Fast Matrix Multiplication by Group Algebras

A Master's Thesis

submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Master of Science

by

Zimu Li

January 24, 2018

Approved

Padraig Ó Catháin
Thesis Advisor

Professor Luca Capogna
Department Head

Abstract

Based on *Cohn* and *Umans'* group-theoretic method, we embed matrix multiplication into several group algebras, including those of cyclic groups, dihedral groups, special linear groups and Frobenius groups. We prove that $SL_2(\mathbb{F}_p)$ and $PSL_2(\mathbb{F}_p)$ can realize the matrix tensor $\langle p, p, p \rangle$, i.e. it is possible to encode $p \times p$ matrix multiplication in the group algebra of such a group. We also find the lower bound for the order of an abelian group realizing $\langle n, n, n \rangle$ is n^3. For Frobenius groups of the form $C_q \rtimes C_p$, where p and q are primes, we find that the smallest admissible value of q must be in the range $p^{4/3} \leq q \leq p^2 - 2p + 3$. We also develop an algorithm to find the smallest q for a given prime p.

Key words: fast matrix multiplication, group-theoretic method, representation theory, cyclic group, dihedral group, special linear group, Frobenius group.

Acknowledgement

This work would not have been possible without the support of the Department of Mathematical Science, Worcester Polytechnic Institute. I am especially indebted to Dr. Padraig Ó Catháin , Assistant Professor of the Department of Mathematical Science, who gave me numerous advises along the whole research. Also, I am grateful to Michael D Malone and Kyle George Dunn who gave me technical support.

Contents

Chapter 1

Introduction

A natural question in computer science is to bound the computational complexity of standard mathematical tasks. This thesis is concerned with the complexity of matrix multiplication. In many models of computation, multiplication is much more expensive then addition. So for simplicity, we only count the number of multiplications when we measure the time complexity of a computational task. We will use the $\mathcal{O}$ notation to help us measure the number of multiplications. For two functions $f(n)$, $g(n)$: $\mathbb{N} \to \mathbb{N}$, we write $f(n) = \mathcal{O}(g(n))$ if there exits $c \in \mathbb{R}$ such that $f(n) \leq cg(n)$ for all sufficiently large n. In our case, f is the number of multiplications in an algorithm for $n \times n$ matrix multiplication.

Given two $n \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$, where a_{ij} and b_{ij} are in field $F = \mathbb{C}$. In general cases, F could be any field, we use $\mathbb{C}$ in our research to simplify the problem. We want to calculate the product of $AB = (c_{ij})$. The naive matrix multiplication algorithm is:

$$c_{ij} = \sum_{m=1}^{n} a_{im} b_{mj}$$

This algorithm takes n multiplication to calculate each entry. Thus it takes n^3 multiplications to compute AB.

Definition 1. Suppose that a matrix multiplication algorithm takes about $\mathcal{O}(n^\omega)$ multiplications, then $\mathcal{O}(n^\omega)$ is the time complexity of this algorithm and ω is the complexity exponent.

By this definition, the time complexity (we also use complexity to refer time complexity below) of the naive algorithm is $n^3 = \mathcal{O}(n^3)$.

In 1969, *Volker Strassen* found the first fast matrix multiplication algorithm in [14] which has complexity exponent smaller then 3.

Assume $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$, $B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$ and $AB = \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}$. Compute the following seven products:

$$\begin{aligned}
p_1 &= (a_1 + a_4)(b_1 + b_4) \\
p_2 &= (a_3 + a_4)b_1 \\
p_3 &= a_1(b_2 - b_4) \\
p_4 &= a_4(b_3 - b_1) \\
p_5 &= (a_1 + a_2)b_4 \\
p_6 &= (a_2 - a_1)(b_1 + b_2) \\
p_7 &= (a_2 - a_4)(b_3 + b_4)
\end{aligned}$$

This seven products surprisingly give us all entries of AB as their linear combination:

$$AB = \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} = \begin{pmatrix} p_1 + p_4 - p_5 + p_7 & p_3 + p_5 \\ p_2 + p_4 & p_1 + p_3 - p_2 + p_6 \end{pmatrix}$$

It uses 7 multiplications instead of 8 to calculate 2×2 matrix multiplication and 7 is also the optimal number for 2×2 matrix multiplication [2]. The optimal number of multiplication for 3×3 matrix multiplication is somewhere between 19 and 23. The larger the matrix is, the harder to find the optimal number of multiplication. However we can apply *Strassen algorithm* to $n \times n$ matrix multiplication by regard $n \times n$ matrix as 2×2 block matrix.

Theorem 1 (Proposition 1.1 in [2]). *One can multiply $n \times n$ matrices with $\mathcal{O}(n^{\log_2 7})$ multiplication.*

Later in 1987, *Strassen* improved the complexity from $\omega < \mathcal{O}(n^{2.81})$ to $\omega < \mathcal{O}(n^{2.48})$ using *laser* method [15]. However it is still not the lowest upper bound of ω. A variant *Strassen's algorithm* from *Coppersmith* and *Winograd* makes a great improvement to $\omega < \mathcal{O}(n^{2.376})$ [5] in 1990. This number stood as the best upper bound of ω for more then 20 years before *Virgina V. Williams* set the new record as $\omega < \mathcal{O}(n^{2.373})$ in 2014 [17]. Many researchers believe that for every $\epsilon > 0$ there exists a $N_\epsilon > 0$ such that matrices of size larger then N_ϵ can be multiplied in $\mathcal{O}(n^{2+\epsilon})$.

All algorithms above are based on *Strassen's Algorithm*, however, *Henry Cohn* and *Christopher Umans* developed a group-theoretic approach to bound the complexity exponent of matrix multiplication [4]. They embedded matrix multiplication into group algebras and accelerated the calculation by decomposing the corresponding representations. They used the *pseudo-exponent* to measure the complexity and they also showed how to match the bound $\omega < \mathcal{O}(n^{2.376})$ using group-theoretic method [3]. Since their approach is relatively simple and almost entirely separate from *Strassen's Algorithm*, our research is based on group-theoretic approach. Instead of focus on finding the bound of the exponent ω, we look into several type of groups and try to find the smallest group to embed matrix multiplication and also try to measure the efficiency.

Chapter 2

Representation theory of abelian groups

In this chapter, we will introduce the representation theory of abelian groups with some basic definitions and some theorems used in our research.

Definition 2. A representation of a group G on a vector space V over a field F is a group homomorphism from G to $GL(V)$. That is, a representation is a map $\rho : G \to GL(V)$ such that,

$$\rho(g_1 g_2) = \rho(g_1)\rho(g_2), \; for \; all \; g_1, g_2 \in G.$$

The dimension of V is called the dimension of the representation.

Definition 3. Let $\mathbb{C}$ be the complex field and G be a finite group. The group algebra $\mathbb{C}G$ is the set of all linear combinations of finitely many elements of G with coefficients in $\mathbb{C}$.

Definition 4. Let $\mathbb{C}G$ be an group algebra and V be a finite dimensional complex-vector space. Suppose for every $v \in V$ and $x \in \mathbb{C}G$ that a unique $vx \in V$ is defined. Assume for all $x, y \in \mathbb{C}G$, $v, w \in V$ and a complex number c that

1. $(v + w)x = vx + wx$
2. $v(x + y) = vx + vy$
3. $(vx)y = v(xy)$
4. $(cv)x = c(vx) = v(cx)$

5. $v1 = v$

Then V is a $\mathbb{C}G$-module.

Let G be a finite group and V be a $\mathbb{C}G$-module, then the map: $G \to GL(V)$ given by $g \to \rho_g$, where $\rho_g(v) = vg$, defines a representation of G on V. On the other hand, given a representation ρ: $G \to GL(V)$ we have a linear action of G on V given by $vg = v\rho(g)$.

Definition 5. A $\mathbb{C}G$-module V is said to be *irreducible* if it is non-zero and it has no $\mathbb{C}G$-module apart from $\{0\}$ and V. If V has an $\mathbb{C}G$-submodule W which is not $\{0\}$ or V, then V is *reducible*. A representation $\rho : G \to GL(n, \mathbb{C})$ is *irreducible* if the corresponding $\mathbb{C}G$-module V given by

$$vg = v(g\rho) \ v \in V, \ g \in G$$

is irreducible; and ρ is reducible if V is reducible.

Definition 6. Let G be a finite group and $\mathbb{C}$ be the complex field. The representation $g \to [g]_B$ obtained by taking B to be the natural basis of $\mathbb{C}G$ is called the *regular representation* of G over $\mathbb{C}$.

Definition 7. Given $\mathbb{C}G$-modules V and W, for $c \in \mathbb{C}$, $v_1, v_2 \in V$ and $w_1, w_2 \in W$, define the operations as follows:

1. $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$

2. $c(v_1, w_1) = (cv_1, cw_1)$

Then $\{(v, w) : v \in V, \ w \in W\}$ is a $\mathbb{C}G$-module called the direct sum of V and W, denoted by

$$V \oplus W.$$

Definition 8. Given groups $(G, *)$ and $(H, \triangle)$, the direct product $G \times H$ is defined as follows:

1. $G \times H = \{(g, h) : g \in G, h \in H\}$

2. The option on $G \times H$ is defined component-wise:

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 * g_2, h_1 \triangle h_2)$$

   where $g_1, g_2 \in G$ and $h_1, h_2 \in H$.

$(G \times H, \cdot)$ satisfies the axioms for group.

In the following paragraphs we will prove that every $\mathbb{C}G-module$ of finite abelian group G with dimension n is direct sum of n *irreducible* $\mathbb{C}G-module$ with dimension 1. Which also means that the regular representation matrix of every $\mathbb{C}G - module$ for finite abelian group G is diagonalizable.

