Proof of Kemnitz' Conjecture and a generalization to higher dimensions

Panwei Hu

Bachelor Thesis

RWTH Aachen University
Univ.-Prof. Dr. Eberhard Triesch
Lehrstuhl II für Mathematik

Contents

1 Abstract

In this thesis, I worked with a lattice point problem which deals with the smallest number $f(n, d)$ of a sequence of lattice points to guarantee that the centroid of n lattice points is also a lattice point in a d-dimensional vector space. The thesis first introduces this concept and gives some examples in the low-dimensional case. The next chapter deals with the *Kemnitz' Conjecture*. The thesis will illustrate a detailed proof which is based on [1]. The last chapter deals with the *Alon-Dubiner Theorem* which gives a general upper bound for $f(n, d)$.

2 Introduction

We define a d-dimensional affine space V and consider the set of points which lie in the set

$$V_d := \Big\{ \sum_{i=1}^{d} a_i v_i \;\Big|\; a_i \in \mathbb{Z}, \quad 1 \le i \le d \Big\},$$

where $\big\{ v_i \;\big|\; 1 \le i \le d \big\}$ are linear independent vectors in V. We also call the point in V_d as *lattice* points. We now want to find out the minimum of the number f s.t. given f sequences in V_d, we can guarantee to find out a subsequence of length n, s.t. the centroid of this subsequence is also a lattice point. We define such minimum number as $f(n, d)$.

Consider the additive group $G := \mathbb{Z}_n^d$. We call a subsequence of length l, which sums to a 0 in $\mathbb{Z}_n^d$ as 0-sum l-subsequence (w.r.t G), where 0 denotes the zero vector in G. It is not hard to see that the above problem can be equivalently formulated as finding the number $f(n, d)$, s.t. for any sequences of elements in G, with length $l \ge f(n, d)$, there exists a 0-sum n-subsequence.

Since given a sequence of elements in G, it is possible that the sequence contains multiple identical elements. We will model a sequence thus as a multiset. From set theory, we know that a multiset whose elements are from the set A can be treated as a function $f : A \to K$, where K is the cardinal number. We thus extend the operation on sets to the multisets as following:

Definition 2.1. Let $f : A \to K, g : B \to K$ be two multisets. Let the domain of a function f be denoted as $\mathrm{Dom}(f)$.
We define the *union*,*intersection*,*difference* as following, using the same notation in set theory:

- $f \cup g$: A multiset whose domain is defined as $\mathrm{Dom}(f \cup g) = \mathrm{Dom}(f) \cup \mathrm{Dom}(g) = A \cup B$ s.t.

$$(f \cup g)(x) = \max\{f(x), g(x)\}$$

- $f \cap g$: A multiset whose domain is defined as $\mathrm{Dom}(f \cap g) = \mathrm{Dom}(f) \cap \mathrm{Dom}(g) = A \cap B$ s.t.

$$(f \cap g)(x) = \min\{f(x), g(x)\}$$

- $f \setminus g$: Let $C := \left\{ x \in A \cap B \;\middle|\; f(x) > g(x) \right\}$ $f \setminus g$ is a multiset whose domain is defined as $\mathrm{Dom}(f \cap g) = (A \setminus B) \cup C$ s.t.

$$(f \setminus g)(x) := f(x) - g(x)$$

We see that the usual operation extends naturally to the multiset, thus for the following discussions I abuse the name 'set' to denote both set and multiset, when the context is clear if multiplicity of an element is allowed. We begin further with our discussion on $f(n, d)$. A natural bound on the number $f(n, d)$ is given in the following Lemma.

Lemma 2.2.

$$(n-1)2^d + 1 \leq f(n, d) \leq (n-1)n^d + 1 \tag{2.2.1}$$

Proof. The left inequality can be seen as following:

We construct $(n-1)2^d$ vectors, which include all the vectors in $\mathbb{Z}_n^d$, which has 0 or 1 in their entry, so there are in all 2^d different vectors. Each vector appear exactly $n-1$ times. It is impossible to find a 0-sum n-subsequence among these vectors. This claim can be proved by induction on the dimension d.

Induction base: $d = 1$. We have 1 and 0 each appear exactly $n-1$ times. We see that to make a $0-sum\mathrm{n}-subsequence$, we need to have either n 0s or n 1s, which is impossible.

Induction step. Assume that the claim is valid for all $k \leq d$. For the dimension $d+1$. We consider the first d components (for a vector $v \in \mathbb{Z}_n^{d+1}$, I use the notation $\tilde{v}$ to denote the truncated). Now due to increased dimension by one. The number of occurence of the vector in the first d components are $2(n-1)$. Given n summands, if each $\tilde{v}$ appear less than n times, then we can reduce to the $d-1$ case and affirm that the sum is not 0 in the d components and thus also not zero in $\mathbb{Z}_n^{d+1}$. If there is $\tilde{v}$ appear more than $n-1$ times, so it could only be possible that all the n summands have the same first d components. But now it falls back to the induction basis, because we do not need to consider the first d component and only focus on the $d+1$ component. We have at most $n-1$ 1s and 0s which is impossible to have a zero sum.

