrix entries of irreducible representations of a group G form a basis of $L^2(G)$, leading to the definition of the Fourier transforms.

Theorem 3.19. *Let $(\rho, \mathbb{C}^a)$ and $(\pi, \mathbb{C}^b)$ be distinct irreducible representations of a finite group G. Let $\rho_{n,m} : G \to \mathbb{C}$ be a function mapping each element $g \in G$ to the element in the n-th row and m-th column of $\rho(g)$, with $\pi_{i,j}$ defined in a similar way. Then,*

$$\langle \rho_{n,m}, \pi_{i,j} \rangle = \sum_{g \in G} \rho_{n,m}(g)\overline{\pi_{i,j}(g)} = 0.$$

Proof. For a linear map $M : \mathbb{C}^b \to \mathbb{C}^a$, define a function N:

$$N := \sum_{g \in G} \rho(g)M\pi(g^{-1}).$$

For a group element y, because ρ is a homomorphism, we have

$$\rho(y)N = \rho(y)\sum_{g \in G} \rho(g)M\pi(g^{-1}).$$

Using a bijection from G to G defined by left multiplication by y, we have:

$$\begin{aligned} \rho(y)N &= \sum_{yg \in G} \rho(yg)M\pi((yg)^{-1}\,y) \\ &= \left(\sum_{yg \in G} \rho(yg)M\pi\left((yg)^{-1}\right)\right)\pi(y) \\ &= N\pi(y) \end{aligned}$$

This implies that the map N intertwines (ρ, V) and (π, W). N is thus either the 0 map or an isomorphism. Because we assume that the two representations are not equivalent, L is not an isomorphism. As M is an arbitrary map from V to W, let it be the map with 1 in the mth row and the jth column, and 0 else. Therefore,

$$\sum_{g \in G} \rho_{n,m}(g)\overline{\pi_{i,j}(g)} = \langle \rho_{n,m}, \pi_{i,j} \rangle = 0.$$

□

Theorem 3.20. *Let $(\rho, \mathbb{C}^a)$ be an irreducible representation of G. Then,*

$$\langle \rho_{nm}, \rho_{ij} \rangle = \begin{cases} 1, & n = i \quad \textit{and} \quad m = j \\ 0, & \textit{else} \end{cases}.$$

Proof. Use the same definition of M and N as above. N intertwines the representation with itself, and by Lemma 3.18, N is a scalar product of the identity matrix. Because the trace of a matrix is preserved under intertwining, we know that

$$|G|\mathrm{Tr}(M) = \mathrm{Tr}(xI) = x \cdot d_\rho.$$

Again, let M be the map with 1 in the mth row and the jth column, and 0 else. Then, we see that $\mathrm{Tr}(M) \neq 0$ if and only if $m = j$, and in that case, the only non-zero entries of N are the diagonal entries. Thus, the inner product is not 0 if and only if $n = i$ and $m = j$. □

Lemma 3.21 (Plancherel's Formula)**.** *Let P be a function from G to $\mathbb{C}$, let $\widehat{G}$ be the set of irreducible representations of G, and let ρ be a representation of G. Then,*

$$\sum_{\pi \in G} |P(\pi)|^2 = \frac{1}{|G|} \sum_{\rho \in \widehat{G}} d_\rho \mathrm{Tr}(\rho(P)\rho(P)^*),$$

where

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

Corollary 3.22. *Let G be a finite group. Then,*

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

The proof of the lemma and its corollary is in [7]. With all these theorems, we have shown that the matrix entries of irreducible representations of a finite group G are orthogonal and have the same size as G. Hence, the matrix entries form a basis of $L^2(G)$. We can now define the Fourier transform of a finite group.

Definition 3.23. Let $f : G \to \mathbb{C}$ be a function, and let $\rho : G \to GL(V)$ be a representation. We now define

$$\widehat{f} : \mathrm{Hom}(G, GL(V)) \to GL(V),$$
$$\widehat{f}(\rho) := \sum_{g \in G} f(g)\rho(g).$$

Remark 3.24. The quantities $\rho(f)$ and $\widehat{f}(\rho)$ are distinct in principle, but are equal in definition. As they represent the same expression, we will treat the two as interchangable in this paper.

We now compute the Fourier transform of the delta function for concreteness.

Example 3.25.

$$\widehat{\delta}_g(\rho) = \sum_{h \in G} \delta_g(h)\rho(h) = \rho(g)$$

because $\delta_g(h)$ is zero unless $g = h$.

We can now go back to our original goal: to calculate convolutions in a simpler way. The following lemma does the job.

