Bounding Random Walks on Groups

Jingxing Wang

August 16, 2024

1 Introduction

The Mathematics Directed Reading Program (or DRP) pairs a graduate student in the department with an undergraduate to assist in learning a particular topic. We settled on the topic of representation theory. We were interested in random walks done on groups because it provides a theoretical basis for the well-known card shuffling trick. We studied the foundations of representation theory, and proved the main theorem related to random walks. Several simulations were done to compare their results with the theorems that follow from the main theorem. A portion of them is included in this paper. This process began during the fall of 2022 and ended before May, 2023. A good portion of the report was done then.

The summer after May 2023, I join a signal processing lab. After several months, after I've gained a deeper understanding on Fourier transforms, by chance I was going back to the theory of representations and characters and I realize that they've got everything to do with Fourier transforms. The algebraic view-point provides me a perfect angle to gaze at signal processing. Thus, the rest of the report was done then, in the spring of 2024.

The reference text: [3].

2 Theory and Interpretations

.

2.1 The Main Theorem.

We start with a simple description of the problem we're interested in.

Main Theorem (Corollary 11.4.8): given finite, abelian group G, denote $\widehat{G}^*$ to be the set of non-trivial irreducible characters of G, and let Q be a probability distribution on G, then

$$||Q^{*k} - U||_{TV}^2 \leq \frac{1}{4} \sum_{\chi \in \widehat{G}^*} |\widehat{Q}(\chi)|^{2k}. \tag{1}$$

There are many confusing symbols in the theorem. We begin with a brief discussion on them. Our object of interest are finite and abelian groups, with the two most important examples $\mathbb{Z}/n\mathbb{Z}$ and $(\mathbb{Z}/2\mathbb{Z})^n$. The author will not repeat the definition of a group.

Random Walks can be done on the elements of a group. A random walk is defined by the group and a symmetric subset of the group denoted as S **(Definition 5.4.3)**. Picture all the elements of the group spread out in space, which there are finitely many. We pick an element from S at each step according to some probability distribution, usually uniformly random. The inverse of any element belonging to S must also be in S.

Probability Distributions on Groups is an intuitive concept. We're simply talking about a discrete mass function that maps each element of the group to a real number while satisfying basic properties of a probability mass function.

At each step of the random walk, we obtain a probability distribution on the group of interest, namely Q^{*k} for the k^{th} step. We're interested in how fast and close this distribution will converge to the uniform distribution. To quantify this interest, we speak of the **Total Variation (Definition 11.1.6)** of a probability distribution of the k^{th} step of a random walk, which takes on two equal forms:

$$||Q^{*k} - U||_{TV} = \max_{A \subseteq G} |Q^{*k}(A) - U(A)| = \max_{A \subseteq G} |Q^{*k}(A) - \frac{|A|}{|G|}|. \tag{2}$$

$$||Q^{*k} - U||_{TV} = \frac{1}{2}\sum_{g\in G}|Q^{*k}(g) - U(g)| = \frac{1}{2}\sum_{g\in G}|Q^{*k}(g) - \frac{1}{|G|}|. \tag{3}$$

This equavalence is established **(Proposition 11.1.8)**. Notice that the absolute value operator takes the same form as the order of a group operator.

Now, these two deinitions, especially (3), seems clear enough, so why do we need to bound the total variation quantity? This is because Q^{*k} is a convoluted probability distribution. If we want to know the probability that we arrive at an element a at the k^{th} step, we must first find all the elements that can lead to a from the $(k-1^{th})$ step, calculate the probability that each one of these elements is transformed to a at the k^{th} step multiplied by the probability that we arrive at these elements in the $(k-1)^{th}$ step, and find the sum. But finding the probability of arriving at each of these elements at the $(k-1)^{th}$ involves more convolution. In conclusion, as k gets large, Q^{*k} becomes increasingly difficult to calculate.

On the other hand, the quantity $\widehat{Q}(\chi)$ is easier to calculate. In fact, for random walks defined on the groups we're interested in and some specific symmetric sets, we're able to express the right side of (1) as some exponential function.

We continue into the next section with a comprehensive explaination of the right side of (1).