Thus we must have

$$(n-1)2^d + 1 \leq f(n, d).$$

To prove the right inequality we use the pigeon hole principle. Since $\left|G\right| = n^d$, given $(n-1)n^d + 1$ elements, there are at least one vector v which has multiplicity

$$\lceil \frac{(n-1)n^d + 1}{n^d} \rceil = n.$$

Choosing these n vectors v, we obtain a 0-sum n-subsequence. We thus deduce that

$$f(n,d) \leq (n-1)n^d + 1.$$

□

It is generally very difficult to examine $f(n,d)$ for every possible $n \in \mathbb{N}$, thus we would like to restrict n to a smaller set of numbers, for example the prime numbers. The following Lemma helps us with such reduction.

Lemma 2.3.

$$f(pq,d) \leq f(p,d) + p(f(q,d) - 1) \tag{2.3.1}$$

Proof. For the convenience of notation, we define

$$f_1 := f(p,d), f_2 := f(q,d), f := f_1 + p(f_2 - 1).$$

Given a sequence S of f elements, since $f > f_1$, we can find a 0-sum p-subsequence, which we denote as A_1. We then consider $S \setminus A_1$, since the rest is still larger than f_1 we can further find a 0-sum p-subsequence, which we denote as A_2. We perform this procedure recursively. Finally, since

$$f = f_1 + p(f_2 - 1),$$

we would obtain f_2 0-sum p-subsequence (w.r.t $\mathbb{Z}_p^d$). According to the definition, each such subsequence A_l sums to a vector w_l of the form

$$w_l = p \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix}, \quad v_i \in \mathbb{Z},\ 1 \leq i \leq d$$

Among all these f_2 vectors, there exists a 0-sum q-subsequence (w.r.t $\mathbb{Z}_q^d$), which means they sum to a vector z, with

$$z = q \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix}, \quad v_i \in \mathbb{Z},\ 1 \leq i \leq d$$

Since z are the sum of vectors of the form w_l, we see that z can also be written as

$$z = pq \cdot \begin{pmatrix} \tilde{v}_1 \\ \vdots \\ \tilde{v}_d \end{pmatrix}, \quad \tilde{v}_i \in \mathbb{Z},\ 1 \leq i \leq d$$

Combining the results we obtain $pq = n$ sequences which sums to a vector z, whose every entry is a multiple of n, thus we obtain a 0-sum n-subsequence. □

Remark 2.4. Due to the symmetry we would obtain similarly:

$$f(pq, d) \leq f(q, d) + q(f(p, d) - 1) \tag{2.3.1'}$$

Combining (2.3.1) and (2.3.1′), we obtain the following upper bound.

Corollary 2.5.

$$f(pq, d) \leq \min\{f(p, d) + p(f(q, d) - 1), f(q, d) + q(f(p, d) - 1)\} \tag{2.5.1}$$

We can use (2.5.1) to find out some values of $f(n, d)$. Before we state the famous theorem first discovered by Erdös ,Ginsburg and Ziv, we state the *Cauchy-Davenport* Theorem which would help us with the proof.

Theorem 2.6 (Cauchy-Davenport)**.** Let p be a prime number. If $A, B \subset \mathbb{Z}_p$ are nonempty, then

$$\big|A + B\big| \geq \min\{p, \big|A\big| + \big|B\big| - 1\},$$

where $A + B$ denotes the set of the addition of elements in A and B, see Definition 4.1.

Proof. See for example [2]. □

Theorem 2.7 (Erdös-Ginsburg-Ziv)**.**

$$f(n, 1) = 2n - 1$$

Proof. From (2.2.1), we obtain

$$f(n, 1) \geq 2n - 1 \tag{2.7.1}$$

So we only need to prove that

$$f(n,1) \leq 2n-1. \tag{2.7.2}$$

From (2.5.1), we could restrict the problem to the case when n is prime, because if we have proved the theorem for all the prime numbers, then for two given prime numbers p, q:

$$\begin{aligned} f(pq,1) &\leq f(p,1) + p(f(q,1)-1) \\ &\leq 2p-1+p(2q-2) \\ &= 2pq-1 \end{aligned} \tag{2.7.3}$$

Applying (2.7.3) recursively to the prime factorization of any integer n leads to (2.7.2).

