Title
Shapes of Finite Groups through Covering Properties and Cayley Graphs

Author
Yang, Yilong

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Shapes of Finite Groups through Covering Properties and Cayley Graphs

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Mathematics

by

Yilong Yang

2017

ABSTRACT OF THE DISSERTATION

Shapes of Finite Groups through Covering Properties and Cayley Graphs

by

Yilong Yang
Doctor of Philosophy in Mathematics
University of California, Los Angeles, 2017
Professor Terence Chi-Shen Tao, Chair

This thesis is concerned with some asymptotic and geometric properties of finite groups. We shall present two major works with some applications.

We present the first major work in Chapter 3 and its application in Chapter 4. We shall explore the how the expansions of many conjugacy classes is related to the representations of a group, and then focus on using this to characterize quasirandom groups. Then in Chapter 4 we shall apply these results in ultraproducts of certain quasirandom groups and in the Bohr compactification of topological groups. This work is published in the Journal of Group Theory [Yan16].

We present the second major work in Chapter 5 and 6. We shall use tools from number theory, combinatorics and geometry over finite fields to obtain an improved diameter bounds of finite simple groups. We also record the implications on spectral gap and mixing time on the Cayley graphs of these groups. This is a collaborated work with Arindam Biswas and published in the Journal of London Mathematical Society [BY17].

The dissertation of Yilong Yang is approved.

Igor Pak

Sorin Popa

Amit Sahai

Terence Chi-Shen Tao, Committee Chair

University of California, Los Angeles

2017

To my wife who always supports my love for mathematics

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I am deeply grateful to my advisor Terence Tao for his support and his knowledge. I am also grateful to Emmanual Breuillard, László Pyber, László Babai for many helpful comments and conversations. I am also grateful to Richard Schwartz and Yahuda Shalom for setting me on the path of group theory. Finally, I am grateful to Arindam Biswas for our enjoyable collaboration.

VITA

2013	B.S. (Mathematics), Brown University.
2017	Graduate Student, Mathematics Department, UCLA.

PUBLICATIONS

Y. Yang. Ultraproducts of quasirandom groups with small cosocles. *Journal of Group Theory*, 19(6): 1137-1164, 2016.

A. Biswas and Y. Yang. A diameter bound for finite simple groups of large rank. *Journal of the London Mathematical Society*, 95(2): 455-474, 2017.

CHAPTER 1

Introduction

1.1 Guiding intuition and summary of the thesis

This thesis tries to explore several properties of the finite groups and their consequences that might be called the "shape" of finite groups.

Given a geometric object, say a rectangle, there are several ways to measure it. We can measure its length, its width, its circumference, its area and so on. When we compare these different ways of measurement to each other, say length with width, or circumference with area, then we start to see the shape of the rectangle. In this thesis, we shall extend this idea to finite groups, explore several measurements and compare these measurements, study their asymptotics, and see what we can derive as consequences.

The first portion of this thesis deals with shapes of finite groups of a more "local" nature. For a finite group, we can pick any element g and compare the conjugacy classes of $g, g^2, g^3, \ldots$. This gives us a sense of "roundness" or "oblateness", measured from the perspective of g.It turns out that these local oblateness are closely tied with representations of these finite groups. For example, if the groups are too "oblate", then they cannot have any embedding into a "round" compact topological group. This can also have applications in ergodic group theory and topological group theory. More backgrounds on these will be presented at the start of Chapter 4. These results are published in the Journal of Group Theory [Yan16].

The second portion of this thesis deals more specifically with finite simple groups. We compare the order of these groups with the diameter of these groups. The diameter is defined as the following.

Definition 1.1.1. Given a finite group G, its diameter is the supremum of the diameters of all possible connected Cayley graphs of G.

This study of this relation between the diameter and the order of a finite group is similar to the study of growth rate in finitely generated groups. A finite group where the diameter is $\log |G|$ (e.g., $\mathrm{SL}_2(\mathbb{F}_p)$) is similar to a finitely generated group with exponential growth (e.g., finitely generated free groups), and a finite group where the diameter is linear in $|G|$ (e.g., $\mathbb{Z}/n\mathbb{Z}$) is similar to a finitely generated group with linear growth (e.g., $\mathbb{Z}$). This has many applications like obtaining expander graphs and establishing nice random walk properties.

Along the intuition above, one would expect that non-abelian fintie simple groups, the "most unabelian" of finite groups, to behave almost like free groups. So we have the following conjecture by László Babai.

Conjecture 1.1.2 (Babai's Conjecture [BS92]). *For a finite simple group G, its diameter should be*

$$\mathrm{diam}(G) = (\log |G|)^{O(1)}.$$

The first class of simple groups verified for Babai's conjecture was $\mathrm{PSL}_2(\mathbb{Z}/p\mathbb{Z})$ with p prime, by Helfgott [Hel08]. Afterwards, a lot of research was done on the diameters and related expansion properties of these Cayley graphs.

The best result to date are those by Pyber and Szabó [PS16], and Breuillard, Green and Tao [BGT11a]. They verified Babai's conjecture for all finite simple groups of Lie type with bounded rank.[3]

For all non-abelian finite simple groups, Breuillard and Tointon [BT16] also obtained a diameter bound of $\max(|G|^\epsilon, C_\epsilon)$ for arbitrary $\epsilon > 0$ and a constant C_ϵ depending only on ϵ. The diameter bounds in all these previous results depend poorly on the rank of the group. It is one of the aims of this thesis to improve the dependency on the rank, in the case of finite simple groups of Lie type.

[3]The preprint of PS [PS10] was published on arXiv in 2010, and proved Babai's conjecture for all finite simple groups of Lie type with bounded ranks. Unlike PS, BGT first announced these results [BGT10] only for finite simple groups not belonging to the Suzuki and Ree family. Their method also applies to these cases, but this only appeared later in [BGT11b] and [BGG15]. We thank László Pyber for this remark.

On the other hand, a lot of research was also done for the symmetric group S_n and the alternating group A_n. In 1988, Babai and Seress showed the following theorem.

Theorem 1.1.3 (Babai and Seress, [BS88]). *Let $G = \mathrm{S}_n$ or $G = \mathrm{A}_n$, then for any generating set,*

$$\mathrm{diam}(G) \leq \exp(\sqrt{n \log n}(1 + o_n(1)) = \exp(\sqrt{\log |G|}(1 + o_n(1)).$$

This was the best known bound for S_n or A_n for over two decades, until Helfgott and Seress recently showed the following.

Theorem 1.1.4 (Helfgott and Seress, [HS14]). *Let $G = \mathrm{S}_n$ or $G = \mathrm{A}_n$, then for any generating set,*

$$\mathrm{diam}(G) \leq \exp(O((\log n)^4 \log \log n)) = \exp((\log \log |G|)^{O(1)}).$$

In this thesis we give a modest upper bound on the diameter for finite simple groups of Lie type, where the dependency on rank is lessened. This is a collaborated work with Arindam Biswas, published in the Journal of London Mathematical Society [BY17].

1.2 Notation

In this thesis, we denote the commutator subgroup or derived subgroup of a group G by G', and we denote the center of a group G by $\mathrm{Z}(G)$. Given an element g of a group G, we denote the conjugacy class containing g by $C(g)$.

Given two subsets A, B of a group G, we define $AB = \{ab : a \in A, b \in B\}$. For any positive integer n, we define $A^n = \{a_1 ... a_n : a_1, ..., a_n \in A\}$.

We adopt the standard convention of the "big O" and "small o" notation. So $g(x) = O_x(f(x))$ means that there are constants C and x_0, such that for all $x > x_0$, $g(x) \leq Cf(x)$. $g(x) = o_x(f(x))$ means that the limit $\lim_{x\to\infty} \frac{g(x)}{f(x)} = 0$.

There is an unfortunate standard notation of dynkin diagrams and alternating groups. There is a family of dynkin diagrams commonly denoted as A_n, and the alternating groups are also commonly denoted as A_n. In this thesis, we shall always use the upright A_n for alternating groups, and the italicized A_n for dynkin diagrams.

For a complex number z, we use $\mathfrak{Re}(z)$ to denote the real part of z.

CHAPTER 2

Preliminary

2.1 Length functions on finite groups

We shall consider two types of geometries on finite groups. One is to embed a finite group in an actual geometric space, say the unitary groups, and consider the induced geometry. This way, we are studying representations of finite groups through their "extrinsic geometries". The other is to consider the Cayley graph of the group. This way, we are studying their "intrinsic geometries". But for both types of geometries, we are essentially studying the various metric structures of a finite group. We are particularly interested in metrics on groups that are compatible with the underlying group structure. Compare the following two definitions:

Definition 2.1.1. A pseudo-metric space is a set X with a non-negative function $d : X \times X \to \mathbb{R}^+$, such that the following is true:

1. $d(x, x) = 0$ for all $x \in X$.
2. For all $x, y \in X$, $d(x, y) = d(y, x)$.
3. For all $x, y, z \in X$, $d(x, y) + d(y, z) \geq d(x, z)$.

If $d(x, y) \neq 0$ whenever $x \neq y$, then d is a metric, and X is a metric space.

Definition 2.1.2. A pseudo-length function of a group G is a non-negative function $\ell : G \to \mathbb{R}^+$, such that the following is true:

1. $\ell(e) = 0$ for the identity element e.

2. $\ell(g) = \ell(g^{-1})$ for all $g \in G$.

3. $\ell(gh) \leq \ell(g) + \ell(h)$ for all $g, h \in G$.

If $\ell(g) \neq 0$ whenever $g \neq e$, then ℓ is a length function.

Definition 2.1.3. A length function ℓ on a group G is invariant if $\ell(ghg^{-1}) = \ell(h)$ for all $g, h \in G$.

Given a group G and a pseudo-length function ℓ on G, then $d(g, h) = \ell(gh^{-1})$ is a pseudo-metric on G, ang it is a metric if ℓ is a length function. So our task now is to study finite groups via their length functions.

2.2 Unitary groups

Since we shall study length functions of finite groups induced from their representations, we shall hereby lay down some basic properties of unitary groups.

Definition 2.2.1. The unitary group $\mathrm{U}_n(\mathbb{C})$ is the group of complex n by n matrices A such that the conjugate transpose of A is the inverse of A.

Note that $\mathrm{U}_n(\mathbb{C})$ is not just a group, but in fact a Lie group with induced smooth structures as a subset of $\mathrm{M}_{n\times n}(\mathbb{C}) = \mathbb{C}^{n\times n}$.

Definition 2.2.2. The ***Hilbert-Schmidt norm*** of an n by n complex matrix A is defined as $||A|| = \sqrt{\mathrm{Tr}(A * A)}$.

Lemma 2.2.3.

1. *The Lie group $\mathrm{U}_n(\mathbb{C})$ has a Riemannian metric $d : \mathrm{U}_n(\mathbb{C}) \times \mathrm{U}_n(\mathbb{C}) \to \mathbb{R}$ such that $d(A, B) = ||B - A||$ for all $A, B \in \mathrm{U}_n(\mathbb{C})$. The norm here is the Hilbert-Schmidt norm.*

2. *This metric is bi-invariant in the sense that $d(AB, AC) = d(BA, CA) = d(B, C)$ for all $A, B, C \in \mathrm{U}_n(\mathbb{C})$.*

3. *This metric induces a Haar measure, and the volume of* $\mathrm{U}_n(\mathbb{C})$ *under this Haar measure is finite, and* $vol(\mathrm{U}_n(\mathbb{C})) = \frac{(2\pi)^{n(n+1)/2}}{1!2!...(n-1)!}$.

4. *Under the metric* d, $\mathrm{U}_n(\mathbb{C})$ *has non-negative Ricci curvature everywhere.*

5. *All geodesic ball of same radius* r *in* $\mathrm{U}_n(\mathbb{C})$ *will have the same volume* $B_n(r)$. *This volume is bounded by* $O(r^{n^2})$, *where the implied constant is independent of* r. *And if* r *is small enough, then the volume* $B_n(r)$ *is also bounde below by* $\Theta(r^{n^2})$, *where the implied constant is independent of* r.

6. *One parameter subgroups of* $\mathrm{U}_n(\mathbb{C})$ *are exactly geodesics through the identity.*

Proof. These are very standard facts. See, e.g., [Sep07] and [CE75]. □

Proposition 2.2.4. *The function* $\ell_{HS} : \mathrm{U}_n(\mathbb{C}) \to \mathbb{R}^+$ *that sends each matrix* A *in* $\mathrm{U}_n(\mathbb{C})$ *to* $||A - I||$ *is an invariant length function on* $\mathrm{U}_n(\mathbb{C})$.

Proof. Let A, B be any unitary matrices. Let d be the Hilbert-Schmidt distance function on $\mathrm{U}_n(\mathbb{C})$, which is a bi-invariant metric.

Positivity: Clearly $\ell_{HS}(A) = d(A, I) \geq 0$. And we have

$$\ell_{HS}(A) = 0 \iff d(A, I) = 0 \iff A = I.$$

Symmetry:

$$\ell_{HS}(A) = d(A, I) = d(AA^{-1}, IA^{-1}) = d(I, A^{-1}) = \ell_{HS}(A^{-1}).$$

Conjugate Invariance:

$$\ell_{HS}(BAB^{-1}) = d(BAB^{-1}, I) = d(BA, B) = d(A, I) = \ell_{HS}(A).$$

Triangle Inequality:

$$\ell_{HS}(AB) = d(AB, I) \leq d(AB, B) + d(B, I) = d(A, I) + d(B, I) = \ell_{HS}(A) + \ell_{HS}(B).$$

□

We call the length function ℓ_{HS} the Hilbert-Schmidt length function.

2.3 Cayley graphs and Schreier graphs

Definition 2.3.1. Given a group G and a symmetric generating subset S, its ***Cayley graph*** is a graph Γ where the vertices of Γ are elements of G, and two vertices $g, h \in G$ are connected by an edge iff $g = sh$ for some $s \in S$.

Definition 2.3.2. Given an action of a group G on a set X, and a symmetric generating subset S of G, the corresponding ***Schreier graph*** is a graph $\Gamma(G, X)$ where the vertices of $\Gamma(G, X)$ are elements of X, and two vertices $x, y \in X$ are connected by an edge iff $x = sy$ for some $s \in S$.

Note that a Cayley graph is the same as the Schreier graph of G acting on itself.

Definition 2.3.3.

1. Given any graph Γ, it has a natural distance function $d : \Gamma \times \Gamma \to \mathbb{R}^+$ such that $d(v, w)$ is the smallest number of edges required to go from the vertex v to the vertex w.

2. The diameter of a graph Γ is defined as the following:

$$diam(\Gamma) := \sup_{v,w \in \Gamma} d(v, w).$$

   We define the diameter to be infinity if the graph Γ is disconnected.

Proposition 2.3.4. *Given a group G acting transitively on a finite set X, and a finite generating set S of G, the diameter of the Schreier graph has a trivial upper bound of $|X|$, the number of elements in X. It has a trivial lower bound by $\log_{|S|} |X|$.*

Proof. The proof is completely trivial, but it illustrates a nice idea about how we shall approach the study of diameters of Cayley graphs. So we shall describe it here.

Pick any element x_0 of X. Let $B_r(x_0)$ be the "balls of radius r centered at x_0", which is the set of all vertices with distance at most r from x_0. Then the sequence $B_0(x_0), B_1(x_0), ...$ gives a increasing sequence of subsets in X. Note that, since S is a generating set of G, and G acts transitively on X, and X is finite, therefore it must follow that $\Gamma(G, X)$ is connected and $X = B_d(x_0)$ where d is the diameter of the Schreier graph.

So to study the diameter of $\Gamma(G, X)$, it is enough to study the growth rate of the sequence of balls around x_0. Now, if $B_r(x_0) = B_{r+1}(x_0)$, then $B_r(x_0)$ must be a non-empty connected component of $\Gamma(G, X)$, so it must in fact be $\Gamma(G, X)$ itself. So whenever $r < d$, we always have $|B_{r+1}(x_0)| \geq |B_r(x_0)| + 1$. So $|X| = |B_d(x_0)| \geq d$.

Conversely, since S contains the identity, each vertex of $\Gamma(G, X)$ has degree at most $|S| - 1$. So $|B_1(x_0)| \leq |S|$, and for all $2 \leq r \leq d$, we have

$$|B_r(x_0)| \leq |B_{r-1}(x_0)| + (|S| - 2)(|B_{r-1}(x_0)| - |B_{r-2}(x_0)|).$$

By induction we have $|B_r(x_0)| \leq |S|^r$ for all $r \leq d$. So $|X| \leq |S|^d$. □

The above proof illustrated that, to study diameters, it is crucial to understant the growth rate of subsets. This shall be the guiding philosophy for all diameter problems considered in this dissertation.

2.4 Formed spaces

In this thesis we shall study linear groups over finite fields from time to time, so it is important to also study the spaces they act on. In this section we shall formally define formed spaces. Throughout this section, we shall fix a field k and a vector space V over k. Let σ be an automorphism of k such that σ^2 is the identity automorphism.

Definition 2.4.1. Give a function $B : V \times V \to k$, we have the following definitions:

1. B is a ***bilinear form*** if it is linear in both the first and the second argument.

2. B is a ***symmetric bilinear form*** if it is bilinear and $B(v, w) = B(w, v)$ for all $v, w \in V$.

3. B is an ***alternating bilinear form*** if it is bilinear and $B(v, w) = -B(w, v)$ for all $v, w \in V$.

4. B is a ***σ-Hermitian form*** if it is linear in the first argument, and $B(v, w) = \sigma(B(w, v))$.

5. B is non-degenerate if $B(v, w) = 0$ for all $w \in V$ implies that $v = 0$, and $B(v, w) = 0$ for all $v \in V$ implies that $w = 0$.

Definition 2.4.2. A function $q : V \to k$ is called a ***quadratic form*** if $Q(av) = a^2 Q(v)$ for all $a \in k$ and $v \in V$, and $q(v + w) - q(v) - q(w)$ is a symmetric bilinear form. $B_q(v, w) = q(v + w) - q(v) - q(w)$ is the called ***the symmetric bilinear form associated to*** q. A quadratic form q is ***non-degenerate*** if $B_q(v, w) = 0$ for all $w \in V$ and $q(v) = 0$ imply $v = 0$.

2.5 Classification of finite simple groups

Definition 2.5.1. A group is simple if its only normal subgroups are the whole group and the trivial subgroup.

In this thesis, we shall pay special attention to finite simple groups. So in this section, we shall briefly introduce a coarse version of the classification of finite simple groups. To do this, we shall first construct a few families of groups used in our coarse classification of finite simple groups.

Definition 2.5.2. Given a dynkin diagram X, let $\overline{G}$ be the corresponding algebraic group over the algebraic closure of any finite field. Let $F : \overline{G} \to \overline{G}$ be any endomorphism of the algeraic group $\overline{G}$. Let $\overline{G}^F$ be the subgroup of fixed points in $\overline{G}$ by F, and let G be the quotient of the derived subgroup of $\overline{G}^F$ by its center. We have the following cases.

1. If F is some power of the Frobenius endomorphism, then G called the ***finite simple Chevalley groups***.

2. If F is the composition of some power of the Frobenius endomorphism with an automorphism of $\overline{G}$ induced by a non-trivial automorphism of the dynkin diagram X, then G, G', $G/\mathrm{Z}(G)$, $G'/\mathrm{Z}(G')$ are called the ***finite simple Steinberg groups***.