2.2 Representations and Characters.

In this section we talk about the foundational results of representation theory that we need to understand the main theorem.

There are two important concepts involved: irreducible characters and the meaning of $\widehat{Q}(\chi)$.

In short, **Irreducible characters** are **Characters** of a group that are irreducible, and characters are defined based on **Representations**. These concepts form some of the most important results in Representation Theory. These results are well proven in the reference text. We need not spend time prove all these results, but it is helpful to look at a summary of some of the results because they give insights on why we define things this particular way. We begin with representations:

- **(Definition 3.1.1)** A representation is a homomorphism or function that maps elements of a group to some invertible matrices, or
$$\varphi : G \to GL(V). \tag{4}$$
- **(Definition 3.1.10)** $W \leq V$ is a G-invarient subspace if for all $g \in G, w \in W$,
$$\varphi_g w \in W. \tag{5}$$
Notice here that both φ_g and w are matrices.
- **(Definition 3.1.15)** Consider φ defined in (4). If V and 0 are the only G-invarient subspaces, then φ is irreducible.
- **(Definition 3.1.21)** Consider φ defined in (4). If
$$V = V_1 \oplus ... \oplus V_n. \tag{6}$$
where V_i are G-invarient and φ_{V_i} is irreducible for all i, then φ is completely reducible.
- **(Theorem 3.2.8)** So why spend all the efforts defining such complicated concepts? This theorem provides an answer. Every representation of a finite group is completely reducible.
- **(Corollary 4.1.8)** For an abelian group G, any of its irreducible representation has degree one, meaning that its corresponding vector space is one-dimensional.

On the other hand, characters provide us a different angle when we try to comprehend representations.

- **(Definition 4.3.1)** The character of a representation with a specified group is defined the following way:
$$\chi_\varphi(g) = Tr(\varphi(g)). \tag{7}$$
This is nice because we get to work with complex numbers not matrices now.

- **(Definition 4.3.6)** Now, characters subset **Class Funtions**, defined as the following: $f : G \to \mathbb{C}$ such that for all $g, h \in G$,

$$f(g) = f(hgh^{-1}). \tag{8}$$

Although not at our center of interest, there are a couple of things on class functions worth mentioning. First, the definition of a class function leads to the intuition that they're constant on a conjugacy class of G. Second, the space of class functions, or $Z(L(G))$, is a subspace of the group algebra. While we know what the space of class functions is, we must define the group algebra first.

- **(Definition 4.2.1)** For a given group G, the **Group Algebra**

$$L(G) = \{f \mid f : G \to \mathbb{C}\}. \tag{9}$$

It must to brought to attention that while $L(G)$ is an inner product space, the space of class functions is the center of some ring.

- **(Theorem 4.3.9)** So why do we care about characters? Because the irreducible characters of G form an orthonormal set of class functions. Fix two irreducible representations of G, φ and ρ, we have

$$\langle \chi_\varphi, \chi_\rho \rangle = 1 \iff \varphi \sim \rho. \tag{10}$$

Another concept crucial to our journey is the idea of a fourier transform on groups. The discussion on this begins with **Dual Groups**.

- **(Definition 5.3.1)** We've previously mentioned that the irreducible characters of a finite Abelian group form an orthonormal set. Denote this set as $\widehat{G}$; this is the dual group.

- **(Proposition 5.3.2)** The dual group must also be a finite abelian group with the corresponding operation defined in the following way: given $\chi, \theta \in \widehat{G}$,

$$(\chi \cdot \theta)(g) = \chi(g)\theta(g). \tag{11}$$

With this in mind,

- **(Definition 5.3.4)** A **Fourier Transform** for a complex-valued function $f : G \to \mathbb{C}$ is $\widehat{f} : \widehat{G} \to \mathbb{C}$:

$$\widehat{f}(\chi) = |G| \langle f, \chi \rangle = \sum_{g \in G} f(g)\overline{\chi(g)}. \tag{12}$$

It follows that given a function in the group algebra of some finite abelian group, intuitively, since we have an orthonormal basis, we must be able to decompose the function into a linear combination of irreducible characters, or elements of the dual group. In other words,