We thus prove (2.7.2) for the case p, p a prime number. Assume there is a sequence with $2p-1$ elements, if there are already p identical elements, we are done. So we could assume that there are no elements which have multiplicity larger than $p-1$. We arrange all the element increasingly (which is possible in the 1 dimensional case), and enumerate the element s.t.

$$a_1 \leq a_2 \leq \cdots \leq a_{2p-1}.$$

We define the set $A_i := \{a_i, a_{i+p-1}\}, 1 \leq i \leq p-1$. Note that

$$\left|A_i\right| = 2, \quad \forall 1 \leq i \leq p-1,$$

because we assume that no elements appear more than $p-1$ times. Apply Theorem 2.6 recursively, we obtain

$$\begin{aligned} \left|A_1 + \cdots + A_{p-1}\right| &\geq \min\{p, \left|A_2 + \cdots + A_{p-1}\right| + \left|A_1\right| - 1\} \\ &\geq \min\{p, \left|A_2 + \cdots + A_{p-1}\right| + 1\} \\ &\geq \min\{p, \left|A_3 + \cdots + A_{p-1}\right| + 2\} \\ &\vdots \\ &\geq \min\{p, \left|A_{p-2}, A_{p-1}\right| + p - 3\} \\ &\geq \min\{p, 2+2-1+p-3\} \\ &\geq \min\{p,p\} = p. \end{aligned}$$

This implies that

$$A_1 + \cdots + A_{p-1} = \mathbb{Z}_p$$

In particular we can pick $p-1$ elements from the first $2p-2$ elements $a_{i_1}, \ldots, a_{i_{p-1}}$, with $a_{i_k} \in A_k, 1 \leq k \leq p-1$. s.t.

$$a_{i_1} + \cdots + a_{i_{p-1}} = -a_{2p-1}.$$

Moving a_{2p-1} to the left handside of the equation, we thus obtain a 0-sum p-subsequence $a_{i_1}, \ldots, a_{i_{p-1}}, a_{2p-1}$, which means

$$f(p,1) \leq 2p-1$$

□

Lemma 2.8.

$$f(2^n, d) = (2^n - 1)2^d + 1$$

Proof. We prove the Lemma through induction on n. For $n = 1$, we know from (2.2.1) that

$$(2-1)2^d + 1 \leq f(2,d) \leq (2-1)2^d + 1$$

Thus $f(2,d) = 2^d + 1$

For the general case n, we have $f(2^{n-1}, d) = (2^{n-1} - 1)2^d + 1$ from induction. Using (2.5.1), we obtain:

$$\begin{aligned} f(2^n, d) = f(2 \cdot 2^{n-1}, d) &\leq f(2^{n-1}, d) + 2^{n-1} \cdot (f(2,d) - 1) \\ &= (2^{n-1} - 1)2^d + 1 + 2^{n-1} \cdot 2^d \\ &= (2^n - 1)2^d + 1 \end{aligned}$$

□

It is proved in [3] that the following holds

Lemma 2.9.

$$f(3,2) = 9$$

Applying the trick as in Lemma 2.8, using the base case in Lemma 2.9, we obtain

Lemma 2.10.

$$f(3^n, 2) = 4 \cdot 3^n - 3$$

By setting $d = 2$ in Lemma 2.8, we see that

$$f(2^n, 2) = 4 \cdot 2^n - 3$$

we see that there is a simlar structure as in $f(3^n, 2)$. A natural question emerges that if this is the case for all $n \in \mathbb{N}$, that is

$$f(n, 2) = 4n - 3$$

This is the well-known *Kemnitz' Conjecture*, which we will discuss in the next chapter.

3 Proof of Kemnitz Conjecture

For the following chapter, I use the notation $\equiv$ to denote the modulo in $\mathbb{Z}_p$. The 0 denotes the usual neutral element of addition in the corresponding abelian group. In particular, 0 denotes the standard 0 in the abelian group $\mathbb{Z}_p$ and $(0,0)$ in the case of $\mathbb{Z}_p \times \mathbb{Z}_p$.

For the proof of the *Kemnitz' Conjecture*, the following theorem is very helpful, which is also known as *Chevalley-Warning Theorem*.

Theorem 3.1 (*Chevalley-Warning Theorem*)**.** Let p be a prime number and $q = p^t, t \in \mathbb{N}$. We use $\mathbb{F}_q$ to denote the finite field of q elements. Let $p_1, \ldots, p_m$ be m polynomials in $\mathbb{F}_q[x_1, x_2, \ldots, x_n]$, with degree d_i. Denote the number of common zeros of the m polynomials as N. If

$$\sum_{i=1}^{m} d_i < n,$$

then

$$N \equiv 0 \pmod{p}$$

Proof. See for example [2]. □

We would like to give a proof of the *Kemnitz' Conjecture*. First we note that from if $f(p,2) = 4p-3$ and $f(q,2) = 4q-3$, then applying (2.5.1)

$$\begin{aligned} f(pq,2) &\leq f(p,2) + p(f(q,2)-1) \\ &\leq 4p - 3 + p(4q-3-1) \\ &= 4pq - 3. \end{aligned}$$

Thus it suffices to consider the case when p is prime. In addition from Lemma 2.8, we already know that

$$f(2,2) = 4 \cdot 2 - 3 = 5,$$

it suffices to consider the odd prime number. Thus we assume in the following paragraph that p is an odd prime number. We denote J, X and other capital alphabets as a multiset of $G := \mathbb{Z}_p \times \mathbb{Z}_p$. We use the notation $(m|J)$ to denote the number of the 0-sum m-subsequence (w.r.t G) in J.

First we obtain the following Corollary from Theorem 3.1.