3. If F is the exceptional endomorphism of $\overline{G}$ such that the square of F is an odd power of the Frobenius endomorphism, then G, G', $G/\mathrm{Z}(G)$, $G'/\mathrm{Z}(G')$ are called the ***finite***

simple Suzuki-Ree groups.

These groups G are called ***finite simple groups of Lie type***. The rank of G is the smallest rank of all dynkin diagrams that give rise to G.

Example 2.5.3. *Let V be a vector space of dimension n over $\mathbb{F}_q$, the field of order q. The following four kinds of groups are called the **finite simple classical groups**:*

1. *Let $\mathrm{SL}_n(V)$ be the group of linear maps with determinant 1. Then the group $\mathrm{PSL}_n(V) = \mathrm{SL}_n(V)/\mathrm{Z}(\mathrm{SL}_n(V))$ is a finite simple Chevalley group for the dynkin diagram A_{n-1}. In particular, they have rank at most $n-1$.*

2. *Let $\mathrm{Sp}_n(V)$ be the group of linear maps fixing a non-degenerate alternating bilinear form on V. Then the group $\mathrm{PSp}_n(V) = \mathrm{Sp}_n(V)/\mathrm{Z}(\mathrm{Sp}_n(V))$ is a finite simple Chevalley group for the dynkin diagram $C_{\frac{n}{2}}$. In particular, they have rank at most $\frac{n}{2}$.*

3. *Let $\mathrm{U}_n(V)$ be the group of linear maps fixing a non-degenerate Hermitian form on V, and let $\mathrm{SU}_n(V)$ be the subgroup of $\mathrm{U}_n(V)$ of determinant 1. Then the group $\mathrm{PSU}_n(V) = \mathrm{SU}_n(V)/\mathrm{Z}(\mathrm{SU}_n(V))$ is a finite simple Steinberg group for the dynkin diagram A_{n-1}. In particular, they have rank at most $n-1$.*

4. *Let $\mathrm{O}_n(V)$ be the group of linear maps fixing a non-degenerate quadratic form on V. Let $\Omega_n(V)$ be its derived subgroup. Then the group $\mathrm{P}\Omega_n(V) = \Omega_n(V)/\mathrm{Z}(\Omega_n(V))$ is a finite simple group of Lie type. If n is odd, then it is a finite simple Chevalley group for the dynkin diagram $B_{\frac{n-1}{2}}$, so G has rank at most $\frac{n-1}{2}$. If n is even, then there are two possible choices of quadratic forms, one gives rise to the finite simple Chevalley group for the dynkin diagram $D_{\frac{n}{2}}$, the other gives rise to the finite simple Steinberg group for the dynkin diagram $D_{\frac{n}{2}}$. Either way, G has rank at most $\frac{n}{2}$.*

Now we are ready to present the following coarse classification, which is essentially a much shorter way to write the famous classification of finite simple groups.

Theorem 2.5.4 (Classification of finite simple groups)**.** *Any finite simple group must belong to at least one of the following families of groups:*

1. *Cyclic groups of prime order.*

2. *Alternating groups.*

3. *Finite simple classical groups.*

4. *Finite simple groups of Lie type of rank at most 8.*

5. *26 Sporadic groups.*

At times, results on finite simple groups can usually be generalized to a class of groups very close to being fintie simple, called quasisimple groups.

Definition 2.5.5. A group G is ***quasisimple*** if it is a finite central extension of a finite simple group, i.e., G is finite and $G/\mathrm{Z}G$ is a finite simple group.

It turns out that each finite simple group has a unique "universal covering group", which serves as the "largest" possible finite central extension of it. And a group is quasisimple if and only if it is a non-trivial quotient of some universal covering group, see e.g., [Asc00].

So in order to study quasisimple groups, it is usually enough to study these universal covering groups. The order of the center of these groups are bounded by the size of the Schur multipliers of the finite simple groups they cover. And the Schur multipliers are classified as part of the classification of finite simple groups, so they are all known.

With finitely many exceptions, the universal covering groups for finite simple classical groups are usually the groups before projectivization.

2.6 The field with one element

The concept of the field with one element is first suggested by Jacques Tits [Tit57]. There is no field with one element. However, the notion "field with one element" refers to a guiding philosophy that combines the study of alternating groups and the study of finite simple groups of Lie type. The general idea is that the symmetric groups S_n should behave like the general linear groups $\mathrm{GL}_n(\mathbb{F}_q)$ with $q = 1$, and the alternating groups A_n should therefore behave like the special linear groups $\mathrm{SL}_n(\mathbb{F}_q)$ with $q = 1$. We shall briefly describe this intuition here in a combinatorial way.

Suppose for now that there really is a field with one element $\mathbb{F}_1$. Then the obvious obstacle is that any affine space over it $\mathbb{A}^n(\mathbb{F}_1)$ must only contain one point, the origin. However, things are much better if one looks at the projective spaces instead.

To start, the projective space of dimension 0 is just one point. For each positive integer n, then $\mathbb{P}^n(\mathbb{F}_1)$, the projective space over $\mathbb{F}_1$ of dimension n, should be $\mathbb{A}^n(\mathbb{F}_1)$ plus the "points at infinity", which form a projective space of dimension $n-1$. So as a set, $\mathbb{P}^n(\mathbb{F}_1) = \mathbb{A}^n(\mathbb{F}_1) \cup \mathbb{P}^{n-1}(\mathbb{F}_1)$. In particular, by induction $\mathbb{P}^n(\mathbb{F}_1)$ should simply be a set of $n+1$ points.

Since projective spaces should be homogeneous, it should not matter which point we pick to be $\mathbb{A}^n(\mathbb{F}_1)$ inside $\mathbb{P}^n(\mathbb{F}_1)$. So in particular, any subset with n elements is a $n-1$ dimensional subspace of $\mathbb{P}^n(\mathbb{F}_1)$. It follows that any subset with k elements is a $k-1$ dimensional subspace of $\mathbb{P}^n(\mathbb{F}_1)$.

Now let us go back to affine spaces. The projective space $\mathbb{P}^{n-1}(\mathbb{F}_1)$ is suppose to be the set of lines through the origin in $\mathbb{A}^n(\mathbb{F}_1)$. So we see that even though $\mathbb{A}^n(\mathbb{F}_1)$ has only 1 point, it nevertheless has n lines through the origin, and it has $\binom{n}{k}$ subspaces of dimension k. And distinct lines are linearly independent. Finally, any quadratic form on $\mathbb{A}^n(\mathbb{F}_1)$ must be trivial, so for visual convenience, one can assume the n lines of $\mathbb{A}^n(\mathbb{F}_1)$ are perpendicular, and are in fact the n "axis" of $\mathbb{A}^n(\mathbb{F}_1)$.

Now a general linear group over $\mathbb{A}^n(\mathbb{F}_1)$ should be a permutation of lines in $\mathbb{A}^n(\mathbb{F}_1)$ that preserves the linear structure. Since the linear structure is trivial, it follows that the general linear group should be the symmetric group S_n. Since the n lines of $\mathbb{A}^n(\mathbb{F}_1)$ are linearly independent, they form a "basis", and the matrix representation of S_n on this "basis" is simply the representation of S_n as n by n permutation matrices. So one can define the transpose, determinant and so forth for permutations in S_n. So we see that A_n corresponds to the special linear groups, and S_n and A_n also corresponds to the orthogonal and special orthogonal groups.

Finally, we conclude this section by a few countings that further support the philosophy of treating S_n and A_n as $\mathrm{GL}_n(\mathbb{F}_q)$ and $\mathrm{SL}_n(\mathbb{F}_q)$ when $q = 1$.

Definition 2.6.1. For a prime power q, we define the following ***q-analogues***:

1. $[n]_q = (q^n - 1)/(q - 1)$.

2. $[n]_q! = [1]_q \cdot [2]_q \cdot \ldots \cdot [n]_q$.

3. $\binom{n}{k}_q = \frac{[n]_q!}{[k]_q! \cdot [n-k]_q!}$

Note that, in the above notions, if one take the limit as q goes to 1, then $[n]_q = n$, $[n]_q! = n!$, and $\binom{n}{k}_q = \binom{n}{k}$.

Proposition 2.6.2. *We have the following analogies:*

1. $|\mathbb{P}^n(\mathbb{F}_q)| = [n+1]_q$, *and* $|\mathbb{P}^n(\mathbb{F}_1)| = n + 1$.

2. *The number of k dimensional subspaces of $\mathbb{A}^n(\mathbb{F}_q)$ is $\binom{n}{k}_q$, and the number of k dimensional subspaces of $\mathbb{A}^n(\mathbb{F}_1)$ is $\binom{n}{k}$.*

3. *There are $[n]_q!$ maximal flags in $\mathbb{A}^n(\mathbb{F}_q)$, and $n!$ maximal flags in $\mathbb{A}^n(\mathbb{F}_1)$.*

4. $|\mathrm{GL}_n(\mathbb{F}_q)| = q^{\frac{n(n-1)}{2}} |\mathrm{GL}_1(\mathbb{F}_q)|^n [n]_q!$, *and* $|\mathrm{S}_n| = 1^{\frac{n(n-1)}{2}} |\mathrm{S}_1|^n n!$.

CHAPTER 3

Conjugacy Expansion

3.1 Expansions of conjugacy class

Recall that, for an element g of a group, we denote its conjugacy class by $C(g)$. In this chapter, we aim to study the shape of a finit group G by looking at $C(g), C(g)^2, C(g)^3, \ldots$, a sequence of subsets of G.

Definition 3.1.1. Given a finite group G and an element $g \in G$, we say g has covering number K if $C(g)^K = G$. If no such positive integer K exists, then g has covering number ∞.

Remark 3.1.2. *If an element has covering number K, then it has covering number K' for all $K' \geq K$.*

Example 3.1.3.

1. *The covering numbers for an element g in finite abelian groups G is $|G|$ only when G is cyclic and g is a generator. Otherwise, the covering number is ∞. So abelian groups have the worst covering numbers.*

2. *On the opposite side, suppose G is a non-abelian finite simple group. By a theorem of Liebeck and Shalov [LS01], if S is a normal subset of G, then $G = S^K$ for any $K \geq O(\frac{\log G}{\log S})$. So these groups have optimal covering numbers.*

In this section, we shall show that the covering number of a finite group is to some degree controlled by its representations.

Definition 3.1.4. Given a finite group G and a representation $\rho : G \to \mathrm{U}_n(\mathbb{C})$, then the ***representation pseudo-length function of*** ρ on G is the length function ℓ_ρ that maps each $g \in G$ to $\ell_{HS}(\rho(g))$.

Since any representation pseudo-length function must be invariant, we can visualize our group as sitting in the unitary group with each conjugacy class contained in the boundary of a ball around the identity matrix.

Proposition 3.1.5. *Given a finite group G and an element $g \in G$ with covering number K, then $K \geq \frac{\sqrt{2\deg(\rho)}}{\ell_\rho(g)}$ for any representation ρ with no trivial subrepresentation.*

Proof. Let χ_ρ be the character for ρ. Then for any $h \in G$, we have the following:

$$\begin{aligned}\ell_\rho(h)^2 &= \mathrm{Tr}((\rho(h) - I)^*(\rho(h) - I)) \\ &= \mathrm{Tr}(\rho(h)^*\rho(h) - \rho(h)^* - \rho(h) + I) \\ &= \mathrm{Tr}(2I - \rho(h) * -\rho(h))) \\ &= 2\deg(\rho) - 2\mathfrak{Re}(\chi_\rho(h)).\end{aligned}$$

Since ρ has no trivial subrepresentation, by orthogonality we have $\sum_{h\in G}\chi_\rho(h) = 0$. So in particular, there is an $h \in G$ such that $\mathfrak{Re}(\chi_\rho(h)) \leq 0$. So $\ell_\rho(h) \geq \sqrt{2\deg(\rho)}$.

Now since $h \in G = C(g)^K$ and ℓ_ρ is an invariant pseudo-length function, we have the following:

$$\sqrt{2\deg(\rho)} \leq \ell_\rho(h) \leq K\ell_\rho(g)$$

$\square$

Conversely, the representations of a finite group is also controlled by the covering number of its elements. In particular, if in a finite groups of large order, all non-trivial elements has small covering number, then the group has no faithful representations of small dimension. This can be easily seen as a corollary of the Camille Jordan's theorem for finite linear groups [Jor78]. We shall present another proof here, which is similar in flavor with the proof of Jordan's theorem, but is slightly more in line with proofs of later sections of this chapter.

Proposition 3.1.6. *Fix positive integers K and D. Then there is a constant $C_{K,D}$, such that for any finite group G with order $|G| \geq C_{K,D}$, if all elements of G has covering number at most K, then G has no faithful representation of degree less than D.*

Proof. Suppose G has a faithful representation ρ of degree less than D, and every non-trivial element of G has covering number at most K. By throwing away trivial subrepresentations, we can assume that ρ is a faithful representation of degree less than D with no trivial subrepresentation.

By Proposition 3.1.5, for each $g \in G$ with covering number at most K, we have

$$\ell_\rho(g) \geq \frac{\sqrt{2\deg(\rho)}}{K}$$

So for any distinct $g, h \in G$, the Hilbert-Schmidt distance of $\rho(g), \rho(h)$ is at least $\frac{\sqrt{2\deg(\rho)}}{K}$.

So we can pack balls of radius $\frac{\sqrt{2\deg(\rho)}}{2K}$ and centered at each element of $\rho(G)$. These balls will be disjoint in $\mathrm{U}_{\deg(\rho)}(\mathbb{C})$, and they all have the same volume depending only on $\deg(\rho)$. So this will run into a contradiction if $|G| \geq C_{K,D}$ for some constant $C_{K,D}$ depending only on K and D. □

3.2 Covering properties and oblateness of a discrete group

In last section, we showed that if all non-trivial elements of a group have small covering number, then its faithful representations are restricted. It turned out that this is an overkill. We only need several related elements to have small covering number, and we do not need to assume that the group is finite or that the representation is faithful. In this section, we shall lay out the concept of "oblateness" of a discrete group, and prove their relation to their representations.

Definition 3.2.1.

1. An element g of a group G is said to have ***symmetric covering number*** K if $C(g)^K C(g^{-1})^K = G$.

2. Let m be a positive integer or ∞. Then an element $g \in G$ has ***the (symmetric) covering property*** (K, m) if g^i has (symmetric) covering number K for all $1 \leq i \leq m$.

3. A group G has ***the (symmetric) covering property*** (K, m) if it has an element $g \in G$ with the (symmetric) covering property (K, m).

4. A group G has ***the (symmetric) covering property*** (K, m) mod N for some normal subgroup N if G/N has the (symmetric) covering property (K, m).

Definition 3.2.2.

1. A pair of elements (g_1, g_2) of a group G is said to have ***symmetric double covering number*** (K_1, K_2) if we have $C(g_1)^{K_1}C(g_1^{-1})^{K_1}C(g_2)^{K_2}C(g_2^{-1})^{K_2} = G$.

2. Let m_1, m_2 be positive integers or ∞. A pair of elements (g_1, g_2) in G has ***the symmetric double covering property*** $[(K_1, m_1), (K_2, m_2)]$ if (g_1^i, g_2^j) has symmetric double covering number (K_1, K_2) for all $1 \leq i \leq m_1, 1 \leq j \leq m_2$.

3. A group G has ***the symmetric double covering property*** $[(K_1, m_1), (K_2, m_2)]$ if it has a pair of elements (g_1, g_2) in G with the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$.

4. A group G has ***the symmetric double covering property*** $[(K_1, m_1), (K_2, m_2)]$ mod N for some normal subgroup N if G/N has the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$.

Remark 3.2.3.

1. *Suppose $K < K'$. Then an element with covering number K has covering number K'. In general, the (symmetric) covering property (K, m) implies the (symmetric) covering property (K', m') when $K' \geq K, m' \leq m$. A similar statement is also true for the symmetric double covering properties.*

2. *Any symmetric covering property is always weaker than the corresponding non-symmetric covering property.*

3. *Any group with the symmetric covering property (K, m) has the symmetric double covering property $[(1, \infty), (K, m)]$. This is easily seen by taking g_1 to be the identity, and taking g_2 to be the element with the symmetric covering property (K, m).*

4. *In our definition of the symmetric double covering properties, since $C(g_1)$ and $C(g_2)$ are conjugate invariant subsets of G, they necessarily commute, i.e., $C(g_1)C(g_2) = C(g_2)C(g_1)$. So the order of (K_1, m_1) and (K_2, m_2) does not matter.*

5. *By imitating the definition of the symmetric double covering properties, one can in fact define the symmetric n-tuple covering properties for groups. As n grows larger and larger, the corresponding covering properties will become weaker and weaker. Note that most results throughout this paper would still hold by replacing the symmetric double covering properties by the symmetric n-tuple covering properties, though for our purpose here, the symmetric double covering properties are enough.*

Now, suppose a finite group G is a subgroup of a compact connected Lie group M. M will have a bi-invariant Riemannian metric unique up to scalar multiples. Under a fixed bi-invariant metric, cyclic subgroups of G lies in closed geodesics of M through the origin, and each conjugacy class of G lies in the boundary of a ball centered around the origin. For visual convenience, assume that M looks like the earth, then cyclic subgroups of G lies in the circles of longitude, while the conjugacy class lies in the circles of latitude.

Now suppose G has an element g with covering property (K, m). The elements $g, g^2, ..., g^m$ are packed in a circle of longitude. So if m is large, then there will be a g^i close to the origin. This g^i has covering number K, so if K is small, $C(g^i)$ must be large. This means that the circle of latitude through g^i must be large. So a circle of latitude close to the north pole has large length. Therefore our earth would look "flat". In fact, there are m circles of latitude with large length. So our earth looks like an oblate spheroid.

For the general case though, the geodesics going through $g, g^2, ..., g^m$ might not be as nice as a circle of longitude, and might have arbitrarily large length. So instead of packing these elements in a geodesic, it is slightly better to pack disjoint geodesic balls around them into M.

In short, we can think of the pair (K, m) as measuring the "oblateness" of a group. If m is large and K is small, then G with the covering property (K, m) is more "oblate". And if m is small and K is large for all possible covering properties (K, m) that G has, then G is more rounded.

The idea is that, if a finite group is too "oblate", and a unitary group is not "oblate" enough, then the finite group cannot fit into the unitary group, and therefore any representation must be trivial. A more rigorous statement is the following proposition:

Proposition 3.2.4. *There is a function $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that, for any positive integers K_1, m_1, K_2, m_2 with $\frac{m_i}{K_i^{n^2}} > f(n)$ for $i = 1, 2$, then any discrete group with the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$ cannot have non-trivial representations of degree less than n.*

Proof. Suppose G is a discrete group with the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$ and a non-trivial representation ρ of degree n.