- **(Theorem 5.3.6)** We have something called a **Fourier Inversion**.

$$f \in L(G) \implies f = \sum_{\chi \in \widehat{G}} \langle f, \chi \rangle \chi = \frac{1}{|G|} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi. \tag{13}$$

- **(Theorem 11.4.6)** After a not so large leap, the following important **Plancherel Theorem** becomes available:

$$a, b \in L(G) \implies \langle a, b \rangle = \frac{1}{|G|} \left\langle \widehat{a}, \widehat{b} \right\rangle. \tag{14}$$

Lastly, we talk about convolutions. This concept is especially useful in the context of probability. The intuition lies in the fact that to reach a point in two steps, say, on a graph, we must consider all possibilities of the middle step. In the representation theory language,

- **(Definition 5.2.1)** The **Convolution** $a * b : G \to \mathbb{C}$ is defined by

$$a * b(x) = \sum_{y \in G} a(xy^{-1})b(y). \tag{15}$$

- **(Theorem 5.3.8)** The Fourier Inversion Theorem moreover tells us that the following equality holds:

$$\widehat{a * b} = \widehat{a} \cdot \widehat{b}. \tag{16}$$

Going back to (1), the right side of the equation says that if we know all the irreducible characters of a given group, then we can calculate the total variation quantity without having to do convolution.

2.3 A Quick Proof

Since we have almost everything we need to prove the main theorem, we might as well try to do so on a less specific level. It must also be made clear that $\widehat{G^*}$ denotes the dual group of G without the trivial character.

First, notice that the right side of (3) can be furthur simplified **(Lemma 11.4.5)**. That is,

$$\frac{1}{2} \sum_{g \in G} |Q^{*k}(g) - \frac{1}{|G|}| \leq \frac{1}{2}|G| \cdot \sqrt{\langle Q^{*k} - U, Q^{*k} - U \rangle}. \tag{17}$$

This relation needs introducing something called the 1-norm, but is less relevant in our case. Next, we observe that the emergence of inner products allows us to use the Plancherel Theorem.

$$\frac{1}{2} \cdot |G| \cdot \sqrt{\langle Q^{*k} - U, Q^{*k} - U \rangle} = \frac{1}{2} \cdot \sqrt{|G| \cdot \left\langle \widehat{Q^{*k} - U}, \widehat{Q^{*k} - U} \right\rangle}. \tag{18}$$

We can split the new inner product into three quantities, each of which we're able to solve:

$$\left\langle \widehat{Q^{*k} - U}, \widehat{Q^{*k} - U} \right\rangle = \left\langle \widehat{Q^{*k}}, \widehat{Q^{*k}} \right\rangle - 2 \cdot \left\langle \widehat{Q^{*k}}, \widehat{U} \right\rangle + \left\langle \widehat{U}, \widehat{U} \right\rangle. \tag{19}$$

The following two equations follow directly from the orthogonality relations of irreducible characters. Intuitively, consider a square matrix with the first column as all 1s and 0s for the rest. Consider the identity matrix. In both of these matracies, the ratios of 1s are both one over the dimension.

$$\left\langle \widehat{Q^{*k}}, \widehat{Q^{*k}} \right\rangle = \frac{1}{|G|} + \frac{1}{|G|} \cdot \sum_{\chi \in \widehat{G^*}} \widehat{Q^{*k}}(\chi) \overline{\widehat{Q^{*k}}(\chi)}. \tag{20}$$

$$\left\langle \widehat{U}, \widehat{U} \right\rangle = \left\langle \widehat{Q^{*k}}, \widehat{U} \right\rangle = \frac{1}{|G|}. \tag{21}$$

Lastly, using (16), we combine everything to obtain (1).

2.4 A Couple of Things That Follow

With (1) and the complete set of irreducible characters of a given group, the total variation as a function of steps becomes a numerical value. The book provides a couple of results.