Theorem 2. *If G is a finite abelian group, then every irreducible $\mathbb{C}G-module$ has dimension 1.*

Proof. Let G be a finite abelian group, and V be an irreducible $\mathbb{C}G-module$. And let ρ be a representation: $\rho\ G \to GL(V)$ such that, $\rho(g_1 g_2) = \rho(g_1)\rho(g_2),\ for\ all\ g_1, g_2 \in G$. Since $\rho(g_1) \in GL(V)$, suppose that $\lambda \in \mathbb{C}$ is the eigenvalue of $\rho(g_1)$ with eigenvector $v \in V$. Left multiplying by $\rho(g_2)$ on both sides, we have

$$\begin{aligned} \rho(g_2)\rho(g_1)v &= \lambda\rho(g_2)v \\ &= \rho(g_1)\rho(g_2)v \end{aligned} \tag{2.1}$$

since G is abelian. Therefore $\rho(g_2)v$ is also a eigenvector of $\rho(g_1)$. Since g_2 can be any element in G, $\rho(g_1)$ act like a complex scalar and $\dim \rho = 1$. It also means that λ-eigenspace is a $\mathbb{C}G$-submodule of V which dimension is equal to 1. Since V is a irreducible $\mathbb{C}G$-module, $\dim V = 1$. □

In the following paragraphs, we will prove that regular representations of finite abelian groups are diagonalizable.

Theorem 3 (Chapter 9 in [7])**.** *Every finite abelian group is isomorphic to a direct product of cyclic groups.*

Theorem 4 (Chapter 8 in [11])**.** *If G is a finite group and field F is $\mathbb{C}$, then the $\mathbb{C}G$-module V can be decompose as:*

$$V = U_1 \oplus U_2 \oplus ... \oplus U_m$$

where U_i are irreducible $\mathbb{C}G$-submodules.

Theorem 5. *Every $\mathbb{C}G - module$ of finite abelian group G with dimension n is a direct sum of n irreducible $\mathbb{C}G$-submodules with dimension 1.*

Proof. Let V be a $\mathbb{C}G - module$ of finite abelian group G. According to *Theorem* 4 We can decompose V as:

$$V = U_1 \oplus U_2 \oplus ... \oplus U_m$$

where U_i are some irreducible $\mathbb{C}G - module$. By *Theorem* 2, $dimU_i = 1$ for $i = 1, 2, ...m$ □

Corollary 1. *The regular representation matrix of every $\mathbb{C}G - module$ for finite abelian group G is diagonalizable.*

Proof. According to *Theorem* 11, we can decompose every finite abelian group G as

$$G = C_{n_1} \times C_{n_2} \times C_{n_3} \times ... \times C_{n_m}$$

where C_{n_i} is a cyclic group generated by c_i of order n_i. Let

$$g_i = (1, 1, ..., c_i, ..., 1) \ where \ c_i \ is \ in \ ith \ position.$$

Then we have $G = \langle g_1, g_2, ..., g_m \rangle$, with $g_i^{n_i} = 1$ and $g_i g_j = g_j g_i$ for all i, j. Let $\theta : G \to GL(n, \mathbb{C})$ be an irreducible representation of G. By *Theorem* 2, $n = 1$. Then for every g_i we have:

$$\theta(g_i) = (\lambda_i) \ where \ \lambda_i \in \mathbb{C}$$

And since $g_i^{n_i} = 1$ and $\lambda_i^{n_i} = 1$. For $\forall g \in G$, we have $g = g_1^{i_1} g_2^{i_2} ... g_m^{i_m} (i_r \ is \ integer)$, which deduce:

$$\theta(g) = \theta(g_1^{i_1} g_2^{i_2} ... g_m^{i_m}) = (\lambda_1^{i_1} \lambda_2^{i_2} ... \lambda_m^{i_m})$$

where λ_i is an n_i^{th} root of unity. There are $n_1 n_2 ... n_m = n$ of such irreducible representations,and no two of them are equivalent. Let θ_i denote such irreducible representations. Let $\rho : V \to GL(n, \mathbb{C})$ be the regular representation of $\mathbb{C}G - module$, then since *Theorem* 5 we have:

$$\rho(v) \ is \ a \ linear \ transformation \ from \ \theta(x_1) \oplus \theta(x_2) \oplus ... \oplus \theta(x_n).$$

Where x_i is distinct elements in G. Which also means $\exists n \times n \ matrix \ M$ such that

$$M^{-1} \cdot \rho(v) \cdot M = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

□

Chapter 3

Embedding polynomial multiplication in a group algebra

In this chapter, we will embed a subset of polynomial ring $\mathbb{C}[x, y]$ into the group algebra $\mathbb{C}G$ for a suitably chosen abelian group G. Efficient multiplication of $\mathbb{C}G$ elements gives an algorithm for multiplication of polynomials in subquadratic time. The matrix multiplication embedding can just analogise the polynomial multiplication embedding.

Let $P_1(x, y)$ and $P_2(x, y)$ be defined as follow:

$$P_1 = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} a_{ij} \cdot x^i \cdot y^j$$

$$P_2 = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} b_{ij} \cdot x^i \cdot y^j$$

And let $G = C_{2m-1} \times C_{2n-1}$ be a finite abelian group and $\mathbb{C}G$ be the group algebra. Assume $C_{2n-1} = \langle c_1 \rangle$ and $C_{2m-1} = \langle c_2 \rangle$. Given the partial embedding $\phi : \mathbb{C}[x, y] \to \mathbb{C}G$ as follow:

$$\phi(P_1) = \sum_{i=0}^{n-1} \sum_{i=0}^{m-1} a_{ij} \cdot c_1^i \cdot c_2^j$$

$$\phi(P_2) = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} b_{ij} \cdot c_1^i \cdot c_2^j$$

We can easily conclude that the coefficient of $x^i y^j$ in $P_1 P_2$ is equal to the coefficient of $c_1^i c_2^j$ in $\phi(P_1)\phi(P_2)$. Therefore, in order to calculate the polynomial multiplication, all we need to do is calculate every coefficient of $\phi(P_1)\phi(P_2)$. We will use regular representation of the group algebra to calculate $\phi(P_1)\phi(P_2)$.

Assume $[c_1^i \cdot c_2^j]_B$ denotes the regular representation of group element $c_1^i \cdot c_2^j$ in G. Then:

$$\rho(\phi(P_1)) = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} a_{ij}[c_1^i \cdot c_2^j]_B$$

$$\rho(\phi(P_2)) = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} b_{ij}[c_1^i \cdot c_2^j]_B$$

Let x_{ij} be an entry of $\rho(\phi(P_1)) \cdot \rho(\phi(P_2))$. Then x_{ij} is equal to the coefficient of the term $c_1^i c_2^j$ in $\phi(P_1)P_2^*$ which is also the coefficient of $x^i y^j$ in P_1P_2.

Corollary 1 shows that $\rho(\phi(P_1))$ and $\rho(\phi(P_2))$ are diagonalizable, and we can use *fast Fourier transform* (*FFT*) to diagonlize them. A *FFT* is an algorithm that can computes *discrete Fourier transform* (*DFT*) in $\mathcal{O}(nlogn)$ time. And we can use the matrix form of *DFT* to diagonalize $\rho(\phi(P_1))$ and $\rho(\phi(P_2))$. Since our work is focus on group-theoretic methods, we will not go into *FFT* and *DFT*. You can find more details about them in [9].

Theorem 6. *If G is a finite abelian group of order n, then we can multiply $\alpha, \beta \in \mathbb{C}G$ in time $\mathcal{O}(n \log n)$.*

Proof. Let G be a finite abelian group and $\alpha, \beta \in \mathbb{C}G$. Suppose ρ is the regular representation in G. By *Corollary* 1, $\rho(\alpha)$ and $\rho(\beta))$ are diagonalizable. And since $\rho(\alpha)$ and $\rho(\beta)$ use the same representation, there is a matrix M diagonalize both of them. Then we can calculate $\rho(\alpha) \cdot \rho(\beta)$ as follow:

$$\begin{aligned}\rho(\alpha) \cdot \rho(\beta) &= M^{-1}\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} M M^{-1} \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix} M \\ &= M^{-1}\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix} M\end{aligned}$$

Using *FFT*, we can diagonalize $\rho(\alpha)$ and $\rho(\beta)$ in $\mathcal{O}(n \log n)$ time. The complexity of multiply two diagonalized matrices is $\mathcal{O}(n)$. Then we can conclude that the complexity of $\alpha \cdot \beta$ is $\mathcal{O}(n \log n)$. □

By *Theorem* 6, the group-theoretic methods of the fast polynomial multiplication reduce two-variable polynomial multiplication complexity from $\mathcal{O}(m^2n^2)$ to $\mathcal{O}(mn\log mn)$

Chapter 4

Embedding Matrix Multiplication in a Group Algebra

We will explain how to embed the $n \times n$ matrices A, B into the group algebra $\mathbb{C}G$. Given subsets S_1, S_2, S_3 of G, $|S_i| = n$, let

$$A^* = \sum_{s_1 \in S_1, s_2 \in S_2} s_1^{-1} \cdot s_2 \cdot A_{s_1 s_2}$$

$$B^* = \sum_{s_2 \in S_2, s_3 \in S_3} s_2^{-1} \cdot s_3 \cdot A_{s_2 s_3}$$

We use elements in S_1, S_2 to label the rows and columns of A and use S_2, S_3 to label the rows and columns of B.

Example 1. Let $G = C_2 \times C_2 \times C_2$, and this three cyclic groups of order 2 are generated by x, y, z respectively. We will give a simple example of embedding 2×2 matrices multiplication into G.

Let $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$, $B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$ and $S_1 = \{1, x\}, S_2 = \{1, y\}, S_3 = \{1, z\}$.

First we label the first row of A as 1 and second row of A as $x \in S_1$, the first column of A as 1 and second column of A as $y \in S_2$. Similarly, we label the first row of B as 1 and second row of B as $y \in S_2$, the first column of B as z and second column of A as $z \in S_3$. Then, for instance, $a_1 = A_{11}$, $a_2 = A_{1y}$, $a_3 = A_{x1}$ and $a_4 = A_{xy}$. We also can label rows of AB as $1, x$ and column of AB as $1, z$(we will prove it later).