Let g_1 and g_2 be the pair of elements of G with the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$. For any $\epsilon_1 > 0$, if m_1 is large enough, then by packing balls of radius $\frac{\epsilon_1}{2}$ into $\mathrm{U}_n(\mathbb{C})$, we see that two points of $\{\rho(g_1), \rho g_1^2, ..., \rho(g_1^{m_1})\}$ will have distance less than ϵ_1. Furthermore, if ϵ_1 is small enough, then geodesic balls in $\mathrm{U}_n(\mathbb{C})$ of radius ϵ will have radius $B_n(\epsilon_1)^{n^2}$ for some constant B_n depending only on n. So if $m_1 \geq \frac{vol(\mathrm{U}_n(\mathbb{C}))}{B_n(\epsilon_1)^{n^2}}$ for ϵ_1 small enough, then two points of $\{\rho(g_1), \rho g_1^2, ..., \rho(g_1^{m_1})\}$ will have distance less than ϵ_1. Then there is an integer $1 \leq i \leq m_1$ such that $\rho(g_1^i)$ is at most ϵ_1 away from the identity matrix I of $\mathrm{U}_n(\mathbb{C})$. Similarly, for any $\epsilon_2 > 0$, there is an integer $1 \leq j \leq m_2$ such that $\rho(g_2^j)$ is at most ϵ_2 away from I.

Now let h be any element of G. Then since the pair (g_1, g_2) has the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$, we conclude that $d_{HS}(\rho(h), I) \leq 2K_1\epsilon_1 + 2K_2\epsilon_2$.

Now, since ρ is non-trivial, its image $\rho(G)$ contains a non-trivial finite cyclic subgroup of $\mathrm{U}_n(\mathbb{C})$. By Lemma 3.2.5, it must contain an element of length at least $\sqrt{2}$. So there is an element h of G such that $\rho(h)$ is at least $\sqrt{2}$ away from I.

To sum up, we have $\sqrt{2} \leq 2K_1\epsilon_1 + 2K_2\epsilon_2$.

So we can choose a function f' such that, if $\frac{m_i}{K_i^{n^2}} > f'(n)$, then we can take ϵ_1 and ϵ_2 to be very small with $2K_1\epsilon_1 + 2K_2\epsilon_2 < \sqrt{2}$. Then G will not be able to have non-trivial representations of dimension n.

To finish our proof, we only need to take $f(n) = \sup_{1 \leq n' < n} f'(n')$. $\square$

Lemma 3.2.5. *Any non-trivial cyclic subgroup of* $\mathrm{U}_n(\mathbb{C})$ *contains an element* $\sqrt{2}$ *away from the identity matrix.*

Proof. Let A be any nontrivial element of $\mathrm{U}_n(\mathbb{C})$ of finite order. Let $\lambda_1, ..., \lambda_n$ be its eigenvalues, and WLOG say $\lambda_1 \neq 1$. Then λ_1 is a primitive k-th root of unity for some k. Replacing A by a proper power of itself, we may assume that λ_1 is an k-th root of unity closest to -1. Then in particular, $|\lambda_1 - 1| > \sqrt{2}$.

Then we know

$$\ell(A)^2 = \mathrm{Tr}(A - I)^*(A - I) = \sum_{i=1}^{n} |\lambda_i - 1|^2 \geq |\lambda_1 - 1|^2 > 2.$$

Now suppose A has infinite order. Let $\lambda_1, ..., \lambda_n$ be its eigenvalues, and WLOG say $\lambda_1 \neq 1$. Then λ_1 is an element of infinite order on the unit circle. Replacing A by a proper power of itself, we may assume that λ_1 is arbitrarily close to -1. Then in particular, $|\lambda_1 - 1| > \sqrt{2}$. Then we are done by the same computation. $\square$

Remark 3.2.6.

1. *As can be seen from the proof of Proposition 3.2.4, for a covering property* (K, m)*, the quantity* m *is essentially compared with the volume of* $\mathrm{U}_n(\mathbb{C})$*, whereas the quantity* K *is essentially compared with the length of the shortest geodesic in* U_n*. The study of the relation between the volume and the length of the shortest geodesic is called systolic geometry. So essentially, a covering property* (K, m) *is like a finite analogue of systolic properties of a finite group.*

2. *In Proposition 3.2.4, we can in fact replace the unitary group* $\mathrm{U}_n(\mathbb{C})$ *by any compact connected Lie group* M *with a fixed bi-invariant metric. Then we can essentially use*

the same proof. The constant $f(n)$ would have to be changed though, depending on the Lie group M.

3.3 Covering properties and cosocles

We shall show that symmetric (double) covering properties ignore cosocles in general.

Definition 3.3.1. For a group G, we define its ***cosocle*** $Cos(G)$ to be the intersection of all maximal normal subgroups of G.

Lemma 3.3.2. *Let G be a group, and let N be a normal subgroup of G contained in its cosocle. Let C be a conjugate invariant symmetric subset of G, such that $CN = G$. Then for any non-empty conjugate invariant subset $S \subseteq G$, $SC = S$ iff $S = G$.*

Proof. Suppose $SC = S$ and $S \neq G$. Then we have $SC^i = S$ for any positive integer i. So S must contain the subgroup generated by C. Since C is conjugate invariant, the subgroup generated by C is a normal subgroup, and it is a proper normal subgroup since it is contained in $S \neq G$. In particular, C is contained in a maximal normal subgroup M of G.

But since N is in the cosocle, it is contained in M. So

$$CN \subseteq MN = M \subsetneq G.$$

This is a contradiction. $\square$

Proposition 3.3.3. *Let G be a group with the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$ mod N for a normal subgroup N contained in the cosocle, and suppose that N contains exactly n conjugacy classes of G. Then G has the symmetric double covering property $[((3n-2)K_1, m_1), ((3n-2)K_2, m_2)]$.*

Proof. Find $g_1, g_2 \in G$ such that (g_1N, g_2N) has symmetric double covering number (K_1, K_2) in G/N. Let $C := C(g_1)^{K_1} C(g_1^{-1})^{K_1} C(g_2)^{K_2} C(g_2^{-1})^{K_2}$. Then by assumption, C is mapped surjectively onto G/N through the quotient map. So $CN = G$.

Now N contains exactly n conjugacy classes of G. I claim that C^{3t} contains at least $t+1$ conjugacy classes of G in N, which would imply that $C^{3n-3} \supseteq N$. Then $C^{3n-2} \supseteq CN = G$, finishing our proof.

We proceed by induction. As a convention we define C^0 to be $\{e\}$. Then the claim is true when $t = 0$.

Now assume the statement is true for some $t < n$. Then C^{3t} contains $t+1$ conjugacy classes of G in N. Let them be $C_1, ..., C_{t+1}$. Then we have $C^{3t+1} \supseteq C(\bigcup_{i=1}^{t+1} C_i)$. Suppose for contradiction that C^{3t+2} is disjoint from $C(N - \bigcup_{i=1}^{t+1} C_i)$. Then we observe that

$$C(N - \bigcup_{i=1}^{t+1} C_i) \supseteq CN - C(\bigcup_{i=1}^{t+1} C_i) = G - C(\bigcup_{i=1}^{t+1} C_i) \supseteq G - C^{3t+1}.$$

So $C^{3t+2} \subseteq C^{3t+1}$. Then Lemma 3.3.2 implies that $C^{3t+2} = C^{3t+1} = G$. This contradicts the assumption that C^{3t+2} is disjoint from $C(N - \bigcup_{i=1}^{t+1} C_i)$.

So, C^{3t+2} intersects with $C(N - \bigcup_{i=1}^{t+1} C_i)$. Let g be an element in this intersection. Then $g \in CC_{t+2}$ for some conjugacy class C_{t+2} of G in N disjoint from $C_1, ..., C_{t+1}$. Find $h \in C_{t+2}$ such that $g \in Ch$. Then since C is symmetric, we have $h \in Cg \subseteq C^{3t+3}$. So C^{3t+3} intersects with C_{t+2}. Since C^{3t+3} is conjugate invariant, we conclude that C^{3t+3} contains C_{t+2}.

Finally, since $e \in C$, we see that C^{3t+3} also contains $C_1, C_2, ..., C_{t+1}$. So C^{3t+3} contains $t+2$ conjugacy classes of G in N. □

Proposition 3.3.4. *Let G be a group with the symmetric covering property (K, m) mod N for a normal subgroup N contained in the cosocle, and suppose that N contains exactly n conjugacy classes of G. Then G has the symmetric covering property $((3n-2)K, m)$.*

Proof. Same strategy as Proposition 3.3.3. □

3.4 Alternating groups with Covering properties

In the following section, we shall show that all non-abelian finite simple groups has good covering properties, as long as the order of the group is large enough. We shall start with alternating groups.

Definition 3.4.1. An even permutation $\sigma \in \mathrm{A}_n$ is ***exceptional*** if its cycles in its decomposition has distinct odd lengths, or equivalently, if its conjugacy class in A_n is different from its conjugacy class in S_n.

Lemma 3.4.2 (Brenner [Bre78, Lemma 3.05])**.** *If an even permutation $\sigma \in \mathrm{A}_n$ is fixed-point free and non-exceptional, then $\mathrm{A}_n = C(\sigma)^4$.*

Proposition 3.4.3. *For any $m \in \mathbb{Z}^+$, A_n has the covering property $(4, m)$ for large enough n.*

Proof. Pick any odd prime $p > m$, and pick another prime $q > p$.

Since p, q are necessarily coprime, for any large enough integer n, we can find positive integers a, b such that $n = ap + bq$. Let $\sigma \in \mathrm{S}_n$ be a permutation composed of a p-cycles and b q-cycles, where all cycles are disjoint.

Since p, q are odd, σ is an even permutation in A_n. Furthermore, for large enough n, a or b can be chosen to be larger than 1, so σ will be non-exceptional. Since σ is also fixed-point free by construction, Lemma 3.4.2 implies that $\mathrm{A}_n = C(\sigma)^4$.

Now clearly σ^i will also have a cycle decomposition of a p-cycles and b q-cycles for all $1 \leq i \leq p-1$, and this implies that $\mathrm{A}_n = C(\sigma^i)^4$ for all $1 \leq i \leq p-1$. So A_n has the covering property $(4, p-1)$. Since $p - 1 \geq m$, A_n has the covering property $(4, m)$. $\square$

Proposition 3.4.4. *Let G be a quasisimple group over an alternating group. Then for any $m < \infty$, G has the symmetric covering property $(16, m)$ if $|G|$ is large enough.*

Proof. Alternating groups A_n with $n > 7$ have Schur multiplier 2. So $|G| \leq 2|G/\mathrm{Z}(G)|$. So if $|G|$ is large enough, then the alternating group it covers must be large enough. Then Proposition 3.4.3 implies that A_n has the covering property $(4, m)$. So G has the covering property $(16, m)$. $\square$

3.5 Finite simple groups of Lie type of bounded rank with Covering properties

Lemma 3.5.1 (Stolz and Thom [ST13, Proposition 3.8])**.** *There is a function $K : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that, in any finite simple group of Lie type of rank $\leq r$, any non-identity element will have covering number $K(r)$.*

We shall fix this function $K(r)$ from now on.

Lemma 3.5.2 (Babai, Goodman and Pyber [BGP97, Proposition 5.4])**.** *Let k be any positive integer. Then for any finite simple group G, if $|G| \geq k^{k^2}$, then $|G|$ has a prime divisor greater than k.*

Proposition 3.5.3. *Let G be a finite simple group of Lie type of rank $\leq r$. For any $m < \infty$, G has the covering property $(K(r), m)$ if $|G|$ is large enough.*

Proof. For any $m \in \mathbb{Z}^+$, suppose G has order larger than m^{m^2}, and thus have an element g of prime order $p > m$. Then g^i are non-identity for all $1 \leq i \leq p-1$. Then Lemma 3.5.1 states that all these elements have covering number $K(r)$. So G has the covering property $(K(r), m)$. □

Corollary 3.5.4. *Let G be a finite quasisimple group of rank $\leq r$. For any $m < \infty$, G has the symmetric covering property $(K(r)\max(3r+1, 10), m)$ if $|G|$ is large enough.*

Proof. There are only finitely many quasisimple groups covering the same finite simple group, and there are only finitely many finite simple groups of a given order. So if $|G|$ is large enough, then the finite simple groups it covers must have large enough order.

Therefore, it is enough to show that, if a finite simple group of Lie type G with rank $\leq r$ has the covering property (K, m), then any perfect central extension G' of it will have the covering property $(K(r)\max(3r+1, 10), m)$.

Let Z be the center of G'. Then Z will be the cosocle of G', and the Schur multiplier of the simple group G would provide an upper bound for $|Z|$. Since G has a rank at most

r, by going through the list of finite simple groups, its Schur multiplier has a size at most $\max(r+1,4)$ if G is large enough. So if G has the covering property (K,m), Proposition 3.3.4 implies that G' has the symmetric covering property $(K(r)\max(3r+1,10),m)$. $\square$

3.6 Jordan length function for finite classical groups

For our purpose of studying covering properties, the best length function for any group is the following one.

Definition 3.6.1. For any group G, the conjugacy length function $\ell_C : G \to \mathbb{R}^+$ maps each g to $\frac{\log |C(G)|}{\log |G|}$.

However, this is annoying to compute from time to time. Here we shall introduce a length function for finite classical groups that is asymptotically the same as the conjugacy length function, but it is much easier to compute in various cases.

Definition 3.6.2. Let g be an $n\times n$ matrix over a finite field F. Let $m_g := \sup_{a\in F^\times} \dim(\ker(a-g))$. Then the ***Jordan length*** of g is $\ell_J(g) := \frac{n-m_g}{n}$

The asymptotic equivalence of Jordan length function and the conjugacy length function can be deduced essentially from the work of Liebeck and Shalev [LS01], and the case of general linear groups and special linear groups is explicitely worked out by Stolz and Thom [ST13]. Here we shall obmit the proof of these facts, but merely cite the following consequence of it which is the most useful to our purpose.

Proposition 3.6.3. *Let G be any subgroup of* $\mathrm{GL}_n(F)$ *for some finite field F. The function ℓ_J on G is a pseudo length function.*

Proof. Non-negativity: For any $g \in G$,

$$m_g = \sup_{a\in F^\times} \dim(\ker(a-g)) \leq n.$$

So $\ell_J(g) = \frac{n-m_g}{n} \geq 0$.

Symmetry: For any $g \in G$, any $a \in F^\times$, and any vector $v \in F^n$, we have

$$v \in \ker(a-g) \iff av = gv \iff g^{-1}v = a^{-1}v \iff v \in \ker(a^{-1}-g^{-1}).$$

As a result,

$$m_g = \sup_{a \in F^\times} \dim(\ker(a-g)) = \sup_{a \in F^\times} \dim(\ker(a^{-1}-g^{-1})) = m_{g^{-1}}.$$

So $\ell_J(g) = \ell_J(g^{-1})$.

Conjugate-invariance: For any $g, h \in G$, any $a \in F^\times$, and any vector $v \in F^n$, we have

$$v \in \ker(a-g) \iff av = gv \iff ahv = (hgh^{-1})hv \iff hv \in \ker(a-hgh^{-1}).$$

As a result,

$$m_g = \sup_{a \in F^\times} \dim(\ker(a-g)) = \sup_{a \in F^\times} \dim(\ker(a-hgh^{-1})) = m_{hgh^{-1}}.$$

So $\ell_J(g) = \ell_J(hgh^{-1})$.

Triangle inequality: For any $g, h \in G$, any $a, b \in F^\times$, and any vector $v \in F^n$, we have

$$v \in \ker(a-g) \cap \ker(a-abh^{-1}) \implies gv = av = abh^{-1}v \implies v \in \ker(abh^{-1}-g).$$

So we know $\ker(a-g) \cap \ker(a-abh^{-1}) \subseteq \ker(abh^{-1}-g)$. As a result, we have

$$\begin{aligned}
m_{gh} &\geq \dim\ker(ab-gh) \\
&\geq \dim\ker(abh^{-1}-g) \\
&\geq \dim(\ker(a-g) \cap \ker(a-abh^{-1})) \\
&\geq \dim(\ker(a-g)) + \dim(\ker(a-abh^{-1})) - n \\
&\geq \dim(\ker(a-g)) + \dim(\ker(b-h)) - n.
\end{aligned}$$

Since this is true for all $a, b \in F^\times$, therefore $m_g + m_h - n \leq m_{gh}$. So $\ell_J(gh) \leq \ell_J(g) + \ell_J(h)$. □

Lemma 3.6.4 (Stolz and Thom [ST13, Lemma 3.11])**.** *There is an absolute constant c_0, such that for any finite classical quasisimple group of Lie type G, and for any $g \in G \setminus Z(G)$, where $Z(G)$ is the center of G, then $C(g)^K = G$ for all $K \geq \frac{c}{\ell_J(g)}$.*

In short, elements of large Jordan length will automatically have small covering number. Now let us relate Jordan length function to the classical finite groups of characteristic 1, i.e., the symmetric and alternating groups. Recall that the matrix representation of these groups are precisely the permutation matrices.

Lemma 3.6.5. *Given an $n_1 \times n_1$ matrix A over a finite field F, and an $n_2 \times n_2$ matrix B over the same finite field, then $\ell_J(A \oplus B) \geq \frac{n_1}{n_1+n_2}\ell_J(A) + \frac{n_2}{n_1+n_2}\ell_J(B)$.*

Proof. For any $a \in F^\times$, we have the following

$$\ker(a - A \oplus B) = \ker((a - A) \oplus (a - B)) = \ker(a - A) \oplus \ker(a - B).$$

So $\dim\ker(a - A \oplus B) \leq m_A + m_B$. Since this is true for all $a \in F^\times$, therefore $m_{A\oplus B} \leq m_A + m_B$. So we have

$$\begin{aligned}\ell_J(A \oplus B) =& \frac{n_1 + n_2 - m_{A\oplus B}}{n_1 + n_2} \\ \geq& \frac{n_1 + n_2 - m_A - m_B}{n_1 + n_2} \\ \geq& \frac{n_1 - m_A}{n_1 + n_2} + \frac{n_2 - m_B}{n_1 + n_2} \\ \geq& \frac{n_1}{n_1 + n_2}\ell_J(A) + \frac{n_2}{n_1 + n_2}\ell_J(B).\end{aligned}$$

□

Lemma 3.6.6. *If P is an $n \times n$ permutation matrix over any finite field where its cycle decomposition has k cycles, then we have $\ell_J(P) \geq \frac{n-k}{n}$.*

Proof. By cycle decomposition, after a change of basis in the vector space, P will be a direct sum of many cyclic permutation matrices. By Lemma 3.6.5, it's enough to prove the case when P is a single cycle of length n, and show that $\ell_J(P) \geq \frac{n-1}{n}$.

Since P is a single cycle of length n, its eigenvalues in the algebraic closure of F are precisely all the n-th roots of unity, with multiplicity 1 for each root of unity. So $\dim\ker(a - P) \leq 1$ for all $a \in F^\times$. So $\ell_J(P) \geq \frac{n-1}{n}$. □

3.7 Classical finite groups of unbounded rank with Covering properties

The main idea here is to embed the alternating groups into the classical finite groups, and extend the conjugacy expansion of the alternating group to the classical finite group that contains it.

Proposition 3.7.1. *There is an absolute constant K_0 such that, for any $m < \infty$, for any finite quasisimple group of Lie type of $n \times n$ matrices, if it contains A_n as permutation matrices, then it will have the covering property (K_0, m) for large enough n.*

Proof. Let $K_0 > 3c_0$ for the absolute constant c_0 in Lemma 3.6.4. Then any element A of Jordan length $\geq \frac{1}{3}$ will have covering number K_0 in any finite quasisimple group of Lie type.