Corollary 3.2. If $\left|J\right| = 3p-3$, then $1-(p-1|J)-(p|J)+(2p-1|J)+(2p|J) \equiv 0$

Proof. We specify J as the multiset $\left\{ (a_n, b_n) \;\middle|\; 1 \leq n \leq 3p-3 \right\}$.
We consider three polynomials

$$p_1 := \sum_{n=1}^{3p-3} x_n^{p-1} + x_{3p-2}^{p-1}, \quad p_2 := \sum_{n=1}^{3p-3} a_n x_n^{p-1}, \quad p_3 := \sum_{n=1}^{3p-3} b_n x_n^{p-1}$$

Since the total sum of degress are $3(p-1) = 3p-3 < 3p-2$, we can apply the *Chevalley-Warning Theorem*. Since three polynomials have 0 as a common zero, the *Chevalley-Warning Theorem* states that there are non-trivial common zeros. We consider the common zeros depending on the term x_{3p-2} in p_1.

Case 1 : $x_{3n-2} = 0$
In this case for the zero of p_1, it needs to satisfy

$$\sum_{n=1}^{3p-3} x_n^{p-1} \equiv 0$$

We know that

$$x^{p-1} \equiv \begin{cases} 1, & \text{if } x \not\equiv 0; \\ 0, & \text{if } x \equiv 0. \end{cases}$$

Since there are in all $3p-3$ variables left, there could only be three cases:

(i) 0 of them are 1:
This corresponds to the case when all x_i are 0.

(ii) p of them are 1:
This corresponds to p of the x_i are non-zero. We enumerate them as $x_{i_1}, \ldots, x_{i_p}$. In order for this particular assignment to be the common zero, it needs to satisfy

$$\sum_{j=1}^{p} a_{i_j} \equiv 0, \quad \sum_{j=1}^{p} b_{i_j} \equiv 0.$$

The total number the set $\{i_1, \ldots, i_p\}$ is exactly $(p|J)$. Combining the fact that there are $p-1$ possible value for each x_i s.t $x_i^{p-1} \equiv 1$, the total number of common zeros with p non-zero entries is $(p-1)^p(p|J)$.

(iii) $2p$ of them are 1:
Arguing similarly as in the second case, we obtain the total number of common zeros with $2p$ non-zero entries is given by $(p-1)^{2p}(2p|J)$.

Case 2 : $x_{3n-2} \neq 0$
Since $x_{3n-2} \neq 0$, then $x_{3n-2}^{p-1} \equiv 1$. The other terms of p_1 must thus satisfy

$$\sum_{n=1}^{3p-3} x_n^{p-1} \equiv p-1.$$

There are two possible cases for the other entries of x.

(i) $p-1$ of them are 1. This corresponds to the case where there needs to be $p-1$ elements in J which form a zero-sum. There are in all

$$(p-1) \cdot (p-1)^{p-1} \cdot (p-1|J) = (p-1)^p(p-1|J),$$

possible choices for x, where the first term corresponds to the choice of x_{3n-2}, the second to the rest of x_i.

(ii) $2p-1$ of them are 1. Similarly, the number of possible common zeros in this case is given by

$$(p-1)^{2p}(2p-1|J)$$

Collecting all the number of common zeros considered in different cases, we obtain, with the usage of *Chevalley-Warning Theorem*

$$1+(p-1)^p(p|J)+(p-1)^{2p}(2p|J)+(p-1)^p(p-1|J)+(p-1)^{2p}(2p-1|J) \equiv 0$$

Simplifying the equation above we obtain

$$1-(p-1|J)-(p|J)+(2p-1|J)+(2p|J) \equiv 0$$

□

Corollary 3.3. If $|J| = 3p-2$, or $|J| = 3p-1$, then $1-(p|J)+(2p|J) \equiv 0$

Proof. w.l.o.g, $|J| = 3p-2$, the case for $\left|J\right| = 3p-1$ is analogous. Consider the polynomials in $\mathbb{F}_p[x_1, x_2, \ldots, x_{|J|}]$,

$$\sum_{n=1}^{3p-3} x_n^{p-1}, \quad \sum_{n=1}^{3p-3} a_n x_n^{p-1}, \quad \sum_{n=1}^{3p-3} b_n x_n^{p-1}$$

since the total degrees are $3p-3 < 3p-2$, the *Chevalley-Warning Theorem* is applicable. Counting common zeros as in Corollary 3.2, we obtain:

$$1+(p-1)^p(p|J)+(p-1)^{2p}(2p-1|J) \equiv 0 \tag{3.3.1}$$

$$\Rightarrow \quad 1-(p|J)+(2p|J) \equiv 0 \tag{3.3.2}$$

□

Remark 3.4. Note that in the case where $\left|J\right| = 3p-1$ or $3p-2$, the three polynomials we examined all have the same number of monomials, i.e. $\left|J\right|$, different from in the case in $\left|J\right| = 3p-3$. This is due to the fact the total degrees of the three polynomials (denoted as $f_i, i \in \{1,2,3\}$) natuarlly satisfy $\sum_i deg(f_i) < \left|J\right|$

Corollary 3.5. If $\left|J\right| = 3p-2$ or $\left|J\right| = 3p-1$, then $(p|J) = 0$ implies $(2p|J) \equiv -1$.