Then we embed A and B into $A^*, B^* \in \mathbb{C}G$:

$$A^* = 1 \cdot 1 \cdot a_1 + 1 \cdot y \cdot a_2 + x^{-1} \cdot 1 \cdot a_3 + x^{-1} \cdot y \cdot a_4 = a_1 + a_2 y + a_3 x + a_4 xy$$

$$B^* = 1 \cdot 1 \cdot b_1 + 1 \cdot z \cdot b_2 + y^{-1} \cdot 1 \cdot b_3 + y^{-1} \cdot z \cdot b_4 = b_1 + b_2 z + b_3 y + b_4 yz$$

We need to find S_1, S_2, S_3 for G before we embed matrices and the following property gives a guideline of finding them.

Definition 9 (Triple-product property)**.** Suppose $|S_1| = n, |S_2| = m, |S_3| = p$ are three subsets of group G. Let $Q_i = \{sv^{-1} | s, v \in S_i\}$ for $i \in \{1, 2, 3\}$. For every $q_i \in Q_i$, $q_1 \cdot q_2 \cdot q_3 = 1$ if and only if $q_i = 1$. If such S_1, S_2, S_3 satisfy *triple-product property* in G, then we say that G realize $\langle n, m, p \rangle$.

If all S_1, S_2, S_3 are subgroups of group G, we can check whether it satisfy *triple-product property* in a more straightforward way.

Theorem 7. *Suppose S_1, S_2, S_3 are three subgroups of group G that satisfy triple-product property. Then for every $x_i \in S_i$, $x_1 \cdot x_2 \cdot x_3 = 1$ if and only if $x_i = 1$.*

Proof. Let S_1, S_2, S_3 be subgroups of G, then $Q_i = \{sv^{-1} | s, v \in S_i\} = S_i$ since S_i are subgroups. By the definition of the *triple-product property*, for every $x_i \in S_i$, $x_1 \cdot x_2 \cdot x_3 = 1$ if and only if $x_i = 1$ implies that S_1, S_2, S_3 satisfy *triple-product property.* □

Corollary 2. *Suppose S_1, S_2, S_3 are three subgroups of group G that satisfy triple-product property. Let $T = \{s_1 s_2 | s_1 \in S_1,\ s_2 \in S_2\}$, then $|T \cap S_3| = 1$*

Proof. It is trivial that $1 \in T \cap S_3$. Assume that $|T \cap S_3| > 1$, let $x \in T \cap S_3$ be an non-trivial element. Then $x^{-1} \in S_3$. Since $x \in T$, there exits $s_1 \in S_1$ and $s_2 \in S_2$ such that $s_1 s_2 = x$ and $s_1 s_2 x^{-1} = 1$. $x^{-1} \neq 1$ contradict *Theorem* 7. Then $|T \cap S_3| = 1$. □

We embed matrices to group algebras since we want that the product of A^* and B^* can somehow give us all the entry of $A \cdot B$. As long as S_1, S_2, S_3 satisfy the *triple-product property*, $A^* \cdot B^*$ will give all the information we need.

Theorem 8. *If subsets S_1, S_2, S_3 of G satisfy the triple-product property, then the entry $A \cdot B_{s_1 s_3}$ is equal to the coefficient of $s_1^{-1} s_3$ in $A^* \cdot B^*$.*

Proof.

$$A^* \cdot B^* = \sum_{s_1 \in S_1} \sum_{s_2, v_2 \in S_2} \sum_{s3 \in S_3} s_1^{-1} \cdot s_2 \cdot v_2^{-1} \cdot s_3 \cdot A_{s_1 s_2} \cdot B_{v_2 s_3}.$$

If $s_1^{-1} \cdot s_2 \cdot v_2^{-1} \cdot s_3 = s_1^{-1} \cdot s_3$, we have term $s_1^{-1} s_3$ in $A^* \cdot B^*$ and the coefficient of $s_1^{-1} s_3$ is

$$\sum_{s_1 \in S_1} \sum_{s_2, v_2 \in S_2} \sum_{s3 \in S_3} A_{s_1 s_2} \cdot B_{v_2 s_3}.$$

By the definition of the *triple-product property*

$$
\begin{aligned}
&s_1^{-1} \cdot s_2 \cdot v_2^{-1} \cdot s_3 = s_1^{-1} \cdot s_3 \\
\Rightarrow &s_1 \cdot s_1^{-1} \cdot s_2 \cdot v_2^{-1} \cdot s_3 \cdot s_3^{-1} = 1 \\
\Rightarrow &s_2^{-1} \cdot v_2 = 1
\end{aligned} \tag{4.1}
$$

Then the coefficient of $v_1^{-1}v_3$ equal to $\sum_{s_1 \in S_1} \sum_{s_2, v_2 \in S_2} \sum_{s3 \in S_3} A_{s_1 s_2} \cdot B_{s_2 s_3}$ which is $A \cdot B_{s_1 s_3}$. □

In the following paragraphs, we will show a example of embedding multiplication of 2×2 matrix into dihedral group.

Definition 10. A dihedral group is the group of symmetries of a regular polygon. The dihedral group of a regular n-side polygon is

$$D_{2n} = \{r, s | r^n = s^2 = 1, srs = r^{-1}\}$$

where r is the rotation symmetry of order n and s is the reflection symmetry. The dihedral group of a regular n-side polygon has order $2n$.

Let $D_8 = \{1, r, r^2, r^3, s, rs, r^2s, r^3s\}$. Let $S_1 = \{1, s\}, S_2 = \{1, rs\}, S_3 = \{1, r^2s\}$, then we have $s_1 \cdot v_1^{-1} \cdot s_2 \cdot v_2^{-1} \cdot s_3 \cdot v_3^{-1} = 1$ if and only if $s_i \cdot v_i^{-1} = 1$ satisfy the *triple-product property*. We can embed matrix multiplication to S_1, S_2, S_3. Let

$$A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$$

$$B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$$

Let elements in S_1 and S_2 represent the row and column of A respectively and S_2 and S_3 represent the row and column of B. Then we have

$$A^* = a_1 \cdot 1 \cdot 1 + a_2 \cdot 1 \cdot rs + a_3 \cdot s \cdot 1 + a_4 \cdot s \cdot rs$$

$$B^* = b_1 \cdot 1 \cdot 1 + b_2 \cdot 1 \cdot r^2s + b_3 \cdot rs \cdot 1 + b_4 \cdot rs \cdot r^2s$$

We can calculate the entries of AB by calculate corresponding coefficient of $A^* \cdot B^*$

$$\begin{aligned}
A^* \cdot B^* &= a_1 \cdot b_1 + r^2 s \cdot a_1 \cdot b_2 + rs \cdot a_1 \cdot b_3 + rs \cdot r^2 s \cdot a_1 \cdot b_4 \\
&\quad + rs \cdot a_2 \cdot b_1 + rs \cdot r^2 s \cdot a_2 \cdot b_2 + rs \cdot rs \cdot a_2 \cdot b_3 + rs \cdot rs \cdot r^2 s \cdot a_2 \cdot b_4 \\
&\quad + s \cdot a_3 \cdot b_1 + s \cdot r^2 s \cdot a_3 \cdot b_2 + s \cdot rs \cdot a_3 \cdot b_3 + s \cdot rs \cdot r^2 s \cdot a_3 \cdot b4 \\
&\quad + srs \cdot a_4 \cdot b1 + srs \cdot r^2 s \cdot a_4 \cdot b_2 + srs \cdot rs \cdot a_4 \cdot b_3 + srs \cdot rs \cdot r^2 s \cdot a_4 \cdot b_4 \\
&= (a_1 \cdot b_1 + a_2 \cdot b_3) + r^2 s \cdot (a_1 \cdot b_2 + a_2 \cdot b_4) \\
&\quad + s \cdot (a_3 \cdot b_1 + a_4 \cdot b_3) + r^2 \cdot (a_3 \cdot b_2 + a_4 \cdot b_4) \\
&\quad + rs \cdot (a_1 \cdot b_3 + a_2 \cdot b_1 + a_3 \cdot b_4 + a_4 \cdot b_2) \\
&\quad + r^3 \cdot (a_1 \cdot b_4 + a_2 \cdot b_2 + a_3 \cdot b_3 + a_4 \cdot b_1)
\end{aligned}$$

Let Φ be a map:$\mathbb{C}G \to \mathbb{C}G$ such that, $\Phi(\sum a_g \cdot g) = \sum_{g \in S_1 \cdot S_3} a_g \cdot g$. Then

$$\begin{aligned}
\Phi(A^* \cdot B^*) &= (a_1 \cdot b_1 + a_2 \cdot b_3) + r^2 s \cdot (a_1 \cdot b_2 + a_2 \cdot b_4) \\
&\quad + s \cdot (a_3 \cdot b_1 + a_4 \cdot b_3) + r^2 \cdot (a_3 \cdot b_2 + a_4 \cdot b_4)
\end{aligned}$$

In this case, terms $(a_1 \cdot b_1 + a_2 \cdot b_3)$, $r^2 \cdot (a_1 \cdot b_2 + a_2 \cdot b_4)$, $r^2 s \cdot (a_1 \cdot b_2 + a_2 \cdot b_4)$ and $s \cdot (a_3 \cdot b_1 + a_4 \cdot b_3)$ are the terms which coefficients provide the entries of $A \cdot B$. Therefore, as long as we have $\Phi(A^* \cdot B^*)$, we will have $A \cdot B$.