Pick any odd prime $p > m$, and pick another prime $q > p$. For any large enough n, we have $n = ap+bq$ for some integers $a > 1$, $0 < b < p+1$. Then find $\sigma \in \mathrm{A}_n$ made up of exactly a p-cycles and b q-cycles, where all cycles are disjoint. This element will be fixed-point free and non-exceptional, and it will have at most $a + b \leq \frac{n}{p} + p$ cycles.

For any finite quasisimple group of Lie type of $n \times n$ matrices, suppose it contains A_n as permutation matrices. Let P be the matrix corresponding to σ. Then we have

$$\ell_J(P) \geq \frac{n - \frac{n}{p} - p}{n} = 1 - \frac{1}{p} - \frac{p}{n} > \frac{1}{3}.$$

The last inequality follows because $p \geq 3$ and $n \geq 2p + q > 3p$.

So this element will have covering number K_0 in G. It clearly has order pq, and all of its powers coprime to pq will also have the same covering number. So G has the covering property $(K_0, p-1)$. □

Corollary 3.7.2. *For any $m < \infty$, all finite special linear groups of rank r for large enough r will have the covering property (K_0, m). Here K_0 is the absolute constant in Proposition 3.7.1.*

Proposition 3.7.3. *There is an absolute constant K_0, such that for any $m < \infty$, we have the following:*

1. *For any finite quasisimple group of Lie type of $2n \times 2n$ matrices, if it contains A_n as $\{P \oplus P : P \in \mathrm{A}_n$ is a permutation $n \times n$ matrix$\}$, then it will have the covering property (K_0, m) for large enough n.*

2. *Let I_1 be the 1 by 1 identity matrix. Then for any finite quasisimple group of Lie type of $(2n+1) \times (2n+1)$ matrices, if it contains A_n as $\{P \oplus P \oplus I_1 : P \in \mathrm{A}_n$ is a permutation $n \times n$ matrix$\}$, then it will have the covering property (K_0, m) for large enough n.*

3. *Let I_2 be the 2 by 2 identity matrix. Then for any finite quasisimple group of Lie type of $(2n+2) \times (2n+2)$ matrices, if it contains A_n as $\{P \oplus P \oplus I_2 : P \in \mathrm{A}_n$ is a permutation $n \times n$ matrix$\}$, then it will have the covering property (K_0, m) for large enough n.*

Proof. The strategy is identical to Proposition 3.7.1. Just take $\sigma \oplus \sigma$, $\sigma \oplus \sigma \oplus I_1$ or $\sigma \oplus \sigma \oplus I_2$ instead of σ, and use Lemma 3.6.5. □

Definition 3.7.4. A vector space V is a ***non-degenerate formed space*** if it has a non-degenerate quadratic form Q (the orthogonal case), or a non-degenerate alternating bilinear form B (the symplectic case), or a non-degenerate Hermitian form B (the unitary case).

Lemma 3.7.5 (Witt's Decomposition Theorem)**.** *Let V be any non-degenerate formed space over a finite field F. Then we have an orthogonal decomposition $V = W \oplus (\bigoplus_{i=1}^{n} H_i)$ where W is anisotropic of dimension at most 2, and H_i are hyperbolic planes.*

Proof. These are standard facts in the geometry of classical groups (e.g., See [Gro02]). □

Proposition 3.7.6. *For a non-degenerate formed space, the special isometry group, i.e., the group of isometries of determinant 1, contains an alternating group in one of the ways described by Proposition 3.7.3.*

Proof. Let V be any finite dimensional non-degenerate formed space over any finite field F. Then we have an orthogonal decomposition $V = W \oplus H$ with an anisotropic space W of dimension at most 2, and an orthogonal sum of hyperbolic planes $H = \bigoplus_{i=1}^{n} H_i$.

Then let (v_i, w_i) be a hyperbolic pair generating H_i for each i. For any $\sigma \in \mathrm{A}_n$, we can let σ act by permutation on the set $\{v_1, .., v_n, w_1, ..., w_n\}$, such that $\sigma(v_i) = v_{\sigma(i)}$ and $\sigma(w_i) = w_{\sigma(i)}$.

Now clearly $\{v_1, ..., v_n, w_1, .., w_n\}$ is a basis of H. So the above action of σ induces a linear transformation $P \oplus P$ on H, where P is the $n \times n$ permutation matrix for σ. And this $P \oplus P$ is clearly an isometry on H by construction. Now taking the direct sum of $P \oplus P$ on H and the identity matrix on W, we shall obtain our desired embedding of A_n into the full isometry group.

Finally, since P is a permutation matrix for an even permutation, it has determinant 1. Therefore the above embedding of A_n is in the special isometry group. □

Corollary 3.7.7. *For any $m < \infty$, any finite symplectic or special unitary group of rank r has the covering property (K_0, m) for large enough r. K_0 is the absolute constant in Proposition 3.7.3.*

Corollary 3.7.8. *For any $m < \infty$, any $\Omega^+_{2n}(\mathbb{F}_q)$, $\Omega_{2n+1}(\mathbb{F}_q)$ or $\Omega^-_{2n}(\mathbb{F}_q)$ has the covering property (K_0, m) for large enough n. K_0 is the absolute constant in Proposition 3.7.3.*

Proof. Embed A_n in $\mathrm{SO}^+_{2n}(q)$, $\mathrm{SO}^-_{2n}(q)$ and $\mathrm{SO}_{2n+1}(q)$ in the ways described by Proposition 3.7.3. After taking the commutator subgroup, the groups $\Omega^+_{2n}(q)$, $\Omega^-_{2n}(q)$ and $\Omega_{2n+1}(q)$ will still contain A_n through this embedding, because A_n is its own commutator subgroup. So we may apply Proposition 3.7.3 to $\Omega^+_{2n}(q)$, $\Omega^-_{2n}(q)$ and $\Omega_{2n+1}(q)$ and obtain the desired result. □

3.8 Combining Covering properties in a uniform way

We have show that for all non-abelian finite quasisimple group, if their order or Lie rank is large enough, then they will have good enough covering property. It is very tempting to combine everything to obtain a covering property for all non-abelian finite quasisimple group of large enough order.

It turns out that this is too optimistic. To have a covering property (K, m), a finite simple group of Lie type must either have a large enough rank to accommodate the large m, according to Corollary 3.7.2, 3.7.7, and 3.7.8, or it must have a small enough rank to accomodate the small K, according to Corollary 3.5.4. So there might be a gap between the "large enough rank" and the "small enough rank", where the finite simple subgroups in the gap would fail to have the covering property (K, m), no matter how large their order are.

In short, the covering properties of finite quasisimple groups are not necessarily uniform. It is uniform when obtained through increasing ranks, and it is uniform when obtained through base fields of increasing sizes. At least with the techniques in this paper, we cannot combine the two uniformity into one. So we must use the double covering properties.

Lemma 3.8.1. *For any integer D and any constant c, we can find integers K_1, K_2, m_1, m_2 such that all finite quasisimple groups with large enough order will have the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$ such that $m_1 > cK_1^{D^2}$, $m_2 > cK^{D^2}$.*

Proof. Let K_1 be $\max(16, K_0)$ where the absolute constant K_0 is as in Proposition 3.7.1 and Proposition 3.7.3. Pick some $m_1 > cK_1^{D^2}$. Find large enough r such that, according to Corollary 3.7.2, 3.7.7, 3.7.8 and Proposition 3.4.4, all finite quasisimple groups of Lie type of ranks $\geq r$ and all alternating groups with large enough order will have the symmetric covering property (K_1, m_1).

Set $K_2 := K(r) \max(3r + 1, 10)$ as in Corollary 3.5.4, and pick some $m_2 > cK_2^{D^2}$. Then all finite quasisimple groups of Lie type of ranks $\leq r$ and with large enough order will have the symmetric covering property (K_2, m_2).

In all cases, a finite quasisimple group with large enough order will have the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$. $\square$

CHAPTER 4

Applications of Covering Properties

4.1 Ultraproducts of Quasirandom groups

As we have seen in the last chapter, groups with good covering properties no representation of small degree. Such a group is called a quasirandom group.

Definition 4.1.1. A group G is D-***quasirandom*** if it has no non-trivial finite dimensional unitary representation of dimension less than D. The largest of such D is called the ***quasirandom degree*** of G.

Quasirandom groups are first introduced by Gowers to find groups with no large product-free subset. They can be seen as stronger versions of perfect groups.

Example 4.1.2 (Gowers [Gow08])**.**

1. *A group (not necessarily finite) is 2-quasirandom iff it is perfect, i.e., it is its own commutator subgroup. The reason is that a non-perfect group has a non-trivial abelian quotient, which in turn has a non-trivial homomorphism into* $\mathrm{U}_1(\mathbb{C})$*. A perfect group, on the other hand, can only have the trivial homomorphism into the abelian group* $\mathrm{U}_1(\mathbb{C})$*.*

2. *A finite perfect group with no normal subgroup of index less than n is at least $\sqrt{\log n}/2$-quasirandom. In fact, using a form of Jordan's theorem [**?**], a finite perfect group with no normal subgroup of index less than n is at least $c \log n$-quasirandom for some constant c.*

3. *In particular, a non-abelian finite simple group G is at least $c \log n$-quasirandom if it has n elements.*

4. *Conversely, any D-quasirandom group must have more than $(D-1)^2$ elements.*

5. *The alternating group A_n is $(n-1)$-quasirandom for $n > 5$, and the special linear group $\mathrm{SL}_2(F_p)$ is $\frac{p-1}{2}$-quasirandom for any prime p.*

Suppose we have a sequence of quasirandom groups with increasing quasirandom degree. Then a natural question to ask is that if we have some sort of a "limit group" for this sequence, then would the limit group be "infinitely quasirandom"? I.e., maybe the limit group will have no non-trivial finite dimensional unitary representation at all. This property turned out to be related to almost periodic functions of this limit group, which we shall not get into.

Definition 4.1.3. A group G is ***minimally almost periodic*** if it has no non-trivial finite dimensional unitary representations.

The limit that we shall consider is the ultraproduct. Let us briefly describe it here.

Definition 4.1.4. A ***filter*** on $\mathbb{N}$ is a collection ω of subsets of $\mathbb{N}$ such that:

1. $\varnothing \notin \omega$;

2. If $X \in \omega$ and $X \subseteq Y$, then $Y \in \omega$;

3. If $X, Y \in \omega$, then $X \cap Y \in \omega$.

An ***ultrafilter*** is a filter that is maximal with respect to the containment order. A ***non-principal ultrafilter*** is an ultrafilter that contains no finite subset of $\mathbb{N}$.

Definition 4.1.5. Given a sequence of groups $(G_i)_{i\in\mathbb{N}}$, let G be their direct product. Given an ultrafilter ω on $\mathbb{N}$, let $N := \{g = (g_i)_{i\in\mathbb{N}} \in G : \{i \in \mathbb{N} : g_i = e\} \in \omega\}$, which is clearly a normal subgroup of G. Then we call G/N the ***ultraproduct*** of the groups $(G_i)_{i\in\mathbb{N}}$ by ω, denoted by $\prod_{i\to\omega} G_i$.

Remark 4.1.6. *An ultrafilter ω is principal (i.e., not non-principal) iff we can find an element $n \in \mathbb{N}$ such that for all subsets $A \subseteq \mathbb{N}$, we have $A \in \omega$ iff $n \in A$. In this case, the* *corresponding ultraproduct of groups $(G_i)_{i\in\mathbb{N}}$ is isomorphic to G_i. Therefore, in practice, the useful ultrafilters are usually non-principal.*

The particular choice of the ultrafilter is not that important. As long as we fix a non-principal ultrafilter, then all the discussion for the rest of the chapter will be true for the ultraproduct of this ultrafilter.

Ultraproducts have an interesting property, given by Łoś' Theorem [Los55]. Given an ultraproduct $G = \prod_{i\to\omega} G_i$ for an ultrafilter ω, any first-order statement ϕ in the language of groups is true for G iff it is true for most of the G_i, i.e., $\{i \in \mathbb{N} : \phi \text{ is true for } G_i\} \in \omega$. In particular, this implies that behaviors at the scale of elements are preserved. We shall not need Łoś' Theorem in this dissertation.

Now, it is very tempting to claim that an ultraproduct of a sequence of groups with increasing quasirandom degree is minimally almost periodic. Unfortunately this is not the case.

Example 4.1.7. *We recall that a group G (not necessarily finite) is 2-quasirandom iff G is perfect. We claim that there is a sequence of D_i-quasirandom groups $(G_i)_{i\in\mathbb{Z}^+}$ with $\lim_{i\to\infty} D_i = \infty$, whose ultraproduct by any non-principal ultrafilter is not even perfect.*

Using the construction of Holt and Plesken [HP89, Lemma 2.1.10], one may construct a finite perfect group $G_{p,n}$ for each prime $p \geq 5$ and positive integer n, such that an element of $G_{p,n}$ cannot be written as a product of less than n commutators, and that the only simple quotient of $G_{p,n}$ is $\mathrm{PSL}_2(\mathbb{F}_p)$, the projective special linear group of 2×2 matrices over the field of p elements. Then by Example 4.1.2 (ii), for any D, $G_{p,n}$ is D-quasirandom for large enough p.

Let G_i be $G_{p_i,i}$, where $(p_i)_{i\in\mathbb{Z}^+}$ is a strictly increasing sequence of primes. Then G_i is D_i-quasirandom for some D_i with $\lim_{i\to\infty} D_i = \infty$. Let $g_i \in G_i$ be an element which cannot be written as a product of less than i commutators. Then $g = (g_i)_{i\in\mathbb{N}}$ corresponds to an element of the ultraproduct $G = \prod_{i\to\omega} G_i$ by any ultrafilter ω. When ω is non-principal, clearly g cannot be written as a product of finite number of commutators in G. So g is not in the commutator subgroup of G, and thus G is not perfect.

However, a recent paper by Bergelson and Tao [BT14] showed the following theorem, which shed some new light on this inquiry:

Theorem 4.1.8 (Bergelson and Tao [BT14, Theorem 49 (i)])**.** *The ultraproduct* $\prod_{i\to\omega} \mathrm{SL}_2(\mathbb{F}_{p_i})$ *by a non-principal ultrafilter* ω *is minimally almost periodic.*

Inspired by this, we can make the following definitions:

Definition 4.1.9. A class $\mathcal{F}$ of groups is a ***q.u.p. (quasirandom ultraproduct property) class*** if for any sequence of groups in $\mathcal{F}$ with quasirandom degree going to infinity, their non-principal ultraproducts will be minimally almost periodic.

Definition 4.1.10. A class $\mathcal{F}$ of groups is a ***Q.U.P. class*** if there is an unbounded non-decreasing function $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that any ultraproduct of any sequence of D-quasirandom groups in $\mathcal{F}$ is $f(D)$-quasirandom.

Remark 4.1.11. *A Q.U.P class is automatically a q.u.p. class. It is like an effective version of q.u.p. class, where we are able to keep track of the amount of quasirandomness passed down to the ultraproduct.*

Now, let $\mathcal{C}_n$ be the smallest class of groups that contains all finite quaisisimple groups and all finite groups with at most n conjugacy classes in its cosocle, and closed under arbitrary direct products (not necessarily finite). I claim that the following is true:

Theorem 4.1.12. *For any sequence of groups in* $\mathcal{C}_n$ *with quasirandom degree going to infinity, their non-principal ultraproducts will be minimally almost periodic. In fact, this class is a Q.U.P. class.*

Corollary 4.1.13. *An ultraproduct of finite simple groups is either finite or minimally almost periodic.*

One should note that the possiblity of taking an infinite direct product and arbirary quotient means that $\mathcal{C}_n$ includes many infinite groups as well.

Proposition 4.1.14. *Let G be a group with the symmetric double covering property for some parameters, and let $(G_i)_{i\in I}$ be an arbitrary family of groups with the symmetric double covering property for some uniform parameters. Then the following are true:*

(i) For any normal subgroup N, G has the symmetric double covering property for the same parameters mod N.

(ii) Any quotient group of G has the symmetric double covering property for the same parameters.

(iii) The group $\prod_{i\in I} G_i$ has the symmetric double covering property for the same parameters.

(iv) As a result of the (ii) and (iii), any ultraproduct $\prod_{i\to\omega} G_i$ has the symmetric double covering property for the same parameters.

Proof. (i), (ii) and (iv) are straightforward.

To see (iii), let $g_{i,1}, g_{i,2} \in G_i$ be the pairs giving G_i the symmetric double covering property. Then I claim that $(g_{i,1})_{i\in I}, (g_{i,2})_{i\in I} \in \prod_{i\in I} G_i$ is the pair giving the desired symmetric double covering property.

For any element $(g_i)_{i\in I} \in \prod_{i\in I} G_i$, then each g_i is in G_i. And by its symmetric double covering property, we know

$$G_i = C(g_{i,1})^{K_1} C(g_{i,1}^{-1})^{K_1} C(g_{i,2})^{K_2} C(g_{i,2}^{-1})^{K_2}.$$

So we can find $a_{i,j}, b_{i,j} \in G_i$ for $i \in I$ and $1 \le j \le K_1$, and $c_{i,j}, d_{i,j} \in G_i$ for $i \in I$ and $1 \le j \le K_2$, such that

$$g_i = (\prod_{1\le j\le K_1} (a_{i,j} g_{i,1} a_{i,j}^{-1})(b_{i,j} g_{i,1}^{-1} b_{i,j}^{-1}))(\prod_{1\le j\le K_2} (c_{i,j} g_{i,2} c_{i,j}^{-1})(d_{i,j} (g_{i,2})^{-1} d_{i,j}^{-1})).$$

Since the above identity is true for all $i \in I$, we have

$$\begin{aligned}(g_i)_{i\in I} =& (\prod_{1\le j\le K_1} ((a_{i,j})_{i\in I}(g_{i,1})_{i\in I}(a_{i,j})_{i\in I}^{-1})((b_{i,j})_{i\in I}(g_{i,1})_{i\in I}^{-1}(b_{i,j})_{i\in I}^{-1}))\\ &(\prod_{1\le j\le K_2} ((c_{i,j})_{i\in I}(g_{i,2})_{i\in I}(c_{i,j})_{i\in I}^{-1})((d_{i,j})_{i\in I}(g_{i,2})_{i\in I}^{-1}(d_{i,j})_{i\in I}^{-1})).\end{aligned}$$

So we have proven (iii). □

Note that for finite quasisimple groups, to have a large enough order is the same as to have a large enough quasirandom degree. So when we apply results from Chapter 3, we can replace all "large enough order" requirements by a "large enough quasirandom degree" requirement.

Corollary 4.1.15. *Let $\mathcal{C}_{QS}$ be the class of finite quasisimple groups. Then $\mathcal{C}_{QS}$ is a Q.U.P. class.*

Proof. For any integer D, and for the constant $c = f(D)$ as in Proposition 3.2.4, we can find D', K_1, K_2, m_1, m_2 as in Lemma 3.8.1, where a group with quasirandom degree D' will have large enough power for us to use Lemma 3.8.1.