Proof. This follows directly from Corollary 3.3 by setting $(p|J) = 0$. □

Corollary 3.6. If J contains exactly $3p$ elements, and $\sum_{x\in J} x \equiv 0$, then $(p|J) > 0$.

Proof. Assume that $(p|J) = 0$, this means that there are no subsets in J, of cardinality p which builds a zero-sum sequence. In particular, if we randomly pick an $x \in J$ and consider the multiset $J-x$, it follows that $(p|J-x) = 0$. Now, since $\left|J-x\right| = 3p-1$, we can apply Corollary 3.5 to obtain that $(2p|J-x) \equiv -1$, in particular, this implies that $(2p|J-x) > 0$.
We observe that $\forall A \subset J$, s.t. $\sum_{a\in A} a \equiv 0$,

$$\sum_{a\in A} a + \sum_{b\in J-A} b = \sum_{j\in J} j \equiv 0$$

$$\Rightarrow \sum_{b\in J-A} b \equiv 0$$

We could see that the map T

$$T : \{\, A \subset J \;\big|\; |A| = p, \sum_{a\in A} a \equiv 0 \,\} \to \{\, A \subset J \;\big|\; |A| = 2p, \sum_{a\in A} a \equiv 0 \,\}$$
$$A \mapsto J - A$$

is a bijection. It follows that:

$$(p|J) = (2p|J) \geq (2p|J - x) > 0,$$

which is a contradiction to the assumption $(p|J) = 0$ □

Corollary 3.7. If $\big|X\big| = 4p - 3$, then

1. $-1 + (p|X) - (2p|X) + (3p|X) \equiv 0$
2. $(p-1|X) - (2p-1|X) + (3p-1|X) \equiv 0$

Proof. Consider the polynomials f_1, f_2, f_3

$$\sum_{n=1}^{4p-3} x_n^{p-1}, \quad \sum_{n=1}^{4p-3} a_n x_n^{p-1}, \quad \sum_{n=1}^{4p-3} b_n x_n^{p-1}$$

The polynomials satisfy $deg(f_i) < 4p-3$, so the *Chevalley-Warning Theorem* is applicable. Performing the similar procedure as in Corollary 3.2, we obtain the first equation.

□

Corollary 3.8. If $\big|X\big| = 4p-3$, then $3-2(p-1|X)-2(p|X)+(2p-1|X)+(2p|X) \equiv 0$.

Proof. We deduce from Corollary 3.2 that:

$$\sum_{I} 1 - (p-1|I) - (p|I) + (2p-1|I) + (2p|I) \equiv 0,$$

where the sum is over $I \subset X$, s.t, $\big|I\big| = 3p-3$. For a given subset $Y \subset X$,s.t $\big|Y\big| = p$ and $\sum_{y\in Y} y \equiv 0$, we want to find out the number of pairs (Y, I), s.t

$Y \subset I, \left|I\right| = 3p-3$.
We could see that

$$\left|\{(Y,I)|Y \subset I, s.t.\ \sum_{y\in Y} y \equiv 0,\ \left|I\right| = 3p-3\}\right| = \binom{3p-3}{2p-3},$$

since once we have chosen p elements Y, we need to further choose $\left|I\right| - p = 3p-3-p = 2p-3$ elements from total $\left|X\right| - p = 4p-3-p = 3p-3$ elements. We consider the sum:

$$\sum_{\substack{Y\subset X, |Y|=p,\\ \sum_{y\in Y} y=0}} \sum_{\substack{I, s.t.\\ Y\subset I,\\ \left|I\right|=3p-3}} 1$$

Note that

$$\begin{aligned}\binom{3p-3}{2p-3}(p|X) = \sum_{\substack{Y\subset X, |Y|=p,\\ \sum_{y\in Y} y=0}} \binom{3p-3}{2p-3} &= \sum_{\substack{Y\subset X, |Y|=p,\\ \sum_{y\in Y} y=0}} \sum_{\substack{I, s.t.\\ Y\subset I,\\ |I|=3p-3}} 1 \\ = \sum_{\substack{I, s.t.\\ |I|=3p-3}} \sum_{\substack{Y\subset I, |Y|=p,\\ \sum_{y\in Y} y=0}} 1 &= \sum_{I} (p|I)\end{aligned}$$

Performing similar calculation, we get

$$\begin{aligned}&\binom{4p-3}{3p-3} - \binom{3p-2}{2p-2}(p-1|X) - \binom{3p-3}{2p-3}(p|X) \\ &\quad + \binom{2p-2}{p-2}(2p-1|X) + \binom{2p-3}{p-3}(2p|X) \equiv 0 \end{aligned} \tag{3.8.1}$$