By taking the matrices relative to the basis $\{1, r, r^2, r^3, s, rs, r^2 s, r^3 s\}$, we can obtain the regular representation of D_8:

$$\rho(s) = \begin{pmatrix} 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&1 \\ 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \end{pmatrix} \qquad \rho(r) = \begin{pmatrix} 0&0&0&1&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \end{pmatrix} \tag{4.2}$$

The regular representation of A^* and B^* are linear combinations of representation of group elements:

$$\rho(A^*) = a_1 \rho(1) + a_2 \rho(r)\rho(s) + a_3 \rho(s) + a_4 \rho(s)\rho(r)\rho(s)$$

$$\rho(B^*) = b_1 \rho(1) + b_2 \rho^2(r)\rho(s) + b_3 \rho(r)\rho(s) + b_4 \rho^3(r)$$

$$\rho(A^*) = \begin{pmatrix} a_1 & 0 & 0 & a_4 & a_3 & 0 & 0 & a_2 \\ a_4 & a_1 & 0 & 0 & a_2 & a_3 & 0 & 0 \\ 0 & a_4 & a_1 & 0 & 0 & a_2 & a_3 & 0 \\ 0 & 0 & a_4 & a_1 & 0 & 0 & a_2 & a_3 \\ a_3 & 0 & 0 & a_2 & a_1 & 0 & 0 & a_4 \\ a_2 & a_3 & 0 & 0 & a_4 & a_1 & 0 & 0 \\ 0 & a_2 & a_3 & 0 & 0 & a_4 & a_1 & 0 \\ 0 & 0 & a_2 & a_3 & 0 & 0 & a_4 & a_1 \end{pmatrix} \quad \rho(B^*) = \begin{pmatrix} b_1 & b_4 & 0 & 0 & 0 & 0 & b_3 & b_2 \\ 0 & b_1 & b_4 & 0 & b_3 & 0 & 0 & b_2 \\ 0 & 0 & b_1 & b_4 & b_2 & b_3 & 0 & 0 \\ b_4 & 0 & 0 & b_1 & 0 & b_2 & b_3 & 0 \\ 0 & 0 & b_2 & b_3 & b_1 & b_4 & 0 & 0 \\ b_3 & 0 & 0 & b_2 & 0 & b_1 & b_4 & 0 \\ b_2 & b_3 & 0 & 0 & 0 & 0 & b_1 & b_4 \\ 0 & b_2 & b_3 & 0 & b_4 & 0 & 0 & b_1 \end{pmatrix} \tag{4.3}$$

As we can see, both $\rho(A^*)$ and $\rho(B^*)$ are 8×8 matrices which are much bigger then original 2×2 matrices. However, both of them have special properties which enable efficient multiplication. This example is too small to give a speed up; we just use it as a illustration of embedding.

Based on 4, we can decompose the $\mathbb{C}G-module$ A^* and B^* into the direct sum of irreducible $\mathbb{C}G-submodule$. However, since D_8 is not an abelian group, we can not decompose A^* and B^* into $\mathbb{C}G-submodules$ of dimension 1. The following Theorems shows how to decompose $\mathbb{C}G-module$ even G is not abelian.

Theorem 9. *Suppose V is a $\mathbb{C}G-module$ such that:*

$$V = U_1 \oplus U_2 \oplus ... \oplus U_r,$$

where U_i are irreducible $\mathbb{C}G-submodules$. If U is any irreducible $\mathbb{C}G-submodule$, then the number of $\mathbb{C}G-submodules$ U_i with $U_i \cong U$ is equal to dim U.

Theorem 10. *Let $V_1, V_2, ..., V_k$ form a complete set of non-isomorphic irreducible $\mathbb{C}G-module$. Then*

$$\sum_{i=1}^{k} (dimV_i)^2 = |G|$$

Then let us decompose A^* and B^* into direct sum of irreducible $\mathbb{C}G-submodule$. Since $|D_8| = 8 = 1 + 1 + 1 + 1 + 2^2$, we can decompose A^* and B^* as following:

$$A^* = A_1 \oplus A_2 \oplus A_3 \oplus A_4 \oplus A_5 \oplus A_6$$

$$B^* = B_1 \oplus B_2 \oplus B_3 \oplus B_4 \oplus B_5 \oplus B_6$$

Where dimension of $A_1, A_2, A_3, A_4, B_1, B_2, B_3, B_4$ is 1 and dimension of A_5, B_5 is 2. Then we can block diagonalize $\rho(A^*)$ and $\rho(B^*)$.

$$\rho(A^*) = \rho(A_1) \oplus \rho(A_2) \oplus \rho(A_3) \oplus \rho(A_4) \oplus \rho(A_5) \oplus \rho(A_6)$$

$$\rho(B^*) = \rho(B_1) \oplus \rho(B_2) \oplus \rho(B_3) \oplus \rho(B_4) \oplus \rho(B_5) \oplus \rho(B_6)$$

Where $\rho(A_1), \rho(A_2), \rho(A_3), \rho(A_4), \rho(B_1), \rho(B_2), \rho(B_3), \rho(B_4)$ is a complex number (1×1 matirx), and $\rho(A_5), \rho(B_5)$, $\rho(A_6)$, $\rho(B_6)$ are 2×2 irreducible representation of irreducible $\mathbb{C}G - submodules$.

Chapter 5

Lower bounds for the complexity of matrix multiplication using a group algebra

In [4], Henry Cohn and Christopher Umans introduced the *pseudo-exponent* of a group G to measure efficiency of the largest possible matrix multiplication which can be embedded in $\mathbb{C}G$.

Definition 11. The *pseudo-exponent* $\alpha(G)$ (or α) of a non-trivial finite group G is the minimum of

$$\frac{3\log|G|}{\log nmp}$$

over all n, m, p(not all 1) such that G realizes $\langle n, m, p\rangle$.

Example 2. We have already show that D_8 can realize $\langle 2, 2, 2\rangle$. Actually, $\langle 2, 2, 2\rangle$ is the largest nmp that D_8 can realize. In other words, assume D_8 can realize $\langle n, m, p\rangle$ then $nmp \le 8$. Therefore, $\frac{3\log|D_8|}{\log nmp} \ge \frac{3\log|D_8|}{\log 8} = 3$. Then $\alpha(D_8) = 3$.

Lemma 1. *Let α be the pseudo-exponent of a finite group G .Then $2 < \alpha \le 3$.*

Proof. Let S_1, S_2, S_3 be subsets of finite group G and $Q_i = \{sv^{-1} : s, v \in S_i\}$ for $i \in \{1, 2, 3\}$. G always realize $\langle 1, 1, |G|\rangle$ through $S_1 = \{1\}$, $S_2 = \{1\}$ and $S_3 = G$. Thus $\alpha \le \frac{3\log|G|}{\log|G|} = 3$.

As for the lower bound, assume G realize $\langle n, m, p\rangle$ ($nmp > 1$) with S_1, S_2, S_3. According to the definition of the *triple product property*, for any $s_1, v_1 \in S_1$ and $s_2, v_2 \in S_2$, $v_1^{-1}s_2 \ne s_1^{-1}v_2$, which implies that $|G| \ge nm$. Let $T = \{q_1q_2 : q_1 \in Q_1, q_2 \in Q_2\}$, then $T \cap Q_3 = \{1\}$. Therefore, $|G| = nm$ only if $p = 1$, otherwise we need more elements in G to avoid non-trivial intersection. Similary, $|G| \geq mp$ and $|G| \geq np$ with equality only if $n = 1$ or $m = 1$. Thus $|G|^3 > n^2m^2p^2$ and $\alpha > 2$. $\square$

Theorem 11. *If G is a finite abelian group, then $\alpha(G) = 3$.*

Proof. Let G be a finite abelian group, and assume that G realize $\langle n, m, p\rangle$ with subsets $|S_1| = n, |S_2| = m, |S_3| = p$. Define map ϕ: $S_1 \times S_2 \times S_3 \to G$ such that $\phi(s_1, s_2, s_3) = s_1s_2s_3$ where $s_i \in S_i$. We will prove that ϕ is an injection by contradiction.

Assume that ϕ is not an injection and $s_1s_2s_3 = v_1v_2v_3$ where $s_i, v_i \in S_i$. Thus,

$$\begin{aligned} 1 &= s_1s_2s_3(v_1v_2v_3)^{-1} && (5.1) \\ &= s_1v_1^{-1}s_2v_2^{-1}s_3v_3^{-1} && (5.2) \end{aligned}$$

which contradicts the definition of the *triple product property.*

Since ϕ is a injection, $|G| \geq nmp$. Then $\frac{3\log|G|}{\log nmp} \geq 3$ and $\alpha(G) = 3$. $\square$

Recall $\mathcal{O}(n^\omega)$ is the time complexity of the fast matrix multiplication. *Lemma 1* shows that the range of α is similar to the range of ω. We can regard the *pseudo-exponent* as an approximation of ω, and *pseudo-exponent* even can bound ω under specific condition. When embedding a matrix multiplication into a group algebra $\mathbb{C}G$, we convert a problem of multiplying matrices of size $|G|^{1/\alpha}$ into a problem of multiplying a collection of matrices($\mathbb{C}G - modules$) of size d_i. The later needs about $\sum_i d_i^\omega$ multiplications while the former takes about $|G|^{\omega/\alpha}$ multiplications. The following theorem shows that $\sum_i d_i^\omega$ is an approximate upper bound for the complexity of multiplying matrices of size $|G|^{1/\alpha}$.

Theorem 12 ([4])**.** *Suppose that G has pseudo-exponent α, and the irreducible representation degrees of G are d_i. Then*

$$|G|^{\omega/\alpha} \leq \sum_i d_i^\omega$$

$|G|^{1/\alpha}$ is the size of the largest matrix multiplication that can be embedded into $\mathbb{C}G$ and $|G|^{\omega/\alpha}$ is roughly the number of multiplication needed. By *Theorem* 10, $\sum_i (d_i)^2 = |G|$, then $\sum_i d_i^\omega \geq \sum_i (d_i)^2 = |G|$. Thus. we can use α as an approximation of ω

Notice that degrees of irreducible representations are essential to control ω. Here we define γ, so that $|G|^{1/\gamma}$ is the maximum character degree of G.