Lemma 3.26 (Convolution-Multiplication Theorem)**.** *Let G be a finite group, let $f, g : G \to \mathbb{C}$, and let ρ be a group representation. Then,*

$$\widehat{f \star g}(\rho) = \widehat{f}(\rho)\widehat{g}(\rho) \tag{3.27}$$

Proof.

$$\begin{aligned}
\widehat{f\star g}(\rho) &= \sum_{k\in G}(f\star g)(k)\rho(k) && \text{by Fourier transform}\\
&= \sum_{k\in G}\left(\sum_{h\in G} f(kh^{-1})g(h)\right)\rho(k) && \text{by convolution}\\
&= \left(\sum_{kh^{-1}\in G} f(kh^{-1})\rho(kh^{-1})\right)\left(\sum_{h\in G} g(h)\rho(h)\right) && \text{by swapping the order of summation}\\
&= \widehat{f}(\rho)\widehat{g}(\rho) && \text{by definition of the Fourier Transform.}
\end{aligned}$$

□

4. Card Shuffling and Random Walks

We now have random walks and Fourier transforms as our mathematical tools, so let's use them in a real-life problem: shuffling cards. Unlike our interest in returning to the origin point in random walks on $\mathbb{Z}^d$, we now want the cards to be as far from the initial state as possible after the shuffle.

Suppose we have n labelled cards to shuffle, using a random transposition model. In this model, one shuffle consists of choosing two cards at random and swapping their positions, and we repeatedly iterate this shuffle k times. The question is: how large should k be to ensure that the cards are randomly arranged?

Definition 4.1. Let G be a group. A *probability distribution* P is a function $P : G \to [0,1]$ such that $\sum_{g\in G} P(g) = 1$.

We can model random transpositions by the probability distribution T on the group S_n, defined as:

$$\begin{aligned}
T(e) &= \frac{1}{n} && \text{if } e \text{ is the identity}\\
T(\pi) &= \frac{2}{n^2} && \text{if } \pi \text{ is a transposition}\\
T(\sigma) &= 0 && \text{otherwise.} \qquad (4.2)
\end{aligned}$$

To formalize the phrase 'randomly arranged', we will compare distributions against the normal distribution U:

$$U(\pi) = \frac{1}{n!} \quad \text{for all } \pi \in S_n, \tag{4.3}$$

Intuitively, as people shuffle more, which we can model as convolving T with itself repeatedly, the cards get more mixed, which should mean $T^{\star k} \to U$ as k gets larger. Actually, Theorem 4.5 shows that after a certain number of transpositions, depending on the number of cards, the probability distribution of the cards will be almost the uniform distribution.

Definition 4.4. Let G be a group, and let P and Q be two probability distributions on G. The *variation distance* is defined as:

$$||P - Q|| := \sum_{g\in G} |P(g) - Q(g)| = 2\sup_{A\subset G} |P(A) - Q(A)|.$$

Theorem 4.5. *Let c be a constant. Then, for $k > \frac{1}{2}n\log n + cn$, we have*

$$||T^{*k} - U|| \leq 6e^{-2c}.$$

Definition 4.6. The *c*haracter χ_ρ of the representation ρ is a function $\chi_\rho : G \to \mathbb{C}$ such that for $g \in G$,

$$\chi_\rho(g) = \mathrm{Tr}(\rho(g)).$$

Note that, due to the properties of trace, equivalent representations have identical characters, and characters are constant on conjugacy classes.

Lemma 4.7. *Let G be a finite group, and let ρ be an irreducible representation of G. Further, let P be a function $P : G \to \mathbb{C}$ that is constant on conjugacy classes. Let P_i be the value of P, let n_i be the cardinality, and let χ_i be the value of χ_ρ on the ith conjugacy class. Then, $\widehat{P}(\rho) = CI$, with*

$$C = \frac{1}{d_\rho}\sum_i P_i n_i \chi_i. \tag{4.8}$$

Proof. Let M_i be the sum of $\rho(g)$ for all g in the ith conjugacy class. We have

$$\widehat{P}(\rho) = \sum_{g\in G} P(g)\rho(g) = \sum_i P_i M_i.$$

By conjugacy, the matrix M_i satisfies $\rho(g)M_i\rho(g^{-1}) = M_i$ for all g. Thus, by Theorem 3.18, $M_i = C_i I$ for some C_i.

To find the value of C_i, we calculate the trace:

$$\mathrm{Tr}(M_i) = n_i\chi_i = C_i d_\rho,$$

which leads to the result. □

Corollary 4.9. *Let T be as defined in 4.2, let ρ be a representation on S_n.Then,*

$$\widehat{T}(\rho) = \left(\frac{1}{n} + \frac{n-1}{n}\frac{\chi_\rho(\sigma)}{d_\rho}\right) I,$$

where σ is any transposition.