We finally prove that

$$\binom{4p-3}{3p-3} \equiv 3, \binom{3p-2}{2p-2} \equiv 2, \tag{3.8.2}$$

because:

$$\begin{aligned}\binom{4p-3}{3p-3} &\equiv \frac{(4p-3)\cdots(4p-(p-1))\cdot 3p\cdot(3p-1)\cdot(3p-2)}{p!}\\ &\equiv \frac{(4p-3)\cdots(4p-(p-1))\cdot 3\cdot(3p-1)\cdot(3p-2)}{(p-1)!}\\ &\equiv 3\cdot\frac{(-3)\cdot(-4)\cdots(-(p-1))\cdot(-1)\cdot(-2)}{(p-1)!}\\ &\equiv 3\cdot\frac{(p-1)!}{(p-1)!}\\ &\equiv 3\end{aligned}$$

and

$$\begin{aligned}\binom{3p-2}{2p-2} &\equiv \frac{(3p-2)\cdots(3p-(p-1))\cdot 2p\cdot(2p-1)}{p!}\\ &\equiv \frac{(3p-2)\cdots(3p-(p-1))\cdot 2\cdot(2p-1)\cdot(2p-2)}{(p-1)!}\\ &\equiv 2\cdot\frac{(-2)\cdot(-4)\cdots(-(p-1))\cdot(-1)}{(p-1)!}\\ &\equiv 2\cdot\frac{(p-1)!}{(p-1)!}\\ &\equiv 2\end{aligned}$$

Note that we have used the fact that p is an odd prime, s.t. $(-1)^{p-1}\equiv 1$. Similarly, one can prove that

$$\binom{3p-3}{2p-3}\equiv 2, \binom{2p-2}{p-2}\equiv 1, \binom{2p-3}{p-3}\equiv 1. \tag{3.8.3}$$

Combining the modulo equivalence in (3.8.2) and (3.8.3), (3.8.1) can be simplified to

$$3-2(p-1|X)-2(p|X)+(2p-1|X)+(2p|X)\equiv 0, \tag{3.8.4}$$

which is what we want to prove. □

Lemma 3.9. If $\left|X\right| = 4p-3$ and $(p|X)=0$, then $(p-1|X)\equiv(3p-1|X)$.

Proof. We consider the partition of $X = A \cup B \cup C$, where

$$\big|A\big| = p-1, \quad \big|B\big| = p-2, \quad \big|C\big| = 2p.$$

and

$$\sum_{a\in A} a \equiv 0, \quad \sum_{b\in B} b \equiv \sum_{x\in X} x, \quad \sum_{c\in C} c \equiv 0$$

Let χ denote the number of such partition. We use two ways to compute the number χ, the first one fixes A and find out the possible set C:

$$\chi \equiv \sum_{A} (2p|X-A) \equiv \sum_{A} -1 \equiv -(p-1|X),$$

where we have used Corollary 3.5, for $J = X - A$, with $\big|J\big| = 3p-2$ and the fact that

$$0 \leq (p|J) \leq (p|X) = 0,$$

which leads to $(p|J) = 0$. Now by fixing B and count the possible set C, we get:

$$\chi \equiv \sum_{B} (2p|X-B) \overset{1}{\equiv} \sum_{B} -1 \overset{2}{\equiv} \sum_{X-B} -1 \overset{3}{\equiv} -(3p-1|X)$$

For the three equivalences, we have used the following facts:

1. We use the similar argumentation as before, since $\big|X-B\big| = 3p-1$ and apply Corollary 3.5 leads to $(2p|X-B) \equiv -1$.

2. Consider the two sets

$$S := \Big\{\, B \subset X \;\Big|\; |B| = p-2, \sum_{b\in B} b \equiv \sum_{x\in X} x \,\Big\}$$
$$W := \Big\{\, J \subset X \;\Big|\; |J| = 3p-1, \sum_{j\in J} j \equiv 0 \,\Big\}$$

   The map T defined by:

$$\begin{aligned} T : S &\to W \\ B &\mapsto X - B \end{aligned}$$

   is a bijection, s.t.

$$\sum_{B} 1 \equiv \sum_{X-B} 1$$

3. Since $\sum_{b\in B} b \equiv \sum_{x\in X} x$, it follows that

$$\sum_{x\in X-B} x \equiv 0,$$

in particular

$$\sum_{X-B} -1 \equiv -1 \cdot (3p-1|X)$$

□

Now we are able to prove the *Kemnitz' Conjecture*:

Theorem 3.10. Any choice of $4p-3$ lattice points in the plane $\mathbb{Z}\times\mathbb{Z}$ contains a subset of cardinality p whose centroid is a lattice point.