Corollary 3. *Let G be a finite group. If $\alpha(G) < \gamma(G)$, then*

$$\omega \leq \alpha(\frac{\gamma - 2}{\gamma - \alpha}).$$

Proof. Let $\{d_i\}$ denote the irreducible representation degrees of G. Recall *Theorem* 10, $\sum_i (d_i)^2 = |G|$,

$$\begin{aligned} |G|^{\omega/\alpha} &\leq \sum_i d_i^{\omega-2} d_i^2 \\ &\leq |G|^{(\omega-2)/\gamma} \sum_i d_i^2 \\ &= |G|^{(\omega-2)/\gamma+1} \end{aligned} \tag{5.3}$$

which also suggests that $\omega/\alpha \leq (\omega-2)/\gamma+1$. Then we conclude that $\omega \leq \alpha(\frac{\gamma-2}{\gamma-\alpha})$, if $\alpha(G) < \gamma(G)$. □

We can strictly bound ω with α. However, the condition $\alpha(G) < \gamma(G)$ require that the maximum degree of irreducible representation smaller then $\sqrt[3]{nmp}$. Since this *Corollary* 3 is not sufficient, we can still use α to approximate ω even if $\alpha(G) \geq \gamma(G)$.

Chapter 6

Cyclic groups and dihedral groups

If a cyclic group G can realize $\langle n, m, p\rangle$, then all groups which contain G as subgroup can realize $\langle n, m, p\rangle$.

Definition 12. Define $C_n = \langle a|a^n = 1\rangle$ as the cyclic group of order n, where a is called the generator of C_n, also denote as $C_n = \langle a\rangle$.

Theorem 13. *For every $\langle n, n, n\rangle$, there exist a integer N such that all cyclic groups of order $\geq N$ realize $\langle n, n, n\rangle$. Also $N = \mathcal{O}(n^3)$.*

Proof. Let G be a cyclic group of order N and a be its generator. Assume q_2, q_3 are primes such that $n < q_2$ and $q_3 > (n-1)(1+q_2)$ Let S_1, S_2, S_3 be subsets of G as following:

$$S_1 = \{1, a, a^2, ..., a^{(n-1)}\}$$
$$S_2 = \{1, a^{q_2}, a^{2q_2}, ..., a^{(n-1)q_2}\}$$
$$S_3 = \{1, a^{q_3}, a^{2q_3}, ..., a^{(n-1)q_3}\}$$

Define Q_1, Q_2, Q_3 as following:

$$Q_1 = \{xy^{-1} : x, y \in S_1\} = \{a^{-(n-1)}, ..., 1, a, ..., a^{(n-1)}\}$$
$$Q_2 = \{xy^{-1} : x, y \in S_2\} = \{a^{-(n-1)q_2}, ..., 1, a^{q_2}, ..., a^{(n-1)q_2}\}$$
$$Q_3 = \{xy^{-1} : x, y \in S_3\} = \{a^{-(n-1)q_3}, ..., 1, a^{q_3}, ..., a^{(n-1)q_3}\}$$

Then in order to prove S_1, S_2, S_3 satisfy *triple product property*, we only need to show that for every $x_i \in Q_i$, $x_1x_2x_3 = 1$ only if $x_i = 1$. Since q_2, q_3 are primes, then $x_ix_j = 1(i, j = 1, 2, 3$ and $i \neq j)$ only if $x_i = x_j = 1$. And since $q_3 > (n-1)(1+q_2)$, let $k_1 \in [-n+1, n-1]$, $k_2 \in [-n+1, n-1]$ be 2 integers , we have $-p_3 < k_1 + k_2p_2 < p_3$.

In order to avoid warp, we need $|G| \geq (n-1) + (n-1)q_2 + (n-1)q_3$. Then we conclude that $x_1x_2x_3 = 1$ only if $x_i = 1$, and G realize $\langle n, n, n \rangle$. By *Theorem* 20, the smallest prime larger then n is about $n + \mathcal{O}(\log n)$. So set $q_2 = n + \mathcal{O}(\log n)$, then $q_3 > (n-1)(1+q_2) = (n-1)(1+n+\mathcal{O}(\log n)) = \mathcal{O}(n^2)$ and

$$\begin{aligned}|G| &\geq (n-1) + (n-1)q_2 + (n-1)q_3 \\ &> (n-1) + (n-1)(n+\mathcal{O}(\log n)) + (n-1)(n-1)(1+n+\mathcal{O}(\log n)) \\ &> n^3 - n + (n^2-n)\mathcal{O}(\log n).\end{aligned}$$

Since $\lim_{n\to\infty} \frac{-n+(n^2-n)\mathcal{O}(\log n)}{cn^2/\log n} = 1$, then $|G| \geq n^3 + \mathcal{O}(n^2/\log n)$ which implies $N = n^3 + \mathcal{O}(n^2/\log n)$. $\square$

We have shown that $|G| \geq n^3$. Since the complexity of multiplication in the group algebra is at least $\mathcal{O}(n^3)$, these embedding will not lead to fast matrix multiplication algorithm. By *Theorem* 11, $\alpha(G) = 3$ is a lower bound on the complexity of matrix multiplication.

As for dihedral group, we find the irreducible representation degrees first.

Lemma 2. *[Corollary 21.20 [11]] Let G be a finite group and ρ an irreducible representation of G. Let N be an abelian normal subgroup of G. Then the degree of ρ divides the index $|G : N|$.*

Theorem 14. *The degree of irreducible representation of dihedral groups are 1 or 2.*

Proof. Let $G = D_{2n} = \{r, s | r^n = s^2 = 1, srs = r^{-1}\}$, $N = C_n$ and ρ be an irreducible representation of G. Then N is an abelian normal group of G. By *Lemma* 2, degree of ρ divide $|G : N| = 2$. Then ρ has degree 1 or 2. $\square$

Lemma 3 (Chapter15 in [11], page 152)**.** *The number of conjugacy classes in a group is equal to the number of irreducible representations.*

Let $G = D_{2n} = \{r, s | r^n = s^2 = 1, srs = r^{-1}\}$ be a dihedral group of order $2n$. If n is odd, r^i conjugates only to r^{-i} for $i \in \{1, 2, ..., (n-1)/2\}$ ($sr^is^{-1} = r^{-i}$). Since $r^isr^{-i} = r^{2i}s$, elements in form of r^is are all in one conjugacy class, where $i \in \{0, 1, 2, ..., n-1\}$. Adding the trivial conjugacy class $\{1\}$, there are $(n+3)/2$ conjugacy classes. If n is odd. If n is even, r^i conjugates only to r^{-i} for $i \in \{1, 2, ..., n/2-1\}$ ($sr^is^{-1} = r^{-i}$). However, there is no element pairing $r^{n/2}$, so $\{r^{n/2}\}$ is also a conjugacy class. Since $r^isr^{-i} = r^{2i}s$ and n is even, there are two conjugacy classes $\{r^{2i}s | 0 \leq i \leq (n-2)/2\}$ and $\{r^{2i+1}s | 0 \leq i \leq (n-2)/2\}$. Adding the trivial conjugacy class $\{1\}$, there are $n/2 + 3$ conjugacy classes if n is even.

By *Lemma* 3, we have the number of irreducible representations of dihedral groups. Then, combine it with *Theorem* 10, we have irreducible representation degree as follow:

Degree	Even n	Odd n
1	4	2
2	n-2/2	n-1/2

character degree of dihedral group

According to the table above, we can reduce a matrix multiplication problem into a collection of 2×2 matrix multiplication and several complex number multiplication, which can not only provide *pseudo-exponent* strictly smaller than three but also strictly bound ω by α.

It is trivial that if C_k can realize $\langle n, n, n\rangle$,then D_{2k} can realize $\langle n, n, n\rangle$. However, $|D_{2k}| = 2k$ is about $\mathcal{O}(n^3)$ which can not lead to any efficient embedding. Therefore, we use the following algorithm to check the *triple-product property* for dihedral groups and try to find the smallest D_{2k} realizing $\langle n, n, n\rangle$.
Let $|S_1| = |S_2| = |S_3| = n$ be subsets of dihedral group G, $Q_i = \{sv^{-1}|s, v \in S_i\}$ for $i \in \{1, 2, 3\}$.

for $non-trivial\ x \in Q_1$ **do**
  for $non-trivial\ y \in Q_2$ **do**
    for $non-trivial\ z \in Q_3$ **do**
      if $xyz = 1$ **then**
        G can not realize $\langle n, n, n\rangle$.
      end if
    end for
  end for
end for

This algorithm is very naive and inefficient. Even if we check *triple-product property* from $k = n^2$ to larger k, it still took hours to find the smallest D_{2k} realizing $\langle 3, 3, 3\rangle$ ($k = 14$). The reason is that there are $\binom{2k}{n} \sim \mathcal{O}(\frac{(2k)^n}{n^n})$ subsets of order n in D_{2k}. Then when $k = n^2$, there are $\binom{2k}{n} \sim \mathcal{O}((2n)^n)$ subsets of order n which means the complexity of traversing the subsets of D_{2k} is $\mathcal{O}((2n)^n)$. Therefore, as long as we can not find some methods which can avoid traversing the subsets of D_{2k}, checking *triple-product property* for group through subsets will be very expensive.

The best embed situation is not for $\langle n, n, n\rangle$ but $\langle n, m, p\rangle$. It is given by Marcus Lang in [12]: G can always realize $\langle m, 2, 2\rangle$, where $m \leq \frac{2n}{3}$. Following is a table of best embed situation and *pseudo-exponent* along with γ for some dihedral groups:

n	Best$\langle m, p, q\rangle$	$\alpha(G)$	γ
12	$\langle 8, 2, 2\rangle$	2.75	4.58
13	$\langle 8, 2, 2\rangle$	2.82	4.70
14	$\langle 9, 2, 2\rangle$	2.79	4.81
15	$\langle 10, 2, 2\rangle$	2.77	4.91
16	$\langle 10, 2, 2\rangle$	2.82	5.00
17	$\langle 11, 2, 2\rangle$	2.80	5.09
18	$\langle 12, 2, 2\rangle$	2.78	5.17
19	$\langle 12, 2, 2\rangle$	2.82	5.25
20	$\langle 13, 2, 2\rangle$	2.80	5.32

pseudo-exponent of dihedral group

Chapter 7

Bounds on the smallest group realizing $p \times p$ matrix multiplication

In the definition of the *triple-product property*, S_i can be any subset of the group. When restrict subsets to subgroups, we can use *Theorem* 7 to decide whether a group can realize $\langle n, n, n\rangle$ which is more straightforward.