Let G_i be a sequence of D'-quasirandom groups in $\mathcal{C}_{QS}$. Then G_i all have the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$. Then any ultraproduct $G = \prod_{i\to\omega} G_i$ will have the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$ by Proposition 4.1.14. Since $m_1 > f(D)K_1^{D^2}$, $m_2 > f(D)K^{D^2}$, G is D-quasirandom by Proposition 3.2.4. □

Corollary 4.1.16 (Quasirandomness implies a Nice Covering Property mod Cosocle)**.** *For any integer D, and any constant c, we can find integers D', K_1, K_2, m_1, m_2 such that all finite D'-quasirandom groups have the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$ mod cosocle, with $m_1 > cK_1^{D^2}$, $m_2 > cK_2^{D^2}$.*

Proof. Let D', K_1, K_2, m_1, m_2 be exactly as in Lemma 3.8.1, where a group with quasirandom degree D' will have large enough power for us to use Lemma 3.8.1. Let G be any finite D'-quasirandom group.

Let N be the cosocle of G. Then G/N is a direct product of D'-quasirandom finite simple groups. These simple groups all have the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$. So by Proposition 4.1.14, their product G/N will have this same symmetric double covering property. □

Corollary 4.1.17. *Let $\mathcal{C}_{CS(n)}$ be the class of finite groups with at most n conjugacy classes in their cosocles. Then $\mathcal{C}_{CS(n)}$ is a Q.U.P. class.*

Proof. Let $c = f(D)(3n-2)^{D^2}$.

For any integer D, and for the constant c, we can find D', K_1, K_2, m_1, m_2 as in Corollary 4.1.16.

Let G_i be a sequence of D'-quasirandom groups in $\mathcal{C}_{CS(n)}$. Then G_i all have the symmetric double covering property $[(K_1, m_1), (K_2, m_2)]$ mod cosocles. Since the cosocles contain at most n conjugacy classes, by Proposition 3.3.3, G_i all have the symmetric double covering property $[((3n-2)K_1, m_1), ((3n-2)K_2, m_2)]$. Then any ultraproduct $G = \prod_{i\to\omega} G_i$ will have the symmetric double covering property $[((3n-2)K_1, m_1), ((3n-2)K_2, m_2)]$ by Proposition 4.1.14.

Since $m_1 > f(D)[(3n-2)K_1]^{D^2}$, $m_2 > f(D)[(3n-2)K]^{D^2}$, G is D-quasirandom by Proposition 3.2.4. □

Proof of Theorem 4.1.12. For any integer D, let $c = f(D)(3n-2)^{D^2}$. As usual, we can find D', K_1, K_2, m_1, m_2 as in Corollary 4.1.16 and Lemma 3.8.1.

Let G_i be a sequence of D'-quasirandom groups in $\mathcal{C}_n$. Then each G_i is a direct product of D'-quasirandom groups in $\mathcal{C}_{QS} \cup \mathcal{C}_{CS(n)}$. These factor groups must then have the symmetric double covering property $[((3n-2)K_1, m_1), ((3n-2)K_2, m_2)]$. By Proposition 4.1.14, G_i must also have this symmetric double covering property $[((3n-2)K_1, m_1), ((3n-2)K_2, m_2)]$. Then any ultraproduct $G = \prod_{i\to\omega} G_i$ will have the symmetric double covering property $[((3n-2)K_1, m_1), ((3n-2)K_2, m_2)]$ by Proposition 4.1.14.

Since $m_1 > f(D)[(3n-2)K_1]^{D^2}$, $m_2 > f(D)[(3n-2)K]^{D^2}$, Proposition 3.2.4 implies that G is D-quasirandom. □

4.2 Self-Bohrifying groups

The application in this section is related to topological groups. We shall treat all groups in previous sections as discrete groups.

Definition 4.2.1. A ***Bohr compactification*** of a topological group G is a continuous homomorphism $b : G \to bG$ such that any continuous homomorphism from G to a compact group factors uniquely through b.

Remark 4.2.2.

1. *The Bohr compactification exists for any group by the work of Holm [Hol64]. It is obviously unique up to a unique isomorphism.*

2. *Clearly, a discrete group is minimally almost periodic iff it has trivial Bohr compactification. Note that for a discrete group, any abstract homomorphism from it to another topological group is automatically continuous.*

Definition 4.2.3. A topological group G is said to be ***self-Bohrifying*** if its Bohr compactification bG is the same abstract group as G, but with a compact topology.

By the results and techniques of this paper, one can find many examples of self-Bohrifying groups. In particular, we have the following theorem.

Theorem 4.2.4. *Let n be a positive integer. Let G_i be a sequence of increasingly quasirandom groups in $\mathcal{C}_n$, the class defined as in Theorem 4.1.12. Then $\prod_{i\in\mathbb{N}} G_i$ is self-Bohrifying as a discrete group.*

Corollary 4.2.5. *Let G_i be a sequence of non-abelian finite simple groups of increasing order. Then $\prod_{i\in\mathbb{N}} G_i$ is self-Bohrifying as a discrete group.*

We will prove Theorem 4.2.4 by first showing that $\prod_{i\in\mathbb{N}} G_i/\coprod_{i\in\mathbb{N}} G_i$ is minimally almost periodic, and then using a lemma by Hart and Kunen [HK02].

Definition 4.2.6. Let G_i be a sequence of groups.

1. Their ***sum*** is the group $\coprod_{i\in\mathbb{N}} G_i = \{g \in \prod_{i\in\mathbb{N}} G_i :$ only finitely many coordinates of g is nontrivial$\}$.

2. Their ***reduced product*** is the group $\prod_{i\in\mathbb{N}} G_i/\coprod_{i\in\mathbb{N}} G_i$.

Lemma 4.2.7 (Hart and Kunen [HK02, Lemma 3.8])**.** *Let $\{G_i\}_{i\in\mathbb{N}}$ be a sequence of finite groups. Then $\prod_{i\in\mathbb{N}} G_i$ is self-Bohrifying if all but finitely many G_i are perfect groups,* *and $\prod_{i\in\mathbb{N}} G_i / \coprod_{i\in\mathbb{N}} G_i$ has trivial Bohr compactification, i.e., $\prod_{i\in\mathbb{N}} G_i / \coprod_{i\in\mathbb{N}} G_i$ is minimally almost periodic.*

Proof of Theorem 4.2.4. All 2-quasirandom groups are perfect. So it is enough to show that the reduced product of G_i is minimally almost periodic, i.e., it is D-quasirandom for all D.

For any integer D, let $c = f(D)(3n-2)^{D^2}$. We can find D', K_1, K_2, m_1, m_2 as in Corollary 4.1.16 and Lemma 3.8.1.

Let G_i be a sequence of increasingly quasirandom groups in $\mathcal{C}_n$. Then all but finitely many G_i will be D'-quasirandom. Since we are interested in the reduced product, which is invariant under the change of finitely many coordinates, we may WLOG assume that all G_i are D'-quasirandom.

Since $G_i \in \mathcal{C}_n$, each G_i is a direct product of D'-quasirandom groups in $\mathcal{C}_{QS} \cup \mathcal{C}_{CS(n)}$. These factor groups must then have the symmetric double covering property $[((3n-2)K_1, m_1), ((3n-2)K_2, m_2)]$. By Proposition 4.1.14, G_i must also have this symmetric double covering property $[((3n-2)K_1, m_1), ((3n-2)K_2, m_2)]$.

Now by Proposition 4.1.14, covering properties are preserved by arbitrary products and quotients. So $\prod_{i\in\mathbb{N}} G_i$ will have this covering property, and the reduced product $\prod_{i\in\mathbb{N}} G_i / \coprod_{i\in\mathbb{N}} G_i$ will also have this covering property.

Since $m_1 > c[(3n-2)K_1]^{D^2}$, $m_2 > c[(3n-2)K]^{D^2}$, the reduced product is D-quasirandom by Proposition 3.2.4. So we are done by Lemma 4.2.7. $\square$

CHAPTER 5

Schreier graphs of non-abelian finite simple groups

Starting from this chapter, we start our discussion of Cayley graphs of non-abelian finite simple groups. It turns out that most of the time, it is much easier to study a "quotient graph" of the Cayley graph, i.e., the Shreier graphs. This amounts to the study of group actions as a tool to the study of groups.

We shall mainly consider two types of Schreier graphs of a group G, one is the Schreier graph for the canonical action or its variants if G is a linear group or an alternating group. The other is the Schreier graph of the action of G on its conjugacy classes. In this chapter, we shall merely establish some crude bounds on the diameters of these Schreier graphs.

5.1 t-transitive subsets of alternating groups

In this section, we establish some very basic results on transitive subsets of symmetric and alternating groups. This serves as a motivation for t-transversal subsets of linear groups in later sections.

Let G be the symmetric group S_n or the alternating group A_n. We let G act on the set $X = \{1, 2, ..., n\}$ as usual.

Definition 5.1.1. A subset T of G is ***t-transitive*** if given any injective function f from any t-element subset Y of X, then there is a permutation in T that extends f.

Lemma 5.1.2. S_n *is t-transitive for all t, and* A_n *is t-transitive for all $t \leq n-2$.*

Proof. This is elementary. $\square$

Lemma 5.1.3. *For any symmetric subset S of S_n, if the subgroup generated by S is t-transitive, then $\cup_{d=1}^{n^t} S^d$ is t-transitive.*

Proof. Let Y be ay t-element subset of X. Let $L(Y)$ be the set of all injective functions from Y to X, and let H be the subgroup generated by S. Then an element g of H acts on $L(Y)$ by sending each function f to $g \circ f$. Let Γ be the corresponding Schreier graph for this action of H with generating set S.

Now, since H is t-transitive, Γ is connected. So the diameter of Γ is trivially bounded by its number of vertices, which is n^t. So $\cup_{d=1}^{n^t} S^d$ can extend any injective function in $L(Y)$. □

5.2 t-Transversal Sets for Special Linear Groups

Let V be a vector space of dimension n over the field $\mathbb{F}_q$. Let the group $\mathrm{GL}_n(\mathbb{F}_q)$ act on it naturally.

Definition 5.2.1. A subset S of $\mathrm{GL}_n(\mathbb{F}_q)$ is called a ***t-transversal set*** if given any embedding X of a t-dimensional subspace W into V, we can find $A \in S$ that extends X on W.

Lemma 5.2.2. $\mathrm{GL}_n(\mathbb{F}_q)$ *is t-transversal for all t, and* $\mathrm{SL}_n(\mathbb{F}_q)$ *is t-transversal for all $t < n$.*

Proof. Let W be any subspace with a basis $w_1, ..., w_t$. We can complete this into a basis of V with new vectors $v_1, ..., v_{n-t}$. Let A be a matrix with column vectors $w_1, ..., w_t, v_1, ..., v_{n-t}$. In the case when $t < n$, we can multiply v_{n-t} by a constant so that $\det(A) = 1$.

For any embedding X of W into V, $X(w_1), ..., X(w_t)$ are linearly independent. We can complete this into a basis of V with new vectors $u_1, ..., u_{n-t}$. Let B be a matrix with column vectors $X(w_1), ..., X(w_t), u_1, ..., u_{n-t}$. In the case when $t < n$, we can multiply u_{n-t} by a constant so that $\det(B) = 1$.

Now $B(A)^{-1}$ is in $\mathrm{GL}_n(\mathbb{F}_q)$ and, if $t < n$, also in $\mathrm{SL}_n(\mathbb{F}_q)$. We also have $(B(A)^{-1})|_W = X$. □

Lemma 5.2.3. *For any symmetric subset S of $\mathrm{GL}_n(\mathbb{F}_q)$, if the subgroup generated by S is t-transversal, then $\bigcup_{d=1}^{d=q^{nt}} S^d$ is t-transversal.*

Proof. Let W be any t-dimensional subspace. Let $L(W)$ be the set of embeddings of W into V. Let H be the subgroup generated by S. Then an element g of H acts on $L(W)$ by $g(X) = (g \circ X)|_W$ for any $X \in L(W)$. Let Γ be the corresponding Schreier graph of this action of H on $L(W)$ with generating set S, i.e., the vertices are elements of $L(W)$, and two vertices X, Y are connected iff $g(X) = Y$ for some $g \in S$.

Now, since H is t-transversal, the graph Γ is connected. So the diameter of Γ is trivially bounded by its number of vertices, which is at most q^{nt}. As a result, the set $\bigcup_{d=1}^{q^{nt}} S^d$ is t-transversal. □

Corollary 5.2.4. *Given any symmetric generating set S for $\mathrm{GL}_n(\mathbb{F}_q)$, the set $\bigcup_{d=1}^{d=q^{nt}} S^d$ is t-transversal. If $t < n$, then the same statement is true with $\mathrm{SL}_n(\mathbb{F}_q)$ replacing $\mathrm{GL}_n(\mathbb{F}_q)$.*

5.3 t-Transversal Sets for Orthogonal Groups, Symplectic Groups, and Unitary Groups

Let's fix some notation for the discussion of the following three sections. Let V be a non-degenerate formed space of dimension n over the field $\mathbb{F}_q$, with a non-degenerate quadratic form Q (the orthogonal case), non-degenerate alternating bilinear form B (the symplectic case), or non-degenerate Hermitian form B with field automorphism σ (the unitary case). In the orthogonal case, we shall let B be the symmetric bilinear form obtained by polarizing Q, i.e., $B(v, w) = Q(v + w) - Q(v) - Q(w)$. Let G be the group of isometries for V.

Definition 5.3.1.

1. A vector $v \in V$ is ***singular*** if $B(v, v) = 0$ and (if applicable) $Q(v) = 0$.
2. A pair of singular vectors $v, w \in V$ is called a ***hyperbolic pair*** if $B(v, w) = 1$.
3. The subspace generated by a hyperbolic pair is a ***hyperbolic plane***.
4. A subspace W of V is ***anisotropic*** if it contains no singular vector other than 0.

5. A subspace is ***totally singular*** if the form B and (if applicable) the quadratic form Q restricted to it is the zero form.

6. Given any subspace W of V, we define its ***orthogonal complement*** to be $W^{\perp} := \{v \in V : B(v, w) = 0 \text{ for all } w \in W\}$. Two subspaces U, W of V are ***orthogonal*** if they are in each other's orthogonal complimant. We denote this as $U \perp W$.

7. The ***radical*** of V is $V^{\perp}$.

8. A subspace W is ***radical-free*** if $W \cap V^{\perp} = \{0\}$.

Theorem 5.3.2 ((Witt's Decomposition Theorem))**.** *The non-degenerate formed space V has an orthogonal decomposition $V = V_{ani} \oplus (\bigoplus_{i=1}^{m} H_i)$, where V_{ani} is anisotropic of dimension at most 2, and H_i are hyperbolic planes. In particular, V has a totally singular subspace of dimension at least $\frac{\dim(V)-2}{2}$, and any anisotropic space in V has dimension at most 2.*

Proof. See [Gro02]. $\square$

Lemma 5.3.3. *Recall that V is a non-degenerate formed space.*

1. $V^{\perp} = \{0\}$ *unless the non-degenerate form for V is a quadratic form, and* $\operatorname{char} \mathbb{F}_q = 2$.

2. $V^{\perp}$ *has dimension at most 1.*

3. *For any subspace W,* $\dim W + \dim W^{\perp}$ *is equal to* $\dim V$ *if W is radical-free, and* $\dim V + 1$ *if W is not.*

4. *For any subspace W, $(W^{\perp})^{\perp} = W + V^{\perp}$.*

5. *A totally singular subspace is always radical-free.*

Proof. See [Gro02]. $\square$

Definition 5.3.4. A subset S of G is called a ***singularly t-transversal set*** if, for any isometric embedding X of a t-dimensional totally singular subspace W into V, we can find $A \in S$ that extends X on W.

Lemma 5.3.5 ((Witt's Extension Lemma)). *G is a singularly t-transversal set for any t.*

Proof. This is a special case of Witt's extension lemma, which states that any bijective isometry of radical-free subspaces of V could be extended to an isometry of the whole formed space. See [Gro02] for a proof. $\square$

Now, since our focus is on the finite simple groups, we don't really use the full isometry group G. Rather, we are interested in its commutator subgroup G'.

Lemma 5.3.6. *For any $t \leq \frac{n-2}{5}$, the commutator subgroup G' of G is singularly t-transversal.*

Proof. Let W be a totally singular space of dimension t. Let $X : W \to V$ be any isometric embedding from W to V.

Step 1: I claim that there is a totally singular subspace W', which is orthogonal to W and $X(W)$, has trivial intersection with W and $X(W)$, and has the same dimension as W.

To see this, we have $\dim W^\perp = \dim X(W)^\perp \geq n - t$. Therefore, $\dim(W^\perp \cap X(W)^\perp) \geq n - 2t$. So in the subspace $W^\perp \cap X(W)^\perp$, we can find a subspace W'' of dimension $n - 3t$ with trivial intersections with W and $X(W)$. Now, since W'' is a formed space (possibly degenerate), it has a totally singular subspace of dimension at least $\frac{\dim W''-2}{2} = \frac{n-3t-2}{2} \geq t$. So, from this totally singular space, we could simply pick any totally singular subspace of dimension t to be the desired W'.

Step 2:

Let $Y : W \to W'$ be any bijective linear map. Since both spaces are totally singular, Y is an isometry. So we could find an extension $A \in G$.

Let $Z : W \oplus W' \to X(W) \oplus W'$ be the linear map that restricts to X on W, and restricts to the identity map on W'. Then by our choice of W', this is a well-defined isometry of totally singular subspaces, and it would have an extension $B \in G$.

Consider $BA^{-1}B^{-1}A \in G'$. This would restrict to X on W. So we are done. $\square$

Lemma 5.3.7. *Let S be any subset of G. If the subgroup generated by S is singularly t-transversal, then $\bigcup_{d=1}^{d=q^{nt}} S^d$ is singularly t-transversal.*

Proof. Let H be the subgroup generated by S. Let W be any t-dimensional totally singular subspace, and let $L(W)$ be the set of isometric embeddings of W into V. Then an element $g \in H$ acts on $L(W)$ by $g(X) = (g \circ X)|_W$ for any $X \in L(W)$. Let Γ be the corresponding Schreier graph of this action of H on $L(W)$ with generating set S.

Any isometric embedding from W to V is a linear map. Therefore, there are at most q^{nt} vertices for Γ, where $t = \dim W$. And since H is singularly t-transversal, the graph Γ must be connected. So Γ must have a diameter at most q^{nt}. □

Corollary 5.3.8. *Given any symmetric generating set S for G or G', the set $\bigcup_{d=1}^{d=q^{nt}} S^d$ is singularly t-transversal for $t \leq \frac{n-2}{5}$.*

5.4 The Conjugacy Expansion Lemmas

In this section, we study Schreier graphs of groups acting on their conjugacy classes. As a result, we shall show that any small degree element will quickly generate the whole group with any symmetric generating set.

Definition 5.4.1. The ***degree*** of a square matrix A is defined to be the rank of $A - I$ where I is the identity matrix with the same dimension as A.

Definition 5.4.2. Give a group G and a symmetric generating set S, then the ***length*** of an element g of G is defined to be the smallest number ℓ such that $g = s_1 s_2 ... s_\ell$ for some $s_1, ..., s_\ell \in S$.