Proof. For the proof of the theorem, we recall the equations we obatined from Corollary 3.7 and Corollary 3.8:

$$-1 + (p|X) - (2p|X) + (3p|X) \equiv 0 \tag{3.10.1}$$
$$(p-1|X) - (2p-1|X) + (3p-1|X) \equiv 0 \tag{3.10.2}$$
$$3 - 2(p-1|X) - 2(p|X) + (2p-1|X) + (2p|X) \equiv 0. \tag{3.10.3}$$

Adding the three above equations, we obtain:

$$2 - (p-1|X) - (p|X) + (3p-1|X) + (3p|X) \equiv 0 \tag{3.10.4}$$

Assume there is a set X, with $\left|X\right| = 4p-3$ which contradicts the theorem, that is $(p|X) = 0$. Using the previous Lemma 3.9, we obtain $(p-1|X) \equiv (3p-1|X)$ Then (3.10.4) simplifies to

$$2 - (p|X) + (3p|X) \equiv 0 \tag{3.10.5}$$

Since p is odd, we see that $(p|X)$ and $(3p|X)$ could not both be 0. Since we assume that $(p|X) = 0$, it follows that $(3p|X) > 0$, i.e., there is a subset $J \subset X$, $\left|J\right| = 3p$ and $\sum_{j\in J} j \equiv 0$. But from Corollary 3.6, we see that $(p|J) > 0$, in particular $(p|X) > 0$, which is a contradiction. □

4 Alon-Dubiner Theorem

In this chapter, I would like to state the proof the *Alon-Dubiner Theorem* which gives a general upper bound of the number $f(n,d)$. I will first discuss some preliminaries from the addtive combinatorics which we will need for the proof. For the following section, the capital letters, $A, B, Y \ldots$ all denote subset of a finite abelian group G.

4.1 Some inequalities from additive combinatorics

Definition 4.1. Let A and B two subsets of a finite abelian group G, define the addition and subtraction of the two sets as following:

$$A + B := \{ a + b \mid a \in A, b \in B \}$$
$$A - B := \{ a - b \mid a \in A, b \in B \}$$

Definition 4.2. Let A be a subset of a finite abelian group G, $g \in G$, the addition and subtraction of the element g and subset A is defined as following:

$$g + A := \{ g + a \mid a \in A \}$$
$$g - A := \{ g - a \mid a \in A \}$$

We have already seen *Cauchy-Davenport* theorem in Chapter 2. The following is another Lemma which characterizes the sum of two subsets which are sufficiently large.

Lemma 4.3. Let G be a finite abelian group, A, B are two subsets of G, satisfying

$$|A| > \frac{|G|}{2}, |B| > \frac{|G|}{2}$$

Then $A + B = G$

Proof. We randomly pick an element $g \in G$. Since $|g - A| = |A| > \frac{|G|}{2}$, and $|B| > \frac{|G|}{2}$. We must have

$$B \cap g - A \neq \emptyset$$

in particular, $\exists a \in A, b \in B$, s.t. $g = a + b$, since g is randomly picked, we obtain $A + B = G$ □

For the following discussion, it is helpful to introduce the concept of hyperplane, which is an analogy to that in the euclidean space.

Definition 4.4. We define a *hyperplane* of $\mathbb{Z}_p^d$ to be the set of all vectors $v \in \mathbb{Z}_p^d$ s.t. $\exists u \in \mathbb{Z}_p^d$, s.t. $u \cdot v = c$ for a constant $c \in \mathbb{Z}_p$. where $\cdot$ denotes the standard scalar product. An *affine basis* is a set $V \subset \mathbb{Z}_p^d$ with $\big|V\big| = d+1$, s.t. V is not contained in a hyperplane.

The sumset of $A + B$ has a lower bound which is given by the following Lemma, the key idea is based on *Plünnecke-Rusza inequality*. I refer the proof of the Lemma to [4] and [5].

Lemma 4.5. Let A be a subset of $\mathbb{Z}_p^d$ and let B be an affine basis of $\mathbb{Z}_p^d$. If $\big|A\big| \leq x^d$ for some integer $x \leq \frac{p}{4d}$ then $\big|A+B\big| \geq \big|A\big|2^{\frac{1}{2x}}$

From the above Lemma, we could derive the following Proposition, which will be used for the proof of the *Alon-Dubiner Theorem*.

Proposition 4.6. Let $x \leq \frac{p}{4d}$ be a power of 2. Let $A_1, \ldots, A_s$ be s affine spaces of $\mathbb{Z}_p^d$, with $s = 4xd$. Then

$$\big|A_1 + \ldots + A_s\big| \geq x^d.$$

Proof. We start with $A_0 := 0$ and set $x_1 = 2$. We apply Lemma 4.5 to the set $A = A_0, B := A_1$, then it follows that $A + B = A_0 + A_1 = A_1$. Besides,

$$\big|A_0 + A_1\big| \geq \big|A_0\big|2^{\frac{1}{2x_1}} \geq 2^{\frac{1}{2x_1}}$$