In this section, we try to find the smallest groups which can realize $\langle p, p, p\rangle$ for prime p with subgroups S_i. The *Sylow Theorems* give fairly detailed information about the maximal *Sylow p-subgroups* of a finite group G. Then we can come up with a lower bound for order of the groups.

Theorem 15 (Sylow Theorems, Theorem 12.1, [6])**.** *Let G be a group of order $p^n m$, where p is prime and $gcd(p, m) = 1$. Let n_p be the number of Sylow p-subgroups of G. Then the following hold:*

1. *n_p divides m, which is the index of the Sylow p-subgroup in G.*
2. *$n_p \equiv 1$ (mod p).*
3. *$n_p = |G : N_G(P)|$, where P is any Sylow p-subgroup of G and N_G denotes the normalizer.*
4. *If P is a Sylow p-subgroup of G and Q is any Sylow p-subgroup of G then there exists $g \in G$ such that $Q \leq gPg^{-1}$, i.e., Q is contained in some conjugate of P In particular, any two Sylow p-subgroup of G are conjugate in G.*

Lemma 4. *If S_1, S_2, S_3 are subgroups in G that satisfy the triple-product property, then $|S_i \cap S_j| = 1$ for distinct $i, j \in \{1, 2, 3\}$.*

Proof. Let S_1, S_2, S_3 be subgroups in G that satisfy the *triple-product property*. Assume $|S_1 \cap S_2| > 1$, $x \in S_1 \cap S_2$ and x is not identity, then $x^{-1} \in S_1 \cap S_2$ since both S_1, S_2 are groups. For $1 \in S_3$, $xx^{-1}1 = 1$ where x and x^{-1} are not identity. This contradicts to *Theorem 6*. Then we can conclude $|S_1 \cap S_2| = 1$. Similarly, $|S_i \cap S_j| = 1$ for distinct $i, j \in \{1, 2, 3\}$. $\square$

By *Lemma* 4, it is safe to say that if a group can realize $\langle p, p, p \rangle$ with subgroups, it needs at least 3 different subgroups of order p.

Theorem 16. *Let S_1, S_2, S_3 be three subgroups of a group G. If S_1, S_2, S_3 satisfy the triple-product property and G realize $\langle p, p, p \rangle$ for prime $p \geq 3$ with subgroups S_1, S_2, S_3, then $|G| \geq p^2$*

Proof. If p^2 divides $|G|$, then the theorem holds trivially So suppose that $|G| = pm$ where $gcd(p, m) = 1$

Let $p \geq 3$ be a prime and S_1, S_2, S_3 be subgroups of order p in group G. Assume G realize $\langle p, p, p \rangle$ with S_1, S_2, S_3. Then G has at least three subgroups S_1, S_2, S_3 of order p. According to part 2 in *Theorem* 15, $n_p \geq p + 1$.

We claim that intersection of any two *Sylow p-subgroups* of G are trivial. Assume P and Q are two *Sylow p-subgroups* of G, and $P \cap Q$ are nontrivial. Then for every $x \neq$ identity, if $x \in P \cap Q$, $x^{-1} \in P \cap Q$ which implies that $P \cap Q$ is a subgroup of P and Q. However, order of P and Q are prime p suggesting that only subgroup of P and Q are trivial group and themselves.

Since the intersection of any two *Sylow p-subgroups* of G are trivial, there are $(p - 1)(p + 1) = p^2 - 1$ elements of order p. Then $|G| \geq p^2$. $\square$

In the following paragraphs, we will look into special linear groups and projective special linear groups. We will prove that both $SL_2(\mathbb{F}_p)$ and $PSL_2(\mathbb{F}_p)$ can realize $\langle p, p, p \rangle$.

Theorem 17. *Let G be the group $SL_2(\mathbb{F}_p)$ of 2×2 matrices with entries in $\mathbb{F}_p$ and determinant is 1. Then $|G| = p^3 - p$ and G realize $\langle p, p, p \rangle$ through subgroups S_1, S_2, S_3 of order p.*

Proof. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$, then $a, b, c, d \in \mathbb{F}_p$ and $ad - bc = 1$. In order to find the order of $|G|$, we only need to calculate the number of possible combinations of a, b, c, d.

1. Assume $a, b, c, d \neq 0$, $ad = m$ and $bc = m - 1$ for $m \in \mathbb{F}_p$ and $m \neq 0, 1$. Thus there are $p - 2$ possible m and for each m there are $(p-1)(p-1)$ possible combinations of a, b, c, d. Then the total number of combinations in this case is $(p-2)(p-1)(p-1)$

2. Assume $ad = 0$ and $bc = p - 1$, there are $2p - 1$ combination of a, d and $p - 1$ combinations of b, c. Then the total number of combinations in this case is $(2p-2)(p-1)$

3. Assume $ad = 1$ and $bc = 0$, just similar to case 2, the total number of combinations in this case is $(2p-2)(p-1)$.

Sum the result of 1, 2 and 3, we conclude that $|G| = (p-2)(p-1)(p-1) + 2(2p-2)(p-1) = p^3 - p$.

Let

$$S_1 = \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} |x \in \mathbb{F}_p \right\} S_2 = \left\{ \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} |y \in \mathbb{F}_p \right\} S_3 = \left\{ \begin{pmatrix} 1+z & z \\ -z & 1-z \end{pmatrix} |z \in \mathbb{F}_p \right\}$$

It is trivial to say that S_1 and S_2 are subgroups of order p in G. Let B, C be any matrices in S_3, then $B = \begin{pmatrix} 1+z_1 & z_1 \\ -z_1 & 1-z_1 \end{pmatrix}$, $C = \begin{pmatrix} 1+z_2 & z_2 \\ -z_2 & 1-z_2 \end{pmatrix}$ where $z_1, z_2 \in \mathbb{F}_p$. Assume $B^{-1} = \begin{pmatrix} b_1 & b_2 \\ b_4 & b_3 \end{pmatrix}$ where $b_1, b_2, b_3, b_4 \in \mathbb{F}_p$ Multiply B, C and calculate B^{-1}:

$$BC = \begin{pmatrix} (1+z_1)(1+z_2) - z_1 z_2 & z_2(1+z_1) + z_1(1-z_2) \\ -z_1(1+z_2) - z_2(1-z_1) & z_1 z_2 + (1-z_1)(1-z_2) \end{pmatrix} = \begin{pmatrix} 1+z_1+z_2 & z_1+z_2 \\ -(z_1+z_2) & 1-(z_1+z_2) \end{pmatrix}$$

$$B^{-1} = \begin{pmatrix} -z_1+1 & -z_1 \\ z_1 & 1+z_1 \end{pmatrix}$$

Therefore $BC, B^{-1} \in S_3$. And since $B = I$ when $z_1 = 0$, S_3 is also a subgroup of G.

By *Theorem 6*, we need to check that for any $s_i \in S_i$, $s_1 s_2 = s_3$ if and only if s_i are all identities.

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} = \begin{pmatrix} 1+xy & x \\ y & 1 \end{pmatrix}$$

Then we have $\begin{pmatrix} 1+xy & x \\ y & 1 \end{pmatrix} = \begin{pmatrix} 1+z & z \\ -z & 1-z \end{pmatrix}$ if and only if $x = y = z = 0$ which implies that s_i are all identity. $\square$

The order of $SL_2(\mathbb{F}_p)$ group is $p^3 - p \leq p^3$ which is good. However, we can still improve it by using the group $PSL_2(\mathbb{F}_p)$.

Corollary 4. *Let G be a $PSL_2(\mathbb{F}_p)$ group of order $\frac{1}{2}(p^3 - p)$, then G realize $\langle p, p, p \rangle$ with subgroups of order p.*

Proof. Let SL be a $SL_2(\mathbb{F}_p)$ group, $Z = \{-I_2, I_2\}$ and $G = SL/Z$. Then G is a $PSL_2(\mathbb{F}_p)$ group. Assume G realize $\langle p, p, p \rangle$ for prime $p \geq 3$ with subgroups S_1, S_2, S_3, then $S_i \cap Z = \{I\}$. Let ρ: $SL \to G$ be an homomorphism. Then $\rho(S_i) \sim S_i/\{I\} \sim S_i$, for $i \in \{1, 2, 3\}$. By the *First isomorphism theorem*, $S_i \sim S_i/\{I\}$ for $i = 1, 2, 3$ where$S_1/\{I\}, S_2/\{I\}, S_3/\{I\}$ are subgroups of order p in G. Then G realize $\langle p, p, p \rangle$ with subgroups $S_1/\{I\}, S_2/\{I\}, S_3/\{I\}$. $\square$

Theorem 18. *Suppose that G is the finite group of smallest order that realize $\langle p, p, p \rangle$ for prime p through subgroups of G. Then $p^2 \leq |G| \leq \frac{1}{2}(p^3 - p)$*

Proof. Let G be the smallest group that realize $\langle p, p, p \rangle$ for prime p through subgroups. By *Theorem* 17, $|G| \geq p^2$. And by *Corollary* 4, $|G| \leq |PSL_2(\mathbb{F}_p)| = \frac{1}{2}(p^3 - p)$ $\square$

Now, we have a lower bound and an upper bound for the smallest group to realize $\langle p, p, p \rangle$. However the lower bound $|G| \geq p^2$ is not that good, since all the groups we discussed have order about $\mathcal{O}(p^3)$ and no way near $\mathcal{O}(p^2)$.

Chapter 8

Matrix multiplication with Frobenius groups

Frobenius groups are an important class of finite groups with a well developed theory. With the properties of Frobenius groups, we developed a special method to check whether it can realize $\langle p, p, p\rangle$ and found the smallest Frobenius groups in form of $C_q \rtimes C_p$ realizing $\langle p, p, p\rangle$ in a efficient way.

Definition 13. A finite group G is a Frobenius group if G is a transitive permutation group on a finite set, such that no non-trivial element fixes more than one point; and some element fixes exactly one point.