Lemma 5.4.3. *Let S be any symmetric generating set for a subgroup H of $\mathrm{GL}_n(\mathbb{F}_q)$. Let A be any matrix in H of degree k, and let B be any matrix conjugate to A in H. Then $B = MAM^{-1}$ for some $M \in H$ of length at most q^{2nk}.*

Proof. Since A has degree k, we know $A = I + A'$ for some matrix A' of rank k. So we can decompose A' as a product XY where X is an n by k matrix of full rank and Y is a k by n matrix of full rank. So $A = I + XY$.

Any conjugates of A can similarly be expressed as $I + X'Y'$ where X' is some n by k matrix of full rank, and Y' is some k by n matrix of full rank. There are at most q^{2nk} possibilities for the pair (X', Y'). So there are at most q^{2nk} conjugates of A.

Now H acts on the conjugacy class of A in H by left conjugation, and the corresponding Schreier graph must be connected. So the Schreier graph has diameter bounded by the number of vertices, i.e., q^{2nk}. □

Theorem 5.4.4 (([LS01])). *Let G be $\mathrm{SL}_n(\mathbb{F}_q)$, $\Omega_n(\mathbb{F}_q)$, $\mathrm{Sp}_n(\mathbb{F}_q)$, or $\mathrm{SU}_n(\mathbb{F}_q)$. Let $A \in G$ be an element of degree k outside the center of G. Then every element of G is a product of at most $O(\frac{n}{k})$ conjugates of A.*

Proposition 5.4.5. *Let G be $\mathrm{SL}_n(\mathbb{F}_q)$, $\Omega_n(\mathbb{F}_q)$, $\mathrm{Sp}_n(\mathbb{F}_q)$, or $\mathrm{SU}_n(\mathbb{F}_q)$. Let S be any symmetric generating set for G. Suppose we have a non-trivial element $A \in G$ of length $d > 0$ and degree $k < n$. Then the diameter of G with respect to S will be $O((2q^{2nk} + d)\frac{n}{k})$.*

Proof. For any B conjugate to A, by the Lemma 5.4.3 above, $B = MAM^{-1}$ for some $M \in G$ of length at most q^{2nk}. So B has length at most $2q^{2nk} + d$. So every conjugate of A in G has length bounded by $2q^{2nk} + d$.

Furthermore, since A has degree $< n$ but non-trivial, it is not a scalar matrix. Then by [Gro02], A is not in the center of G.

Now by the resulf of Liebeck and Shalev [LS01], we know that every element of G can be written as a product of $O(\frac{n}{k})$ conjugates of A. So the whole group G has a diameter bound of $O((2q^{2nk} + d)\frac{n}{k})$. □

CHAPTER 6

Degree Reduction in Finite Linear Groups

The goal of this chapter is to establish an algorithm to reach an element of small degree, starting from any generating set of a linear group.

6.1 An Inequality on Primes

In this section, we shall establish an inequality on primes to be used in the next section.

Throughout this section, we shall fix a prime power q, which in the next section shall become the characteristic and the order of a finite field.

Let $p_1, ..., p_r$ be the first r primes coprime to $q(q-1)$. Let $\ell_q(p_i)$ be the multiplicative order of q in the field $\mathbb{Z}/p_i\mathbb{Z}$. Let M be the least common multiple of $\ell_q(p_1), ..., \ell_q(p_r)$. Let S be the sum of $\ell_q(p_1), ..., \ell_q(p_r)$. Our goal for this section is the following proposition:

Lemma 6.1.1. *There exist absolute constants c_1 and c_2 such that, if $p_r \geq c_1 \log q_0$, then*

$$S \leq (p_r)^2 \leq c_2(\log M)^3.$$

Before we prove this, let us first set up some properties of these multiplicative orders.

Lemma 6.1.2. *Let $E(x) := \{p \leq x : \ell_q(p) \leq \frac{\sqrt{p}}{logp}\}$. Then there is an absolute constant c_E such that*

$$|E(x)| \leq c_E \frac{x \log q}{(\log x)^3}.$$

Proof. For each $p \in E(x)$, there is a number $1 \leq r \leq \frac{\sqrt{p}}{\log p}$ such that p divides $q^r - 1$.

Then by pigeon hole principal, there is a number $1 \leq m \leq \frac{\sqrt{x}}{\log x}$, such that $q^m - 1$ is divided by $\frac{|E(x)|}{\sqrt{x}/\log x}$ primes.

Then we have

$$\frac{|E(x)|}{\sqrt{x}/\log x} \ll \frac{\log(q^m-1)}{\log\log(q^m-1)} \ll \frac{m\log q}{\log m} \ll \frac{\sqrt{x}\log q}{(\log x)^2}.$$

□

Now let us first set up more notations. Let P^+ be the function that sends each positive integer to its largest prime factor. Let $P = \{p_1, ..., p_r\}$. For any $\delta > 0$, let $P_\delta = \{\text{prime number } p : 3 \leq p \leq p_r, P^+(p-1) \geq (p_r)^\delta\}$, and let $P_\delta^* = P_\delta \cap P - E(p_r)$.

We start by citing an important theorem of Fouvry.

Lemma 6.1.3 ((Fouvry [Fou85])). *There is an absolute constant $\delta > \frac{2}{3}$, and an absolute constant c_0, c_p, such that for $p_r \geq c_p$, we have*

$$|P_\delta| \geq c_0 \frac{p_r}{\log p_r}.$$

Corollary 6.1.4. *There is an absolute constant $\delta > \frac{2}{3}$, and absolute constants c_0 and c_3, such that $|P_\delta^*| \geq c_0 \frac{p_r}{\log p_r} - c_3 \frac{\log q}{\log\log q} - c_E \frac{p_r \log q}{(\log p_r)^3}$ if $q > e^e$, and $|P_\delta^*| \geq c_0 \frac{p_r}{\log p_r} - 3 - c_E \frac{p_r \log q}{(\log p_r)^3}$ if $q \leq e^e$.*

Proof. By prime number theorem, when $\log\log q > 1$ (i.e., $q > e^e \approx 15.15$), the number of prime factors of $q(q-1)$ is bounded by $c_3 \frac{\log q}{\log\log q}$ for some absolute constant c_3. If $q \leq e^e$, then there are at most 3 prime factors of $q(q-1)$. □

Lemma 6.1.5. *Let $p \geq (p_r)^\delta$ be some prime. Then*

$$|((P^+)^{-1}(p)+1) \cap P_\delta^*| \leq \frac{2p_r}{(\log 2)(p_r^\delta - 1)}.$$

Proof. We assume that the left hand side of the inequality is non-zero, because otherwise the inequality is trivial. Then p divides some $p_i - 1$, which is not a prime. So in particular, $2p < p_r$.

The set $((P^+)^{-1}(p)+1)\cap P_\delta^*$ is contained in the set of primes $\le p_r$ that are congruent to 1 mod p. By the Brun-Titchmarsh theorem, combined with the fact that $p \ge (p_r)^\delta$, we have

$$\begin{aligned}|(P^+)^{-1}(p)\cap P_\delta^*| \le & \frac{2p_r}{\phi(p)\log\frac{p_r}{p}} \\ \le & \frac{2p_r}{(\log 2)(p-1)} \\ \le & \frac{2p_r}{(\log 2)((p_r)^\delta - 1)}.\end{aligned}$$

Here ϕ is the Euler totient function. $\square$

Lemma 6.1.6. *All primes in $P^+(P_\delta^* - 1)$ are factors of M.*

Proof. It is enough to show that, if $p \in P_\delta^*$, then $\ell_q(p)$ contains $P^+(p-1)$ as a factor.

Since $p \in P_\delta^* \subseteq P_\delta$, we know $P^+(p-1) \ge (p_r)^\delta \ge p^{\frac{2}{3}}$. On the other hand, since $p \in P_\delta^* \in P - E(x)$, we know $\ell_q(p) \ge \frac{\sqrt{p}}{\log p} > \frac{p-1}{P^+(p-1)}$.

Therefore, $\ell_q(p)$ must contain $P^+(p-1)$ as a factor. $\square$

Now we have enough to prove Lemma 6.1.1.

of Lemma 6.1.1. The first inequality is straightforward

$$S \le \sum_{\text{prime } p \le p_r} p \le (p_r)^2.$$

All the primes in $P^+(P_\delta^* - 1)$ are factors of M, and they are all larger than $(p_r)^\delta$. Furthermore, when $q_0 > e^e$, we have

$$\begin{aligned}|P^+(P_\delta^* - 1)| \ge & \frac{|P_\delta^*|}{\max_{p\ge(p_r)^\delta} |((P^+)^{-1}(p)+1)\cap P_\delta^*|} \\ \ge & (c_0\frac{p_r}{\log p_r} - c_3\frac{\log q_0}{\log\log q_0} - c_E\frac{p_r \log q_0}{(\log p_r)^3})/(\frac{2p_r}{(\log 2)((p_r)^\delta - 1)}) \\ \ge & \frac{\log 2}{2}((p_r)^\delta - 1)(\frac{c_0}{\log p_r} - \frac{c_3}{p_r}\frac{\log q_0}{\log\log q_0} - c_E\frac{\log q_0}{(\log p_r)^3}).\end{aligned}$$

So, if $p_r > c_1 \log q_0$ for an absolute constant c_1 such that $c_3\frac{1+\log c_1}{c_1} + \frac{c_E}{\log c_1} < \frac{c_0}{2}$, then we have

$$\begin{aligned}
\log M \geq & |P^+(P^*_\delta - 1)| \log((p_r)^\delta) \\
\geq & \frac{\log 2}{2}\delta((p_r)^\delta - 1)(c_0 - c_3\frac{\log q_0}{\log\log q_0}\frac{\log p_r}{p_r} - c_E\frac{\log q_0}{(\log p_r)^2}) \\
\geq & \frac{\log 2}{2}\delta((p_r)^\delta - 1)(c_0 - c_3\frac{1+\log c_1}{c_1} - \frac{c_E}{\log c_1}) \\
\geq & \frac{\log 2}{4}\delta c_0((p_r)^\delta - 1).
\end{aligned}$$

Since $\delta > \frac{2}{3}$, we can pick some constant such that $c_2(\log M)^3 \geq (p_r)^2$.

Now, suppose $q_0 \leq e^e$. Then similarly we have

$$\begin{aligned}
|P^+(P^*_\delta - 1)| \geq & \frac{|P^*_\delta|}{\max_{p \geq (p_r)^\delta} |((P^+)^{-1}(p) + 1) \cap P^*_\delta|} \\
\geq & (c_0\frac{p_r}{\log p_r} - 3 - c_E\frac{p_r \log q_0}{(\log p_r)^3})/(\frac{2p_r}{(\log 2)((p_r)^\delta - 1)}) \\
\geq & \frac{\log 2}{2}((p_r)^\delta - 1)(\frac{c_0}{\log p_r} - \frac{3}{p_r} - c_E\frac{\log q_0}{(\log p_r)^3}).
\end{aligned}$$

So, if $p_r > c_1 \log q_0$ for a sufficiently large absolute constant c_1 such that $\frac{3\log p_r}{p_r} + \frac{c_E}{\log c_1} < \frac{c_0}{2}$, then we have

$$\begin{aligned}
\log M \geq & |P^+(P^*_\delta - 1)| \log((p_r)^\delta) \\
\geq & \frac{\log 2}{2}\delta((p_r)^\delta - 1)(c_0 - \frac{3\log p_r}{p_r} - c_E\frac{\log q_0}{(\log p_r)^2}) \\
\geq & \frac{\log 2}{4}\delta c_0((p_r)^\delta - 1).
\end{aligned}$$

Since $\delta > \frac{2}{3}$, we can again pick some constant such that $c_2(\log M)^3 \geq (p_r)^2$. $\square$

As a side note, for any improved value of δ in the Fouvry's theorem, our diameter bound in this paper would improve to $q^{O(n(\log n + \log q)^{\frac{2}{\delta}})}$ for finite simple groups of Lie type of rank n over $\mathbb{F}_q$.

If one were to assume the Hardy-Littlewood conjucture on prime tuples, the δ could be improved to $1 - o(1)$. Combine this with the more efficient estimate $S \leq \frac{(p_r)^2}{\log p_r}$, the diameter bound of this paper would improve to $q^{O(n(\log n + \log q)^2)}$ for finite simple groups of Lie type of rank n over $\mathbb{F}_q$.

6.2 P-Matrices and Degree Reduction

This section aims to show that, given a P-matrix, we can reduce its degree by raising it to a large power.

Definition 6.2.1. Let $\mathbb{F}_q$ be a finite field of characteristic p, and let $p_1, p_2, ..., p_r$ be the first r primes coprime to $p(q-1)$. Then a matrix A over $\mathbb{F}_q$ is called a ***P(r)-matrix*** if, for each $i \leq r$, it has a primitive p_i-th root of unity in the algebraic closure of $\mathbb{F}_q$ as an eigenvalue.

Lemma 6.2.2. *Let A be a matrix over $\mathbb{F}_q$, a field with characteristic p. Let m be any number coprime to p. Then if A has a primitive m-th root of unity as an eigenvalue, A must have degree at least $\ell_q(m)$.*

Proof. Let $\Phi_m(X)$ be the m-th cyclotomic polynomial. Then by Galois theory over finite field, the polynomial $\Phi_m(X)$ factors into distinct irreducible polynomials of degree $\ell_q(m)$. See, e.g., [Lan02].

Therefore, if A has a primitive m-th root of unity as an eigenvalue, then A must have at least $\ell_q(m)$ primitive m-th roots of unity as eigenvalues, and the result follows. $\square$

Lemma 6.2.3. *Let n be an integer, and let q be a power of the prime p. Then we can find an integer r and an absolute constant c, such that the following is true:*

1. *Let $p_1, p_2, ..., p_r$ are the first r primes coprime to $p(q-1)$. Then $\mathrm{lcm}_{i=1}^{r} \ell_q(p_i) > n^4$, and $\sum_{i=1}^{r} \ell_q(p_i) < c(\log n + \log q)^3$.*
2. *Let $A \in \mathrm{GL}_n(\mathbb{F}_q)$ where the field has characteristic p, and $\deg A = k$. If A is a P(r)-matrix, then there exists $\ell \in \mathbb{N}$ such that A^ℓ will be a non-identity matrix of degree at most $\frac{k}{4}$, and 1 is the only eigenvalue of A^ℓ lying in $\mathbb{F}_q$.*

Proof.

The First Statement:

Let M be the least common multiple, and let S be the sum. Let c_1 be the constant as in Lemma 6.1.1.

Pick p_r to be the smallest prime such that $M > n^4$ and $p_r > c_1 \log q$. Then the second condition guarantees that $S < c_2(\log M)^3$, according to Lemma 6.1.1.

Now, if $p_r \leq 2c_1 \log q$, then for some absolute constant c_4 by the Prime Number Theorem, we have

$$\begin{aligned}\log M \leq &\sum_{i=1}^{r} \log p_r \\ \leq &c_4 p_r \\ \leq &2c_1 c_4 \log q.\end{aligned}$$

So $S \leq c(\log q)^3$ for some absolute constant c.

Suppose $p_r > 2c_1 \log q$. Then by the Bertrand-Chebyshev Theorem, $p_{r-1} > c_1 \log q$. Let $M' = \mathrm{lcm}_{i=1}^{r-1} \ell_q(p_i)$. Then by the minimality of p_r, we must have $M' \leq n^4$. In particular, we have

$$\begin{aligned}4 \log n \geq &\log M' \\ \geq &(\frac{p_{r-1}}{\sqrt{c_2}})^{\frac{2}{3}}.\end{aligned}$$

So, we have $p_{r-1} \leq 8\sqrt{c_2}(\log n)^{\frac{3}{2}}$. Then $p_r \leq 16\sqrt{c_2}(\log n)^{\frac{3}{2}}$.

Furthermore, we have

$$\begin{aligned}\log M \leq &\log(M' p_r) \\ \leq &\log(n^4(16\sqrt{c_2}(\log n)^{\frac{3}{2}})) \\ < &6 \log n + \log(16\sqrt{c_2}).\end{aligned}$$

So $S < c_1(\log M)^3 < c(\log n)^3$ for some absolute constant c.

The Second Statement:

Let M_i denote the least common multiple of $\ell_q(p_1), \ell_q(p_2), ..., \ell_q(p_i)$. Let $t_1 = \ell_q(p_1)$ and $t_i = \frac{M_i}{M_{i-1}}$ for $i > 1$. Then $\prod_{i=1}^{r} t_i = M_r > n^4$.

Let $N = \{1, 2, ..., n\}$. Let $d_1, ..., d_n$ be the eigenvalues of A in the algebraic closure of $\mathbb{F}_q$.

For each $j \in N$, let P_j be the set of prime factors of the multiplicative order of d_j among $p_1, ..., p_r$. Then by Lemma 6.2.2, for each $j \in N$,

$$\prod_{p_i \in P_j} t_i \leq \mathrm{lcm}_{p_i \in P}\, \ell_q(p_i) \leq k.$$

Now let $n(i)$ denote the number of P_j that contain p_i.

We take the weighted average T of these $n(i)$ with weight $\log t_i$. The sum of the weights is $\sum_{i=1}^{r} \log t_i > 4 \log n$.

$$\begin{aligned}
T &= \frac{\sum_{1\leq i\leq r} n(i) \log t_i}{\sum \log t_i} \\
&= \frac{\sum_{1\leq i\leq r} \sum_{j\in N} (\log t_i) 1_{p_i \in P_j}}{\sum \log t_i} \\
&= \frac{\sum_{j\in N} \sum_{1\leq i\leq r} (\log t_i) 1_{p_i \in P_j}}{\sum \log t_i} \\
&\leq \frac{\sum_{j\in N} \sum_{p_i \in P_j} \log t_i}{4 \log n} \\
&\leq \frac{k \log k}{4 \log n} \\
&\leq \frac{k}{4}.
\end{aligned}$$

So there is a p_i such that $n(i) \leq \frac{k}{4}$. So if A has order $m(A)$, then $A^{\frac{m(A)}{p_i}}$ is the desired non-identity matrix of degree at most $\frac{k}{4}$. Every eigenvalue of the latter matrix not equal to 1 is a primitive p_i-th root of unity, which would be outside of $\mathbb{F}_q$. $\square$

6.3 Commutators and Degree Reduction

In this section, we shall use repeated commutators with elements of a t-transversal set. This way, we repeatedly create P-matrices and raise them to a large power, and would eventually end up with a matrix of very small degree.

Definition 6.3.1. Given any element g of a group G and a symmetric generating set S for G, the ***length*** of g is $\ell(g) = \min\{d \in \mathbb{N} : g = s_1 ... s_d \text{ for some } s_1, ..., s_d \in S\}$.

Proposition 6.3.2. *For any matrices A, B, $\deg(ABA^{-1}B^{-1}) \leq 2\min(\deg A, \deg B)$.*

Proof.

$$\begin{aligned}
\deg(ABA^{-1}B^{-1}) &= \operatorname{rank}(ABA^{-1}B^{-1} - I) \\
&= \operatorname{rank}(AB - BA) \\
&= \operatorname{rank}((A - I)(B - I) - (B - I)(A - I)) \\
&\leq \operatorname{rank}(A - I)(B - I) + \operatorname{rank}(B - I)(A - I) \\
&\leq \operatorname{rank}(A - I) + \operatorname{rank}(A - I) \\
&= 2\operatorname{rank}(A - I).
\end{aligned}$$

Similarly, we also have $\deg(ABA^{-1}B^{-1}) \leq 2\operatorname{rank}(B - I)$. So we are done. □

Lemma 6.3.3. *Fix any matrix $A \in \mathrm{GL}(V)$ of degree k, such that the eigenvalues of A are either 1 or outside of $\mathbb{F}_q$. For any $t \leq \frac{k}{2}$, we can find a subspace W of V with the following properties:*

1. $\dim W = t$;
2. $W \cap AW = \{0\}$.