Take the example of $(\mathbb{Z}/2\mathbb{Z})^n$. **(Theorem 11.4.10)** For $c > 0$, let $k \geq (n+1)(\log n + c)/4$, our theorem tells us that

$$||Q^{*k} - U||_{TV}^2 \leq \frac{1}{2}(e^{e^{-c}} - 1). \tag{22}$$

A lower bound is moreover provided. For $0 < c < \log n$ and a large n, let $k \leq (n+1)(\log n - c)/4$, then

$$||Q^{*k} - U||_{TV}^2 \geq 1 - 20e^{-c}. \tag{23}$$

Another useful example is $\mathbb{Z}/n\mathbb{Z}$. In this case, given that the group order is odd and the operation of the random walk is either adding 1 or substracting 1 with equal probability, **(Theorem 11.4.13)** for $n \geq 6$ and $k \geq 36$,

$$\frac{1}{2} e^{-\frac{\pi^2 k}{2n^2} - \frac{\pi^4 k}{2n^4}} \leq ||Q^{*k} - U||_{TV} \leq e^{-\frac{\pi^2 k}{2n^2}}. \tag{24}$$

3 Fourier Transforms

Suppose we have some Pseudo-Boolean function $f : (\mathbb{Z}/2\mathbb{Z})^n \to \mathbb{R}$. These functions are useful in signal processing applications, especially the ones related to my research, which aims at learning these functions under sparse Fourier[1] or Mobius[2] basis. In the lens that we're equipped with, these functions themselves are representations. In 2.2, we've already defined the Fourier transform of these functions that we're interested in as a linear combination of the function outputs multiplied by the conjugate of the corresponding element in the dual group. Let's first straighten two things up: first, how are Fourier transform usually defined in signal processing, and second, what is the dual group of our support of interest.

In [1], the Fourier transform and the inversion formula are given as $G = (\mathbb{Z}/2\mathbb{Z})^n$

$$F[k] = \frac{1}{2^n} \sum_{m \in G} f[m](-1)^{-\langle m,k \rangle}. \tag{25}$$

$$f[m] = \sum_{k \in G} F[k](-1)^{-\langle m,k \rangle}. \tag{26}$$

Notice that we're using slightly different notations here: $\widehat{f}$ is used to denote the Fourier transform in the previous section and F is used to denote the same thing here. We'll stick with F.

Let's try to answer the second question now: that is, let's try to find the dual group of G, or $\widehat{G}$. This is the same as finding, first, every possible irreducible representation of G. Here, we invoke the following fact:

(Proposition 4.5.1) Let $\chi_1, ...\chi_m$ be the set of irreducible representation of $H = \mathbb{Z}/2\mathbb{Z}$. If we can write our group as a product of abelian groups, which we can easily in this case: $G = H^n = H \times ... \times H$, then consider the set of functions $\alpha_{b_1,...,b_n} : G \to \mathbb{C}^*$ with $b_i \in \{1, ..., m\}$ defined as

$$\alpha_{b_1,...,b_n}(g_1, ..., g_n) = \chi_{b_1}(g_1)...\chi_{b_n}(g_n). \tag{27}$$

This is a complete set of irreducible representations of G. In other words, we're just saying that if we know all the irreducible representation of a set of abelian groups, then we know the same for their product.

Moreover, as mentioned in 2.2, the irreducible representations of any abelian group has degree of at most one, so **(Remark 4.3.2)** the complete set mentioned above is actually the same set as the set of irreducible characters!

Lastly, **(Example 5.3.3)** the book gives us that the two characters of H are:

$$\chi_k([m]) = e^{i\pi km} = (-1)^{km} \text{ for } k \in \{0, 1\}, m \in \{0, 1\}. \tag{28}$$

Combining all of these together and along with the (**Theorem 5.3.6**), we arrive back at (26) exactly. This is not a surprising result and in fact we've not even dive too deep. When I think about this now, it all makes sense, but it does deepen my understanding on Fourier Transforms.

References

[1] Yigit Efe Erginbas et al. “Efficiently Computing Sparse Fourier Transforms of q-ary Functions”. In: (2023). arXiv: 2301.06200 [eess.SP].

[2] Justin S. Kang et al. “Learning to Understand: Identifying Interactions via the Mobius Transform”. In: (2024). arXiv: 2402.02631 [cs.LG].

[3] Benjamin Steinberg. *Representation Theory of Finite Groups. An Introductory Approach*. Springer, 2012.