Definition 14. Let G be the set of ordered pairs (h, k) with $h \in H$ and $k \in K$, and let Φ be a homomorphism from K into $Aut(H)$. Then define the following mutiplication on G:

$$(h_1, k_1)(h_2, k_2) = (h_1 h_2^{\Phi(k_1)}, k_1 k_2)$$

where $\Phi(k_1)$ denote the (left) action of K on H determined by Φ. Then G is called the *semidirect product* of H and K denoted by $G = H \rtimes K$.

The proof of $G = H \rtimes K$ is a group can be find in [6] page 176.

An alternate definition can be found in [11] page 286 which implies that $G = C_q \rtimes C_p$, where q, p are primes and $q = kp + 1$, is a type of Frobenius group. We check groups with small order and find out that $C_3 \rtimes C_2$ (also known as symmetric group of three points) is the smallest group to realize $\langle 2, 2, 2\rangle$ and $C_7 \rtimes C_3$ is the smallest group to realize $\langle 3, 3, 3\rangle$. Therefore, we want to look into this type of groups and try to find the smallest $C_q \rtimes C_p$ to realize $\langle p, p, p\rangle$

According to *Dirichlet's prime number theorem*, there are infinite primes in form $kp+1$ since $gcd(1,p)=1$. Combined with the Chebotarev Density Theorem, we can approximate the number of primes in form $kp+1$ in a interval.

Theorem 19 (Dirichlet's Theorem on Primes in Density version,[16])**.** *Let $a, n \geq 1$ be positive integers with $(a,n)=1$. Then the natural (resp. Dirichlet)density of primes p such that $p \equiv a(mod\ n)$ in the set of all primes of $\mathbb{Z}$ is $\frac{1}{\phi(n)}$.*

Theorem 20 (Prime Number Theorem,[13])**.** *Let $\pi(x)$ be the prime-counting function that gives the number of primes less or equal than x. for any real number x. Then $\frac{x}{\log x}$ is a good approximation to $\pi(x)$:*

$$\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\log x}} = 1$$

By *Theorem* 19 and *Theorem* 20, there are about $(\frac{p^2}{\log\ p^2} - \frac{p}{\log\ p})\frac{1}{\phi(p)} = \frac{p(p-2)}{(2\log p)(p-1)} \approx \frac{p}{2\log p}$ primes between p and p^2 ensuring that there are many groups of the form $C_q \rtimes C_p$ in the interval $p^2 \leq |G| \leq p^3$.

The order of $C_q \rtimes C_p$ is pq, thus we need to express smallest q in terms of p or at least find a bound of q in terms of p.

Theorem 21. *Let G be a Frobenius Group of $C_q \rtimes C_p$ where $p > 5$ and q are primes, G always have subsets $S_1, S_2, S_3 \in G$ such that $|S_i| = p$ and S_1, S_2, S_3 satisfy triple-product property if $q > p^2 - 2p + 3$.*

Proof. Consider G as a permutation group and each element of G is a permutation of p points. Let $S_1 = G_0, S_2 = G_1$ and $S_3 = G_t$, then S_1, S_2, S_3 are all Sylow p-subgroups of order p and we can check *triple-product property* by *Theorem* 7. Also we have $S_1 \cap S_2 = \{1\}$.

Let $T = \{x_1x_2 | x_1 \in S_1, x_2 \in S_2\}$, then $|T| \leq p^2$. Every non-trivial element of S_1 fixes point 0 and every non-trivial element of S_2 fixes point 1. Since $|T \backslash \{1\}| \leq p^2 - 1$, $|S_1 \backslash \{1\}| = p-1$, $|S_2 \backslash \{1\}| = p-1$ and $S_1, S_2 \subseteq T$, then non-trivial elements in T fix no more then $p^2 - 1 - 2(p-2) = p^2 - 2p + 3$ points.

Let S_3 be the stabilizer of a point that is not fixed by any nontrivial element of T, then $S_1 \cap S_3 = S_2 \cap S_3 = T \cap S_3 = \{1\}$. □

In the following paragraphs, we will find a lower bound for q such that $C_q \rtimes C_p$ can realize $\langle p, p, p \rangle$.

Lemma 5. *Let $G = C_q \rtimes C_p$, where p, q are primes, be a Frobenius group, and let $G_{poly} = \{a^{ik}x + t | a^{ik} \in C_p, t \in C_q\}$ be a group under composition of function but NOT multiplication of polynomial. Then G and G_{poly} are isomorphic.*

Proof. Let $\mathbb{F}_q$ be a finite field, and let be the additive group of $\mathbb{F}_q$. The multiplicative group of $\mathbb{F}_q$ is cyclic of order $q - 1 = kp$. Let C_p be the unique subgroup of $\mathbb{F}_q^*$ of order p. With respect to a generator a of $\mathbb{F}_q^*$, we have $C_p = \langle a^k \rangle$.Then we can write G as $G = \{(a^{ik}, t) | a^{ik} \in C_p, t \in C_q\}$.

Let ϕ: $G \to G_{poly}$ be a map such that for $g = (a^{ik}, t) \in G$:

$$\phi(g) = a^{ik}x + t.$$

It is trivial that ϕ is bijective. For any $g_1, g_2 \in G$, let $g_1 = (a^{ik}, t_1)$, $g_2 = (a^{jk}, t_2)$, thus $g_1 g_2 = (a^{(i+j)k} + t_1 a^{jk}, t_1 + t_2)$. $\phi(g_1) = a^{ik}x + t_1$ and $\phi(g_2) = a^{jk}x + t_2$, thus $\phi(g_1) \circ \phi(g_2) = a^{jk}(a^{ik}x + t_1) + t_2 = a^{(i+j)k}x + t_1 a^{jk} + t_1 + t_2$. Then we can conclude that $\phi(g_1 g_2) = \phi(g_1) \circ \phi(g_2)$, G is isomorphic to G_{poly}. $\square$

By *Lemma* 5 we can represent $C_q \rtimes C_p$ as a group of polynomials. And the *Lemma* below shows a polynomial-wise *triple-product property*.

Lemma 6. *Let p and $q = kp + 1$ be primes, and let $G = C_q \rtimes C_p$. For $t \in \mathbb{F}_q$, the subgroups G_0, G_1, G_t realize have the triple-product property if and only if*

$$x^k t - y^k + (1 - t) = 0 \tag{8.1}$$

has a unique solution.

Proof. By *Lemma* 5 G is isomorphic to $G_{poly} = \{a^{ik}x + t | a^{ik} \in C_p, t \in C_q\}$. Define $G_t = \{a^{ik}(x - t) + t | a^{ik} \in C_p, t \in C_q\}$. Then G_t is the subgroup of G_{poly} which fixes the point t. It has order p by the *Orbit Stabilizer Theorem*.

Let $S_1 \cong G_0, S_2 \cong G_1$ and $S_3 \cong G_t$ where S_i are *Sylow p-subgroups* of G. Define $T = \{x_1 x_2 | x_1 \in G_0, x_2 \in G_1\}$ then

$$T = \{a^{(i+j)k}x - a^{jk} + 1 | a^{ik}, a^{jk} \in C_p\}$$

.

Since S_i are *Sylow p-subgroups* of G, $|S_i \cap S_j| = 1$ for distinct $i, j \in \{1, 2, 3\}$. So do G_0, G_1 and G_t. By *Theorem* 7, if $T \cap G_t = \{1\}$ then S_1, S_2, S_3 satisfy *triple-product property*.

Let $p(x) \in T \cap G_t$, then $p(t) = a^{(i+j)k}t - a^{jk} + 1 = t$. Regard $a^{(i+j)k}$ as x^k and a^{jk} as y^k, then $p(t) = t$ is just same as equation 8.1. Thus, the order of $T \cap G_t$ is equal to the number of solution of equation 8.1. We conclude that if equation 8.1 has unique solution then G realize $\langle p, p, p \rangle$. $\square$

The *Hasse-Weil Bound* is a famous result from number theory which bound the number of solutions to polynomial over a finite field.

Theorem 22 (Hasse Weil Bound [8])**.** *If the number of points on the curve C of genus g over the finite field $\mathbb{F}_q$ of order q is N,then*

$$|N-(q+1)| \leq 2g\sqrt{q}$$

In this case, we can regard $a^{(i+j)k}t - a^{jk} + 1$ as a curve, then the number of solutions of the equation 8.1 is just the number of points on the curve over $\mathbb{F}_q$. And genus

$$\begin{aligned} g &= \frac{(k-1)(k-2)}{2} - d \\ &= \frac{(\frac{q-1}{p}-1)(\frac{q-1}{p}-2)}{2} - d \\ &< \frac{(\frac{q}{p})(\frac{q}{p})}{2} = \frac{1}{2}\frac{q^2}{p^2}. \end{aligned}$$

Theorem 23. *Let G be a Frobenius Group of $C_q \rtimes C_p$, then G can realize $\langle p,p,p\rangle$ only if $q \geq p^{\frac{4}{3}}$.*

Proof. Let N be the number of solution of the equation 8.1. By *Theorem* 22, N should satisfy the following inequality:

$$|N-(q+1)| \leq \frac{q^2}{p^2}\sqrt{q}.$$

Let $N=1$, we have $q \geq p^{\frac{4}{3}}$. Then we can conclude that if $q < p^{\frac{4}{3}}$, then $N > 1$ for all t. In this case, we can not find S_1, S_2, S_3 satisfy *triple-product property.* □

By *Theorem* 21, $C_q \rtimes C_p$ realize $\langle p,p,p\rangle$ if $q > p^2 - 2p + 3$, while by *Theorem* 23, $C_q \rtimes C_p$ can not realize $\langle p,p,p\rangle$ if $q < p^{\frac{4}{3}}$. The proof of *Theorem* 23 also gives us an efficient method to check *triple-product property.* Then we developed the following algorithm to check *triple-product property* for given $C_q \rtimes C_p$:
Let $\mathbb{F}_q$ be finite feild of order q and $(\mathbb{F}_q,\cdot) = \langle a\rangle$.Let $X = \{a^{ik} | i \in [1,2,...,p-1]\}$ where $q = kp+1$.

for $t \in [2,3,...,q-1]$ **do**
  $Y = \{tx | x \in X\}$
  $Z = \{t+x-1 | x \in X\}$

$n = Y \cap Z$
if $n = 1$ **then**
$C_q \rtimes C_p$ realize $\langle p, p, p \rangle$.
end if
end for

In this algorithm, $Y \cap Z = 1$ lead to $a^{ik}t - a^{jk} + 1 = t$ which suggest that $|T \cap S_3| = 1$. Thus we conclude that $C_q \rtimes C_p$ realize $\langle p, p, p \rangle$ if $n = 1$.