Proof. We shall proceed by induction on the dimension of W. Let V_A be the subspace of fixed points of A in V.

Initial Step: Suppose $t = 1$. Simply pick any vector v outside of V_A, and let W be the span of v. We have $W \cap V_A = \{0\}$ by choice of v. Since A has no eigenvalue in $\mathbb{F}_q$ other than 1, v and Av must be linearly independent. So $W \cap AW = \{0\}$.

Inductive Step: Suppose we have found a subspace W of dimension $t - 1$ such that $W \cap AW = \{0\}$. I claim that, when $t \leq \frac{k}{2}$, we can find another vector v, such that the desired subspace is the span of v and W.

To prove the existence of v, let us count the number of vectors to avoid. We want v to avoid $V_A + W + AW$. Afterwards, it is enough to let Av avoid any linear combination of v

and $W + AW$. So we need v to avoid $\bigcup_{x\in\mathbb{F}_q}(A-x)^{-1}(W+AW)$. Here we shall interpret $(A-x)^{-1}$ as the pullback map of subsets.

Now, since A has no eigenvalue in $\mathbb{F}_q$ other than 1, therefore $A - x$ is invertible when $x \neq 1$. And $A - 1$ has kernel exactly V_A, which has dimension $n - k$. So, we have

$$\begin{aligned}
&|(V_A + W + AW) \cup (\bigcup_{x\in\mathbb{F}_q}(A-x)^{-1}(W+AW))| \\
\leq& q^{n-k+2t-2} + q^{n-k+2t-2} + (q-1)q^{2t-2} \\
<& q^{n-k+2t}.
\end{aligned}$$

So as long as $2t \leq k$, we have $q^{n-k+2t} \leq q^n$. So it is possible to choose a vector v as desired. $\square$

Lemma 6.3.4. *Let A be a matrix in the group $\mathrm{GL}_n(\mathbb{F}_q)$. Then the order of A is less than q^n.*

Proof. By the rational canonical form of a matrix A (see, e.g., [Her75]), it is enough to prove the case when A is a single rational jordan block.

In this case, the characteristic polynomial $f(x)$ of A is a power of an irreducible polynomial $g(x)$. Say $f = g^t$ and the characteristic of $\mathbb{F}_q$ is p. Then the order of A is $(q^{\frac{n}{t}} - 1)p^{\lceil \frac{\log t}{\log p} \rceil} < q^{\frac{n}{t}} p^{\lceil \frac{\log t}{\log p} \rceil}$.

Then it is enough to show that $p^{\lceil \frac{\log t}{\log p} \rceil} \leq q^{(1-\frac{1}{t})n}$. And since $p^{\lceil \frac{\log t}{\log p} \rceil}$ is the smallest p-power that is larger than or equal to t, and since q is a power of p, it is enough to show that $t \leq q^{(1-\frac{1}{t})n}$. And this is true since $q \geq 2$ and $n \geq t$. $\square$

Proposition 6.3.5. *For any symmetric generating set S of $\mathrm{GL}_n(\mathbb{F}_q)$, $\mathrm{GL}_n(\mathbb{F}_q)$ has a non-trivial element of degree at most $C(\log n + \log q)^3$ for some absolute constant C, of length less than $q^{C'n(\log n+\log q)^3}$ for some absolute constant C'. The same statement is true with $\mathrm{SL}_n(\mathbb{F}_q)$ replacing $\mathrm{GL}_n(\mathbb{F}_q)$.*

Proof. We pick r and c according to Lemma 6.2.3. We may assume that $c(\log n + \log q)^3 < n$, because otherwise the statement is trivial.

Let $p_1, ..., p_r$ be the first r primes coprime to $p(q-1)$. Let $f_i(x)$ be the irreducible polynomial over $\mathbb{F}_q$ for all the primitive p_i-th roots of unity, and let C_i be the companion matrix of $f_i(x)$.

Initial Step:

Let us find our first P(r)-matrix. Let T be a $c(\log n + \log q)^3$-transversal set. Then by definition, we can find $A_0 \in T$ that maps some subspace W of dimension $c(\log n + \log q)^3$ onto itself, and that its restriction to this subspace is the matrix $(\bigoplus_{i=1}^r C_i) \oplus I$ for some arbitrary choices of basis on W, where I is some identity matrix of suitable size.

In particular, A_0 is a P(r)-matrix. Since $A_0 \in T$, by choosing T as in Corollary 5.2.4, A_0 has length bounded by $q^{cn(\log n+\log q)^3}$.

By using Lemma 6.2.3, we can raise A_0 to a large power, and obtain a non-identity matrix A_1 of degree $\leq \frac{\deg(A_0)}{4} \leq \frac{n}{4}$, with eigenvalues either 1 or outside of $\mathbb{F}_q$. Since the order of A_0 is bounded by q^n by Lemma 6.3.4, the length of A_1 is bounded by $q^{cn(\log n+\log q)^3+n}$.

Inductive Step:

Suppose we have obtained a non-identity matrix A_j with eigenvalues either 1 or outside of $\mathbb{F}_q$, degree at most $\frac{n}{2^{j+1}}$, and length at most $q^{2cn(\log n+\log q)^3+j(n+2)}$. If $\deg A_j \leq 2c(\log n + \log q)^3$, then we stop. If not, then let us construct a non-identity matrix A_{j+1} of even smaller degree.

First we shall transform A_j into a P(r)-matrix. Find a subspace W_j of dimension at least $c(\log n + \log q)^3$ as in Lemma 6.3.3 using A_j. In particular, W_j has trivial intersection with A_jW_j. Let T' be a $2c(\log n + \log q)^3$-transversal set, then we can find $M_j \in T'$ that fixes A_jW_j, and restricts to a map from W_j to W_j as $\bigoplus_{i=1}^r C_i \oplus I$ for an arbitrary basis of W_j and some identity matrix I of suitable size.

Consider the commutator $M_jA_j^{-1}M_j^{-1}A_j$. Since M_j fixes A_jW_j, we see that $M_jA_j^{-1}M_j^{-1}A_j$ restricted to W_j is identical to M_j restricted to W_j.

In particular, $M_jA_j^{-1}M_j^{-1}A_j$ is a P(r)-matrix, and it has degree at most $2\deg(A_j)$. Now we use Lemma 6.2.3 again, raising $M_jA_j^{-1}M_j^{-1}A_j$ to a large power, and we would obtain a matrix A_{j+1} of degree at most $\frac{2\deg(A_j)}{4}$, with eigenvalues either 1 or outside of $\mathbb{F}_q$.

Since $M_j \in T'$, by choosing T' as in Corollary 5.2.4, M_j have length bounded by $q^{2cn(\log n+\log q)^3}$. And since the order of $M_j A_j^{-1} M_j^{-1} A_j$ is bounded by q^n by Lemma 6.3.4, the length of A_{j+1} is at most

$$q^n(2q^{2cn(\log n+\log q)^3} + 2q^{2cn(\log n+\log q)^3+j(n+2)}) \le q^{2cn(\log n+\log q)^3+(j+1)(n+2)}.$$

We repeat the above induction $\frac{\log n}{\log 2} - 1$ times, or stop early if we hit degree $2c(\log n + \log q)^3$. The last A_j we obtained is the desired matrix of small degree and small length. $\square$

6.4 Degree Reducing for Orthogonal, Symplectic, Unitary Groups

Lemma 6.4.1. *For any non-zero singular $v \in V$, there is a vector $w \in V$ such that v, w form a hyperbolic pair.*

Proof. Recall that V is a non-degenerate formed space with an alternating bilinear, symmetric bilinear or Hermitian form B. In the case of characteristic 2, a symmetric bilinear form B might be degenerate, even though the formed space itself is not. Let σ be the field automorphism of the base field F for the Hermitian form B, or identity if B is bilinear.

For any element $k \in F$, we define $\mathrm{Tr}(x) = x + \sigma(x)$.

Now given a singular $v \in V$, since V is a non-degenerate formed space, we can find a vector $w' \in V$ such that $B(v, w') \neq 0$. By scaling w', we can assume that $B(v, w') = 1$.

Suppose we can find an element $k \in F$ such that $\mathrm{Tr}(k) = B(w', w')$, then $w = w' - kv$ is the desired vector forming a hyperbolic pair with v: $B(v, w) = B(v, w') = 1$, and

$$\begin{aligned} & B(w' - kv, w' - kv) \\ =& B(w', w') - \mathrm{Tr}(k) \\ =& 0. \end{aligned}$$

Now, it remains to show that such k always exists.

Let E be the subfield of F fixed by σ. Then obviously $B(w', w') \in E$. So it is enough to show that either $E = \mathrm{Tr}(F)$, or $B(w', w') = 0$ for all w'.

Now, $\mathrm{Tr}(F)$ is closed under addition, and it is also closed under multiplication by elements of E. So $\mathrm{Tr}(F)$ is a E-vector space contained in E. So either $E = \mathrm{Tr}(F)$, or $\mathrm{Tr}(F) = 0$.

In the case that $\mathrm{Tr}(F) = 0$, then $\sigma(x) = -x$ for all $x \in F$. But since σ is a field automorphism, we must conclude that the field F has characteristic 2, and σ is the identity. Then the form B is alternating, and $B(w', w') = 0$ for any $w' \in V$. □

Lemma 6.4.2. *If a subspace H of V is an orthogonal sum of hyperbolic planes, then $H \cap H^{\perp} = \{0\}$.*

Proof. The subspace H is an orthogonal sum of hyperbolic planes. Then let us assume that these planes are the linear span of hyperbolic pairs $(v_1, w_1), (v_2, w_2), (v_3, w_3), ..., (v_t, w_t)$.

Suppose $v \in H \cap H^{\perp}$. Then for some scalars $a_i, b_i \in F$, we have

$$v = \sum_{i=1}^{t} a_i v_i + \sum_{i=1}^{t} b_i w_i.$$

Now, since $B(v, v_i) = 0$, we can deduce that $b_i = 0$. Similarly, since $B(v, w_i) = 0$, we can deduce that $a_i = 0$. So $v = 0$. □

Lemma 6.4.3. *Fix any nonzero elements $a, b, c \in \mathbb{F}_q$. Then the equation $ax^2 + by^2 + cz^2 = 0$ has a non-trivial solution in $\mathbb{F}_q$.*

Proof. If $\mathrm{char}(\mathbb{F}_q) = 2$, then $(\mathbb{F}_q)^*$ is a multiplicative group of odd order. So every nonzero element of $\mathbb{F}_q$ is a square.

Find x, y, z such that $x^2 = a^{-1}$, $y^2 = b^{-1}$ and $z = 0$. This is a non-trivial solution of the equation.

Suppose q is odd. Let S be the set of squares in $\mathbb{F}_q$. Then $|S| = \frac{q+1}{2}$. Then $|aS| + |-c - bS| > |\mathbb{F}_q|$. As a result, we have $aS \cap (-c - bS) \neq \varnothing$. So $-c \in aS + bS$.

Pick $x, y \in \mathbb{F}_q$ such that $ax^2 + by^2 = -c$. Then the triple $(x, y, 1)$ is a non-trivial solution to the equation. □

Lemma 6.4.4. *Fix any $A \in G$. Then given any totally singular subspace $W \in V$ of dimension d, we can find a subspace W' of W such that W' is perperdicular to AW', and W' has dimension at least $\frac{d}{4} - \frac{3}{2}$.*

Proof. We proceed by induction on the dimension of W.

Initial Step: For the base case of the induction, suppose the dimension of W is 7 or 8 or 9 or 10. Then all we need is to find a nonzero vector $v \in W$ such that $v \perp Av$. Suppose for contradiction that there is no such vector.

Pick any non-zero $v_1 \in W$. Let W_1 be the intersection of W and $\text{span}\{v_1, Av_1, A^{-1}v_1\}^{\perp}$. Since v_1 is not perpendicular to Av_1, it is not in $\text{span}\{v_1, Av_1, A^{-1}v_1\}^{\perp}$. So $v_1 \notin W_1$, and W_1 has dimension at least $\dim W - 3 \geq 4$.

Pick any non-zero $v_2 \in W_1$. Let W_2 be the intersection of W_1 and $\text{span}\{v_2, Av_2, A^{-1}v_2\}^{\perp}$. Then W_2 has dimension at least $\dim W_1 - 3 \geq 1$ and similarly $v_2 \notin W_2$. Pick any non-zero $v_3 \in W_2$.

Now, we know $B(v_1, Av_1), B(v_2, Av_2), B(v_3, Av_3)$ are all in $\mathbb{F}_q^*$. We shall divide our discussion into two cases:

Orthogonal or Symplectic Case: Let $a = B(v_1, Av_1), b = B(v_2, Av_2), c = B(v_3, Av_3)$. Then by Lemma 6.4.3, we can find a nontrivial triple $x, y, z \in \mathbb{F}_q$ such that $ax^2+by^2+cz^2 = 0$. Let $v = xv_1 + yv_2 + zv_3$, then we have

$$\begin{aligned}
B(v, Av) =& x^2 B(v_1, Av_1) + y^2 B(v_2, Av_2) + z^2 B(v_3, Av_3) \\
=& ax^2 + by^2 + cz^2 \\
=& 0.
\end{aligned}$$

Unitary Case: If B is a Hermitian form for a field automorphism σ of order 2, then let F be the fixed subfield of σ. Let $N : \mathbb{F}_q \to F$ be the field norm, which is surjective.

Now, $\mathbb{F}_q$ is an F-vector space of dimension 2. So $B(v_1, Av_1), B(v_2, Av_2), B(v_3, Av_3)$ cannot be F-linearly independent in $\mathbb{F}_q$. So one can find non-trivial triple $a_1, a_2, a_3 \in F$ such that $a_1 B(v_1, Av_1) + a_2 B(v_2, Av_2) + a_3 B(v_3, Av_3) = 0$.

Since the norm map is surjective, find $x_1, x_2, x_3 \in \mathbb{F}_q$ such that $N(x_i) = a_i$. Let $v = x_1v_1 + x_2v_2 + x_3v_3$. Then we have

$$\begin{aligned} B(v, Av) =& N(x_1)B(v_1, Av_1) + N(x_2)B(v_2, Av_2) + N(x_3)B(v_3, Av_3) \\ =& a_1B(v_1, Av_1) + a_2B(v_2, Av_2) + a_3B(v_3, Av_3) \\ =& 0. \end{aligned}$$

So in either case, we could find the desired non-trivial vector $v \in W$ such that $v \perp Av$.

Inductive Step: Now let us proceed for general W of larger dimension. Since the dimension of W is at least 7, by the argument in the base case of the induction, we can find $v_1 \in W$ such that $B(v_1, Av_1) = 0$. Let W_1 be the intersection of W and $\operatorname{span}\{v_1, Av_1, A^{-1}v_1\}^{\perp}$. Then W_1 has dimension at least $d-3$. Pick any subspace W_2 of W_1 linearly independent from v_1. Then W_2 has dimension at least $d-4$ and at most $d-1$. Then by induction hypothesis, we can find W_2' a subspace of W_2, such that W_2' is perpendicular to AW_2', and W_2' has dimension at least $\frac{d-4}{4} - \frac{3}{2}$.

Let W' be the span of W_2' and v_1. Then W' will be perpendicular to AW', and has dimension at least $\frac{d-4}{4} - \frac{3}{2} + 1 = \frac{d}{4} - \frac{3}{2}$. So we are done. $\square$

Lemma 6.4.5. *Fix any $A \in G$ where all eigenvalues of A are outside of $\mathbb{F}_q$. Then there is a t-dimensional totally singular subspace W of V such that $W \cap AW = \{0\}$, for any $t \leq \frac{n}{6}$.*

Proof. Fix n, which we assume to be at least 3, so that V has at least one singular vector. We shall proceed by induction on the dimension of W.

Initial Step: Suppose $t = 1$. Simply pick any singular vector v, and let W be the span of v. Since A has no eigenvalue in $\mathbb{F}_q$, v and Av must be linearly independent. So $W \cap AW = \{0\}$.

Inductive Step: Suppose we have found a totally singular subspace W of dimension $t-1$ such that $W \cap AW = \{0\}$. I claim that, when $t \leq \frac{n}{6}$, we can find another singular vector v, such that the desired subspace is the span of v and W.

First of all, we want v to be a singular vector perpendicular to W. We know $W^{\perp}$ has dimension $n - t + 1$, and by Witt's decomposition theorem, V has a totally singular space of dimension at least $\frac{n-2}{2}$. This totally singular space will intersect $W^\perp$ in a subspace of dimension at least $\frac{n-2}{2}-t+1 = \frac{n}{2}-t$. So there are at least $q^{\frac{n}{2}-t}$ singular vectors perpendicular to W.

Among these vectors, to prove the existence of a good v, we should count the number of vectors to avoid. We need v to avoid $W + AW$. Afterwards, it is enough to have Av avoiding any linear combination of v and $W + AW$. To satisfy the second requirement, we need v to avoid $\bigcup_{x\in\mathbb{F}_q}(A - x)^{-1}(W + AW)$. Here we shall interpret $(A - x)^{-1}$ as the pullback map of subsets.

Now, since A has no eigenvalue in $\mathbb{F}_q$, therefore $A - x$ are all invertible. So, we have

$$
\begin{aligned}
&|(W + AW) \cup (\bigcup_{x\in\mathbb{F}_q}(A - x)^{-1}(W + AW))| \\
\leq &q^{2t-2} + q \times q^{2t-2} \\
< &q^{2t}.
\end{aligned}
$$

So as long as $2t \leq \frac{n}{2} - t$, i.e., $t \leq \frac{n}{6}$, then it is possible to choose a vector v as desired. □

Lemma 6.4.6. *Fix any matrix $A \in G$ of degree k, such that the eigenvalues of A are either 1 or outside of $\mathbb{F}_q$. Then we can find a subspace W of V with the following properties:*

1. $\dim W \geq \frac{k}{32} - \frac{7}{4}$
2. *W is totally singular;*
3. *$W \cap AW = \{0\}$;*
4. *$W \perp AW$.*

Proof. Let V_A be the subspace of fixed points of A in V. Let $V_r = V_A \cap (V_A)^\perp$. Choose any positive number a to be determined later. Then either V_r has dimension $< a$, or it has dimension $\geq a$.

Case of Large V_r:

Suppose V_r has dimension $\geq a$. Pick any non-zero singular $v_1 \in V_r$, then we can find $w_1 \in V$ such that v_1, w_1 form a hyperbolic pair. Let V_{r1} be the intersection of V_r with $\text{span}\{v_1, w_1\}^\perp$. Pick any non-zero singular $v_2 \in V_{r1}$, then we can find w_2 in $\text{span}\{v_1, w_1\}^\perp$, such that v_2, w_2 form a hyperbolic pair. Then let V_{r2} be the intersection of V_{r1} with $\text{span}\{v_1, w_1, v_2, w_2\}^\perp$, and repeat.