Proof. By the definition of T, we know that $T = \frac{1}{n}$ on the conjugacy class $[e]$ of the identity, $\frac{2}{n^2}$ on the conjugacy class $[\sigma]$ of transpositions, and 0 on all other conjugacy classes. We can then plug in the values of d_ρ and χ_ρ to Equation 4.8 to obtain the result. □

Lemma 4.10 (Upper Bound Lemma)**.** *Let G be a finite group, P be a probability distribution, and ρ be any G-representation. Then,*

$$4||P^{*k} - U||^2 \leq \sum_{\rho\neq 1} d_\rho ||\widehat{P}(\rho)||^{2k}. \tag{4.11}$$

Proof.

$$\begin{aligned} 4||P^{*k} - U||^2 &= \Big(\sum_{g\in G} |P^{*k}(g) - U(g)|\Big)^2 && \text{by definition} \\ &\leq |G| \sum_{g\in G} |P^{*k}(g) - U(g)|^2 && \text{by Cauchy-Schwarz inequality} \\ &= \sum_{\rho} d_\rho \mathrm{Tr}(\widehat{P}(\rho)^k (\widehat{P}(\rho)^k)^*) && \text{by Plancherel's Formula} \\ &\leq \sum_{\rho} d_\rho ||\widehat{P}(\rho)||^{2k}. \end{aligned}$$

□

Definition 4.12. A *partition* $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)$ of n is a sequence $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ of positive integers with $n = \lambda_1 + \lambda_2 + \cdots + \lambda_m$.

Remark 4.13. There exists a one-on-one correspondence between each partition of n and an irreducible group representation of S_n. The exact representation can be found in [5], but will not be constructed here considering the length of this paper.

Lemma 4.14. *The character of the irreducible representation of S_n corresponding to the partition λ, evaluated at a transposition g satisfies:*

$$\frac{\chi_\rho(g)}{d_\rho} = \frac{1}{n(n-1)} \sum_{j=1}^{m} (\lambda_j^2 - 2j\lambda_j + \lambda_j).$$

The proof of this lemma is in [4].

Proof of Theorem 4.5. We will only prove that the difference between T^{*k} and U is bounded, but will not show the detailed explanation for its precise value. More explicit calculations can be found in [5].

By its definition, we know that T is invariant under conjugation: $T(g) = T(g^{-1}hg)$. As $\widehat{T}$ is the Fourier transform of T, it satisfies

$$\rho(g^{-1})\widehat{T}(\rho)\rho(g) = \widehat{T}(\rho)$$

for all g. By Theorem 3.18, $\widehat{T}(\rho) = cI$ for some c. We can compute

$$c = \frac{1}{n} + \frac{n-1}{n}\frac{\chi_\rho(g)}{d_\rho}$$

by taking traces of matrices. Plugging this value of c into Equation 4.11, we have

$$4||T^{\star k} - U||^2 \leq \sum_{\rho\neq 1} d_\rho \left(\frac{1}{n} + \frac{n-1}{n}\frac{\chi_\rho(g)}{d_\rho}\right)^{2k}.$$

We then take a partition of n and the corresponding irreducible representations of S_n. The case that the cards are farthest from the uniform distribution is when they are in the original order, which means they preserve the identity map in S_n. This leads us to consider the partition $(n-1, 1)$. By Lemma 4.14, the corresponding term in the previous summation is less than

$$(n-1)^2 \left(1 - \frac{2}{n}\right)^{2k}.$$

In all other possible partitions, the difference can only be less than this value. Therefore, by the fact that $1-x \leq e^{-x}$, the result is that $||T^{\star k}-U||$ is bounded. □

Acknowledgments

I would like to thank my mentor Calder Sheagren for providing essential advice on the content and the structure of the paper. Meanwhile, I would like to thank Pallav Goyal for offering many valuable comments on this paper. I would also like to express my gratitude to Peter May for organizing this REU program and admitting me to it.

References

[1] Derek Johnston An Introduction to Random Walks
https://www.math.uchicago.edu/ may/VIGRE/VIGRE2011/REUPapers/Johnston.pdf

[2] Rohan Dandavati The Fourier Transform on Finite Groups: Theory and Computation
http://math.uchicago.edu/ may/REU2018/REUPapers/Dandavati.pdf

[3] Fourier analysis on finite Abelian groups: some graphical applications
https://iuuk.mff.cuni.cz/ andrew/dwajgApr06.pdf

[4] Persi Diaconis Random Walks on Groups: Characters and Geometry
http://statweb.stanford.edu/ cgates/PERSI/papers/randomwalksongroup.pdf

[5] Persi Diaconis, Mehrdad Shahshahani Generating a Random Permutation with Random Transpositions
https://statistics.stanford.edu/research/generating-random-permutation-random-transpositions

[6] H.Dym, H.P.McKean Fourier Serier and Integrals Academic Press, New York

[7] Terras, A. Fourier Analysis on Finite Groups and Applications. Cambridge University Press.