Following table show some result including the smallest q for some prime p and $\alpha(G)$ for $G = C_q \rtimes C_p$.

p	q	$\alpha(G)$
101	3637	2.78
103	2267	2.67
107	2141	2.64
109	2399	2.66
113	2713	2.67
127	3049	2.66
131	3407	2.67
137	4933	2.73
139	5839	2.76
149	7451	2.78
151	4229	2.66
157	3769	2.63
163	5869	2.70
167	5011	2.66
173	6229	2.70
179	4297	2.61
181	5431	2.65
191	6113	2.66
193	6563	2.67
197	7487	2.69
199	11941	2.77
211	8863	2.70

pseudo-exponent of Frobenius group

According the proof of *Theorem* 21, if p^2 nontrivial elements in T fix no more then $q - 1$ points, then $C_q \rtimes C_p$ realize $\langle p, p, p \rangle$. The well-known *Coupon collector's problem* can be a good analogy of this problem.

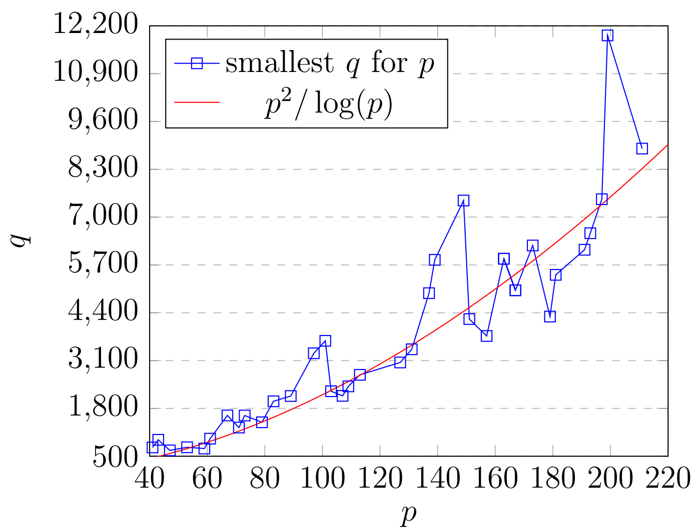
Theorem 24 (Coupon collector's problem, chapter 8.4 in [10])**.** *Suppose that you throw balls into n distinguishable bins. After throwing $\mathcal{O}(n \log n)$ balls, every bin is non-empty with high probability.*

We can regard q points as different bins and p^2 as number of balls thrown. By *Theorem* 24, if

$$p^2 \leq q \log q \tag{8.2}$$

there is a high chance that at least one bin is empty, which also implies that at least one points can not be fixed by nontrivial elements in T. Rewrite 8.2, $q \geq \frac{p^2}{\log q} \geq \frac{p^2}{\log p}$. Then we can conclude that the smallest q such that $C_q \rtimes C_p$ realize $\langle p, p, p \rangle$ is about $\mathcal{O}(\frac{p^2}{\log p})$. The following graph also implies the same result.

smallest q for $41 \leq p \leq 211$ such that $C_q \rtimes C_p$ realize $\langle p, p, p \rangle$



In this graph, p is x-axis and q is y-axis. The blue broken line links point (p, q), where $41 \leq p \leq 211$ and q is the smallest prime such that $C_q \rtimes C_p$ realize $\langle p, p, p \rangle$. The red curve is the graph of $q = \frac{p^2}{\log p}$.

Chapter 9

Future work

Question: Smallest groups realizing $\langle n, n, n \rangle$ through subgroups

In Chapter 6, we have found bound for the smallest group realizing $\langle p, p, p \rangle$. If we want to implement this methods to a practical algorithm, the next step is finding the smallest groups realizing $\langle n, n, n \rangle$. We believed that the smallest groups realizing $\langle p, p, p \rangle$ with subgroups are:

- $SL_2(\mathbb{F}_p)$, when $p = 5$.
- $PSL_2(\mathbb{F}_p)$, when $p = 7, 11, 19, 23, 43$.
- $C_q \rtimes C_p$, when $p \geq 13$ and $p \neq 19, 23, 43$.

Finding the smallest group realizing $\langle p, p, p \rangle$ and $\langle q, q, q \rangle$ gives a bound or even the smallest group realizing $\langle pq, pq, pq \rangle$ via the following result.

Lemma 7. *[4] If N is a normal subgroup of G that realizes $\langle n_1, n_2, n_3 \rangle$ and G/N realizes $\langle m_1, m_2, m_3 \rangle$, then G realizes $\langle n_1m_1, n_2m_2, n_3m_3 \rangle$.*

Lemma 7 provides a great property which can be used to prove the *triple-product property* for large groups with complicated structure. Combining it with *Sylow Theorems*, we may prove the *triple-product property* for groups such that $|G| = p^k m$.

Question: Improve the bound for $C_q \rtimes C_p$

In Chapter 7, we represent group elements of $C_q \rtimes C_p$ by polynomials and check the *triple-product property* by counting the number of solutions. Besides the *Hasse-Weil Theorem*, *Laszlo Babai* also gave us the following *Theorem*.

Theorem 25 ([1],page 19)**.** *Let $k|q-1$ be an integer, A_1 $A_2 \subseteq F_q$, and let N denote the number of solutions of the equation*

$$x + y = z^k \quad (x \in A_1, y \in A_2, z \in \mathbb{F}_q^{\times}).$$

Then

$$|N - \frac{|A_1|A_2|(q-1)|}{q}| < k\sqrt{|A_1||A_2|q}$$

He combined Fourier transform and characters theory to discuss the number of solution of equations over finite abelian group. Although, in our case, his result is not as suitable as the *Hasse-Weil Theorem*, if we could modify this conclusion to suit our case, it might yields better bounds for q.

Question: Smallest groups realizing $\langle n, n, n \rangle$ through subsets

In Chapter 5, we show the embedding to cyclic groups which are normal subgroups of some larger groups. Based on *Lemma* 7, we can discuss the *triple-product property* of $C_m \times K$ or $C_m \rtimes K$. However, since the *pseudo-exponent* of cyclic groups are 3, there is a great chance that groups which have cyclic normal subgroups are not the smallest groups realizing $\langle n, n, n \rangle$. Also we develop a very naive algorithm to check whether G realize $\langle n, n, n \rangle$ through subsets. If we can avoid traverse subsets of G, there will be a efficient algorithm to find smallest groups realizing $\langle n, n, n \rangle$.

Question: Better upper bounds for ω

In our research, we did not focus on the upper bound on the complexity exponent ω which is the hottest topic of fast matrix multiplication among researchers. In Chapter 3, we show how this group-theoretic embedding converts matrix multiplication into $\mathbb{C}G - modules$ multiplication. And also the relations between *pseudo-exponent* and ω is based on the representation theory which gives the decomposition of $\mathbb{C}G - modules$. If there are more efficient algorithms to multiply $\mathbb{C}G - modules$, we will find a better method to bound ω which might lead to a better upper bound.

Bibliography

[1] L. Babai. The fourier transform and equations over finite abelian groups. lecture note, Department of Computer Science, University of Chicago, 2002.

[2] M. Bläser. *Fast Matrix Multiplication.* Number 5 in Graduate Surveys. Theory of Computing Library, 2013.

[3] H. Cohn, R. Kleinberg, B. Szegedy, and C. Umans. Group-theoretic algorithms for matrix multiplication. *Proceedings of the 46th Annual Symposium on Foundations of Computer Science, 23-25 October 2005, Pittsburgh, PA, IEEE Computer Society, pp. 379-388*, 2005.

[4] H. Cohn and C. Umans. A group-theoretic approach to fast matrix multiplication. *Proceedings of the 44th Annual Symposium on Foundations of Computer Science, 11-14 October 2003, Cambridge, MA, IEEE Computer Society, pp. 438-449*, 2003.

[5] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *J. Symbolic Comput.*, 9(3):251–280, 1990.

[6] D. S. Dummit and R. M. Foote. *Abstract algebra.* John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2004.

[7] J. B. Fraleigh. *A first course in abstract algebra.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1967.

[8] H. Hasse. Zur Theorie der abstrakten elliptischen Funktionenkörper. I. Die Struktur der Gruppe der Divisorenklassen endlicher Ordnung. *J. Reine Angew. Math.*, 175:55–62, 1936.

[9] M. Huhtanen and A. Perämäki. Factoring matrices into the product of circulant and diagonal matrices. *J. Fourier Anal. Appl.*, 21(5):1018–1033, 2015.

[10] R. Isaac. *The pleasures of probability.* Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1995. Readings in Mathematics.

[11] G. James and M. Liebeck. *Representations and characters of groups.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1993.

[12] M. Lang. Group theoretical methods of matrix multiplication and new upper bounds on the triple product capacity of dihedral and generalized dicyclic groups. http://theory.stanford.edu/~virgi/matrixmult-f.pdf, 2014.

[13] D. J. Newman. Simple analytic proof of the prime number theorem. *Amer. Math. Monthly*, 87(9):693–696, 1980.

[14] V. Strassen. Gaussian elimination is not optimal. *Numer. Math.*, 13:354–356, 1969.

[15] V. Strassen. Relative bilinear complexity and matrix multiplication. *J. Reine Angew. Math.*, 375/376:406–443, 1987.

[16] N. G. Triantafillou. The chebotarev density theorem. *Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts*, 2015.

[17] V. V. Williams. Multiplying matrices in $\mathcal{O}(n^{2.373})$ time. Stanford University, http://theory.stanford.edu/~virgi/matrixmult-f.pdf, 2014.