As long as $\dim V_{ri} > 2$, then V_{ri} cannot be anisotropic. So we can keep going at least $\lfloor \frac{a-2}{2} \rfloor$ times. Thus we obtained $w_1, ..., w_{\lfloor \frac{a-2}{2} \rfloor}$. They span a totally singular space W_r of dimension at least $\frac{a-3}{2}$. Then by Lemma 6.4.4, we can find a subspace W of W_r, such that $W \perp AW$ and W has dimension at least $\frac{a-3}{8} - \frac{3}{2}$.

I claim that, ignoring the dimension requirement, this W satisfies all the desired properties. By construction of W, we have W totally singular and $W \perp AW$. We only need to show that $W \cap AW = \{0\}$.

For any vector $w = \sum_{i=1}^{\lfloor \frac{a-2}{2} \rfloor} a_i w_i \in W$, suppose it is perpendicular to V_r. Then for each i, since $B(v_i, w) = 0$, we see that $a_i = 0$. So $w = 0$. To sum up, W has trivial intersection with $(V_r)^\perp$.

Suppose $w \in W \cap AW$. Then $w - A^{-1}w \in W$, and for any $v \in V_r$ we have

$$\begin{aligned} B(v, w - A^{-1}w) =& B(v, w) - B(v, A^{-1}w) \\ =& B(v, w) - B(Av, w) \\ =& B(v, w) - B(v, w) \\ =& 0. \end{aligned}$$

So $w - A^{-1}w \in W \cap V_r^\perp = \{0\}$. So $w = Aw$, and $w \in W \cap V_A \subseteq W \cap (V_r)^\perp = \{0\}$.

To sum up, this W is the space we desired, with dimension at least $\frac{a-3}{8} - \frac{3}{2}$.

Case of Small V_r:

Suppose V_r has dimension $< a$.

Step 1: We want to first find a subspace W_A of $(V_A)^\perp$ where $W_A \perp W_A$, and $W_A \cap AW_A = V_r$, and the codimension of V_r in W_A is at least $\frac{k-a-5}{6}$.

Now, the bilinear or Hermitian form B restricted to $(V_A)^\perp$ is still bilinear or Hermitian, with exactly V_r as the radical. So the space $V' = (V_A)^\perp/V_r$ has an induced bilinear or Hermitian form B', and now B' is non-degenerate.

So V' is a non-degenerate formed space with dimension at least $k-a$. Furthermore, since V_r and $(V_A)^\perp$ are both A-invariant, A induces a linear map A' on V'. Clearly A' has no non-trivial fixed point in V', so all eigenvalues of A' are ouside of $\mathbb{F}_q$. So by Lemma 6.4.5, V' has a totally singular subspace W' of dimension at least $\lfloor \frac{k-a}{6} \rfloor \geq \frac{k-a-5}{6}$, such that $W' \cap A'W' = \{0\}$.

Let W_A be the pullback of W' through the projection map $(V_A)^\perp \to V'$. Since W' is totally singular under B', the form B vanishes on W_A. (Note that in the orthogonal case, the quadratic form Q might not vanish on W_A, so W_A might not be totally singular.)

Step 2: Now let us find a totally singular subspace W_r of W_A, which intersects trivially with V_r and has dimension at least $\frac{k-a-5}{6}$.

If $\operatorname{char} \mathbb{F}_q \neq 2$, or if we are not in the orthogonal case, then W_A is totally singular. Pick any subspace W_r of W_A which has trivial intersection with V_r and has dimension at least $\frac{k-a-5}{6}$, and we are done.

Suppose $\operatorname{char} \mathbb{F}_q = 2$, and we are in the orthogonal case, and Q vanishes on V_r. Then the space $V' = (V_A)^\perp/V_r$ would have an induced non-degenerate quadratic form Q' that corresponds to the non-degenerate bilinear form B'. So by Lemma 6.4.5, when we picked W' to be totally singular, we can pick it to be totally singular with respect to the quadratic form Q'. This way the subspace W_A would be totally singular. Again pick any subspace W_r of W_A which has trivial intersection with V_r and has dimension at least $\frac{k-a-5}{6}$, and we are done.

Finally, suppose now that $\operatorname{char} \mathbb{F}_q = 2$, and we are in the orthogonal case, and we have a vector $v_0 \in V_r$ such that $Q(v_0) \neq 0$. Since $\operatorname{char} \mathbb{F}_q = 2$, the squaring map is bijective on $\mathbb{F}_q$, we can assume that $Q(v_0) = 1$ by scaling v_0.

Define a map $X : W_A \to W_A$ such that $X(v) = v + \sqrt{Q(v)}v_0$. Here the square root is well defined because $\operatorname{char} \mathbb{F}_q = 2$. Then we have $Q(X(v)) = 0$ for all $v \in W_A$.

Furthermore, X is linear. To see this, first we notice that for any v, w in W_A, since B

vanishes on W_A,

$$Q(v+w) = Q(v) + Q(w) + B(v,w) = Q(v) + Q(w).$$

So we have

$$\begin{aligned} X(v+w) =& v + w + \sqrt{Q(v+w)}v_0 \\ =& v + w + \sqrt{Q(v) + Q(w)}v_0 \\ =& v + w + (\sqrt{Q(v)} + \sqrt{Q(w)})v_0 \\ =& X(v) + X(w). \end{aligned}$$

For any scalar $a \in \mathbb{F}_q$, we also easily have $X(av) = aX(v)$.

Now, since X is linear, $X(W_A)$ is a subspace of W_A. So $X(W_A)$ is a totally singular subspace.

Now pick any subspace W_r of W_A which has trivial intersection with V_r and has dimension at least $\frac{k-a-5}{6}$. Then $X(W_r)$ is a totally singular subspace of W_A. It remains to show that this $X(W_r)$ intersects V_r trivially and has the correct dimension.

For any vector v, if $X(v) \in V_r$, then $v = \sqrt{Q(v)}v_0 + X(v) \in V_r$. So $X(W_r)$ only has trivial intersection with V_r. And since the kernel of X is entirely in V_r, $X(W_r)$ has the same dimension as W_r.

So replace W_r by $X(W_r)$, and we are done.

Step 3: Now we construct the desired subspace W.

By Lemma 6.4.4, we find a subspace W of W_r such that $W \perp AW$, and W has dimension at least $\frac{k-a-5}{24} - \frac{3}{2}$.

I claim that, ignoring the dimension requirement, this W satisfies all the desired properties.

First of all, we know W is totally singular and $W \perp AW$. By construction, W is in $(V_A)^{\perp}$ but has trivial intersection with V_r. Then since $W_A \cap AW_A = V_r$, we know that $W \cap AW = \{0\}$.

To sum up, this W is the space we desired, with dimension at least $\frac{k-a-5}{24} - \frac{3}{2}$.

Find the Optimal a:

Picking the optimal $a = \frac{k}{4} + 1$ for both cases above, we eventually find the desired subspace W of dimension at least $\frac{k}{32} - \frac{7}{4}$. $\square$

Proposition 6.4.7. *For any symmetric generating set S of G or G', there is a non-trivial element of degree at most $C(\log n + \log q)^3$ for some absolute constant C, of length less than $q^{C'n(\log n+\log q)^3}$ for some absolute constant C'.*

Proof. We pick r and c according to Lemma 6.2.3. Let us assume that $2c(\log n+\log q)^3 < \frac{n-2}{5}$, because otherwise the statement is trivial.

Let $p_1, ..., p_r$ be the first r primes coprime to $p(q-1)$. Let $f_i(x)$ be the irreducible polynomial over $\mathbb{F}_q$ for all the primitive p_i-th roots of unity, and let C_i be the companion matrix of $\mathbb{F}_q$.

Initial Step:

Let us find our first P(r)-matrix. Let T be a singularly $c(\log n + \log q)^3$-transversal set. Let W be any totally singular subspace of dimension $c(\log n + \log q)^3$, which exists by Witt's decomposition theorem. Note that any bijective linear map from W to W is an isometry, and is therefore subject to Witt's extension lemma.

By definition of a singularly transversal set, we can find $A_0 \in T$ that maps the totally singular subspace W onto itself, and that its restriction to this subspace is the matrix $(\bigoplus_{i=1}^{r} C_i) \oplus I$ for some arbitrary choices of basis on W, where I is some identity matrix of suitable size.

In particular, A_0 is a P(r)-matrix. Since $A_0 \in T$, by choosing T as in Corollary 5.3.8, A_0 has length bounded by $q^{cn(\log n+\log q)^3}$.

By using Lemma 6.2.3, we can raise A_0 to a large power, and obtain a non-identity matrix A_1 of degree $\leq \frac{\deg(A_0)}{4} \leq \frac{n}{4}$, with eigenvalues either 1 or outside of $\mathbb{F}_q$. Since the order of A_0 is bounded by q^n by Lemma 6.3.4, the length of A_1 is bounded by $q^{cn(\log n+\log q)^3+n}$.

Inductive Step:

Suppose we have obtained a non-identity matrix A_j with eigenvalues either 1 or outside of $\mathbb{F}_q$, degree at most $\frac{n}{2^{j+1}}$, and length at most $q^{2cn(\log n+\log q)^3+j(n+2)}$. If $\deg A_j \leq 56+32c(\log n+\log q)^3$, then we stop. If not, then let us construct a non-identity matrix A_{j+1} of even smaller degree.

First we shall transform A_j into a P(r)-matrix. Find a totally singular subspace W_j of dimension $c(\log n + \log q)^3$ as in Lemma 6.4.6. In particular, $W_j \oplus A_j W_j$ is a well-defined totally singular space. Let T' be a singularly $2c(\log n + \log q)^3$-transversal set, then we can find $M_j \in T'$ that fixes $A_j W_j$, and restricts to a map from W_j to W_j as $\bigoplus_{i=1}^r C_i \oplus I$ for any arbitrary basis of W_j and some identity matrix I of suitable size.

Consider the commutator $M_j A_j^{-1} M_j^{-1} A_j$. Since M_j fixes $A_j W_j$, we see that $M_j A_j^{-1} M_j^{-1} A_j$ restricted to W_j is identical to M_j restricted to W_j.

In particular, $M_j A_j^{-1} M_j^{-1} A_j$ is a P(r)-matrix, and it has degree at most $2\deg(A_j)$. Now we use Lemma 6.2.3 again, raising $M_j A_j^{-1} M_j^{-1} A_j$ to a large power, and we would obtain a matrix A_{j+1} of degree at most $\frac{2\deg(A_j)}{4}$, with eigenvalue wither 1 or outside of $\mathbb{F}_q$.

Since $M_j \in T'$, by choosing T' as in Corollary 5.3.8, M_j have length bounded by $q^{2cn(\log n+\log q)^3}$. And since the order of $M_j A_j^{-1} M_j^{-1} A_j$ is bounded by q^n by Lemma 6.3.4, the length of A_{j+1} is at most

$$q^n(2q^{2cn(\log n+\log q)^3} + 2q^{2cn(\log n+\log q)^3+j(n+2)}) \leq q^{2cn(\log n+\log q)^3+(j+1)(n+2)}.$$

We repeat the above induction $\frac{\log n}{\log 2} - 1$ times, or stop early if we hit degree $2c(\log n + \log q)^3$. The last A_j we obtained is the desired matrix of small degree and small length. $\square$

6.5 Diameter bounds, spectral gaps and mixing time

Using results from this chapter, we can obtain the following result on diameter bounds of finite simple group of Lie type.

Corollary 6.5.1. *The diameter of a finite simple group of Lie type of rank n over $\mathbb{F}_q$ are at most $O(q^{O(n(\log n+\log q)^3)})$, independent of the choice of generating sets. The implied constants are absolute.*

Proof. Combine Proposition 5.4.5 with Proposition 6.3.5 or Proposition 6.4.7. $\square$

Given a group G and its generating set S, let $\Gamma(G,S)$ be its Cayley graph, and let A be the normalized adjacency matrix of the graph. Then A has real eigenvalues $\lambda_1, ..., \lambda_{|G|}$, ordered from the largest one to the smallest one. Then the ***spectral gap*** of $\Gamma(G,S)$ is $\lambda_1 - \lambda_2$.

Let μ be the random distribution $\frac{1}{2}1_{\{e\}} + \frac{1}{2|S|}1_S$. Then a lazy random walk of length k is the random outcome of the distribution $\mu^{(k)} = \mu * \mu * \mu * ... * \mu$. Using the definition of Helfgott, Seress and Zuk [HSZ15], the ***strong mixing time*** of $\Gamma(G,S)$ is the least number k such that $\mu^{(k)}$ is at most $\frac{1}{2|G|}$ away from the uniform distribution on $\Gamma(G,S)$, in the ℓ^∞ norm.

One can bound the spectral gap using a diameter bound.

Proposition 6.5.2 (([DS93], Corollary 3.1))**.** *Given a finite group G and a symmetric generating set S, let Γ be the Cayley graph. Then the spectral gap of the Cayley graph is bounded from below by $\frac{1}{(\operatorname{diam}\Gamma)^2|S|}$*

In turn, one can bound the strong mixing time by the spectral gap.

Proposition 6.5.3 (([Lov93], Theorem 5.1))**.** *Given a finite group G and a symmetric generating set S, let Γ be the Cayley graph, and let λ be the spectral gap. Then the strong mixing time of the Cayley graph is bounded by $O(\frac{\log|\Gamma|}{\lambda})$.*

Then our main result implies the following corollary:

Corollary 6.5.4. *Let G be a finite simple group of Lie type of rank n over $\mathbb{F}_q$. The spectral gap of $\Gamma(G,S)$ is bounded by $|S|^{-1}q^{-O(n(\log n+\log q)^3)}$, and the mixing time of $\Gamma(G,S)$ is bounded by $|S|q^{O(n(\log n+\log q)^3)}$.*

REFERENCES

[Asc00] M. Aschbacher. *Finite group theory*, volume 10. Cambridge University Press, 2000.

[BGG15] E. Breuillard, B. Green, R. Guralnick, and T. Tao. "Expansion in finite simple groups of Lie type." *J. Eur. Math. Soc. (JEMS)*, **17**, 2015.

[BGP97] L. Babai, A. J. Goodman, and L. Pyber. "Groups without Faithful Transitive Permutation: Representations of Small Degree." *J. Algebra*, **195**:1–29, 1997.

[BGT10] E. Breuillard, B. Green, and T. Tao. "Linear Approximate Groups." *arXiv:1001.4570*, 26 Jan 2010.

[BGT11a] E. Breuillard, B. Green, and T. Tao. "Approximate subgroups of linear groups." *Geom. Funct. Anal.*, **21(4)**:774–819, 2011.

[BGT11b] E. Breuillard, B. Green, and T. Tao. "Suzuki groups as expanders." *Groups Geom. Dyn.*, **5**, 2011.

[Bre78] J. L. Brenner. "Covering Theorems for FINASIGs. VlII. Almost All Conjugacy Classes in A_n Have Exponent ≤ 4." *J. Aust. Math. Soc. Ser. A*, **25(02)**(210-214), 1978.

[BS88] L. Babai and Á. Seress. "On the diameter of Cayley graphs of the symmetric group." *J. Combin. Theory*, **49(1)**:175–179, 1988.

[BS92] L. Babai and Á. Seress. "On the diameter of permutation groups." *European J. Combin.*, **13(4)**:231–243, 1992.

[BT14] V. Bergelson and T. Tao. "Multiple Recurrence in Quasirandom Groups." *Geom. Funct. Anal.*, **24(1)**:1–48, 2014.

[BT16] E. Breuillard and M. Tointon. "Nilprogressions and groups with moderate growth." *Adv. Math.*, **289**:1008–1055, 2016.

[BY17] A. Biswas and Y. Yang. "A diameter bound for finite simple groups of large rank." *J. Lond. Math. Soc.*, **95(2)**:455–474, 2017.

[CE75] J. Cheeger and D. G. Ebin. *Comparison Theorems in Riemannian Geometry.* American Mathematical Society, 1975.

[DS93] P. Diaconis and L. Saloff-Coste. "Comparison techniques for random walk on finite groups." *Ann. Probab.*, **21(4)**:2131–2156, 1993.

[Fou85] É. Fouvry. "The Brun-Titchmarsh theorem: application to the Fermat theorem." *Invent. Math.*, **79(2)**:383–407, 1985.

[Gow08] W. T. Gowers. "Quasirandom Groups." *Combin. Probab. Comput.*, **17(3)**:363–387, 2008.

[Gro02] L. C. Grove. *Classical groups and geometric algebra.* American Mathematical Soc., 2002.

[Hel08] H. A. Helfgott. "Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$." *Ann. of Math. (2)*, **167(2)**:601–623, 2008.

[Her75] I. N. Herstein. *Topics in Algebra 2nd Edition.* Xerox College Publishing, 1975.

[HK02] J. E. Hart and K. Kunen. "Bohr Compactifications of Non-Abelian Groups." *Topology Proc.*, **26(2)**:593–626, 2002.

[Hol64] P. Holm. "On the Bohr Compactification." *Math. Ann.*, **156**:34–46, 1964.

[HP89] D. Holt and W. Plesken. *Perfect Groups.* Oxford: Clarendon Press, 1989.

[HS14] H. A. Helfgott and Á. Seress. "On the diameter of permutation groups." *Ann. of Math. (2)*, **179(2)**:611–658, 2014.

[HSZ15] H. A. Helfgott, Á. Seress, and A. Zuk. "Random generators of the symmetric group: diameter, mixing time and spectral gap." *J. Algebra*, **421**:349–368, 2015.

[Jor78] C. Jordan. "Mémoire sur les equations différentielle linéaire à intégrale algébrique." *J. reine angew. Math.*, **84**:89–215, 1878.

[Lan02] S. Lang. *Algebra revised third edition.* GTM 211, 2002.

[Los55] J. Łos. "Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres." *Mathematical Interpretations of Formal Systems*, pp. 98–113, 1955.

[Lov93] L. Lovász. "Random walks on graphs: a survey." *Combinatorics, Paul Erdős is eighty*, **2**:353–397, Keszthely, 1993.

[LS01] M. W. Lieback and A. Shalev. "Diameters of finite simple groups: sharp bounds and applications." *Ann. of Math. (2)*, **154(2)**:383–406, 2001.

[PS10] L. Pyber and E. Szabó. "Growth in finite simple groups of Lie type." *arXiv:1001.4556*, 25 Jan 2010.

[PS16] L. Pyber and E. Szabó. "Growth in finite simple groups of Lie type." *J. Amer. Math. Soc.*, **29(1)**:95–146, 2016.

[Sep07] M. R. Sepanski. *Compact Lie Group*, volume 235. Springer Science and Business Media, 2007.

[ST13] A. Stolz and A. Thom. "On the lattice of normal subgroups in ultraproducts of compact simple groups." *Proceedings of the London Mathematical Society*, **108(1)**:73–102, 2013.

[Tit57] J. Tits. "Sur les analogues algébriques des groupes semi-simples complexes." *Colloque d'algèbre supérieure*, pp. 261–289, 1957.

[Yan16] Y. Yang. "Ultraproducts of quasirandom groups with small cosocles." *J. Group Theory*, **19(6)**:1137–1164, 2016.