We then update B to A_2 , A to $A+B$ and do the same procedure recursively, thus

$$\begin{aligned}
\big|A_1 + A_2\big| &\geq \big|A_1\big| \cdot 2^{\frac{1}{2x_1}} \geq 2^{\frac{2}{2x_1}}, && A \leftarrow A_1 + A_2 \\
\big|A_1 + A_2 + A_3\big| &\geq \big|A_1 + A_2\big| \cdot 2^{\frac{1}{2x_1}} \geq 2^{\frac{3}{2x_1}}, && A \leftarrow A_1 + A_2 + A_3 \\
& && \vdots \\
\big|A_1 + \ldots + A_m\big| &\geq \big|A_1 + \ldots + A_{m-1}\big| \cdot 2^{\frac{1}{2x_1}} \geq 2^{\frac{m}{2x_1}}, && A \leftarrow A_1 + \ldots + A_m
\end{aligned}$$

Because

$$\frac{m}{2x_1} = d \Leftrightarrow m = 2x_1 d$$

It follows that after $j \leq 2x_1 d$ steps, we could guarantee that $\left|A_1 + \ldots + A_j\right| \geq x_1^d$. Choose $i_1 := \inf\left\{ j \;\middle|\; \left|A_1 + \ldots + A_j\right| \geq x_1^d \right\}$ and set $A = A_1 + \ldots + A_{i_1}$. Now we put $x_2 = 2^2$. Note that at present stage, $\left|A\right| \geq x_1^d$.
Applying again Lemma 4.5, we obtain

$$\begin{aligned} \left|A + A_{i_1+1}\right| &\geq x_1^d \cdot 2^{\frac{1}{2x_2}}, \quad A \leftarrow A + A_{i_1+1} \\ &\vdots \\ \left|A + A_m\right| &\geq x_1^d \cdot 2^{\frac{j-i_1}{2x_2}}, \quad A \leftarrow A + A_m \end{aligned}$$

Because

$$\frac{m}{2x_2} = d \Leftrightarrow m = 2(i_1 + x_2)d \leq 2(x_1 + x_2)d,$$

where we have used the fact that $i_1 \leq 2dx_1$ for the inequality. It follows that after $j \leq 2(x_1 + x_2)d$ steps in total, we could guarantee that

$$\left|A_1 + \ldots + A_j\right| \geq x_1^d \cdot 2^d = x_2^d.$$

Choose $i_2 := \inf\left\{ j \;\middle|\; \left|A_1 + \ldots + A_j\right| \geq x_2^d \right\}$ and set $A = A_1 + \ldots + A_{i_2}$. Now we put $x_3 = 2^3$ and do the similar procedure. Note that now $\left|A\right| \geq x_2^d$. Since x is a power of 2, we assume $x = 2^l$. From the discussion above, we could see that after i_j steps, $\left|A\right| \geq x_j^d = 2^{jd}$. In particular after i_l steps, $\left|A\right| \geq 2^{ld} = x^d$. In addition, we observe that $i_j \leq 2(x_1 + \ldots + x_j)d$, so at the l-th recursion, we obtain

$$\begin{aligned} i_l &\leq 2(x_1 + \ldots + x_l)d \\ &= 2(2 + 2^2 + \cdots + 2^l)d \\ &\leq 2 \cdot 2^{l+1} \cdot d \\ &= 4 \cdot 2^l \cdot d = 4xd = s \end{aligned}$$

In particular we see that $\left|A_1 + \ldots + A_s\right| \geq x^d$. □

We state the following Lemma which gives the estimation of the cardinality of the set $(ia + Y) \setminus Y$. A proof can be found in [5].

Lemma 4.7. Let Y be a finite subset of an abelian group G. Choose an $g \in G$ and $i \in \mathbb{N}$. Then

$$\left|(-g + Y) \setminus Y\right| = \left|(g + Y) \setminus Y\right| \tag{4.7.1}$$

$$\left|(ig + Y) \setminus Y\right| \leq i \cdot \left|(g + Y) \setminus Y\right| \tag{4.7.2}$$

4.2 characters

Throughout this section, we assume G is a finite abelian group, using additive notation. Let S a multiset of G, with S being symmetric, which means:

$$\forall x \in S \Rightarrow -x \in S$$

Definition 4.8. Let G be an abelian group and S a multiset of G, S symmetric. The *Cayley graph* $\Gamma_{G,S}$ is defined as the graph (G, E), with vertices be the group element of G and the edge E defined as:

$$E := \left\{ (g, s) \;\middle|\; g \in G, s \in S \right\}.$$

We denote the adjacency matrix of the graph $\Gamma_{G,S}$ as $A_{G,S}$.

From the property of S, we could see that the adjacency matrix $A_{G,S}$ has the following property.

Proposition 4.9. $A_{G,S}$ is symmetric.

We know that since $A_{G,S}$ is symmetric, all its eigenvalues are real. Moreover, it turns out that we can precisely find out each eigenvalue of $A_{G,S}$. For that purpose, I would like to introduce the concept of characters. I first introduce the general concept of a character of a general group, then confine to the case of finite abelian groups.

Definition 4.10. Let Γ be a group, a function $f : \Gamma \to \mathbb{C}^*$ is a *character* of Γ if f is a group homomorphism, where $\mathbb{C}^*$ denotes the multiplicative group of the units of $\mathbb{C}$. We denote $\widehat{\Gamma}$ to be the set of all possible characters of Γ.

We now turn to the case of finite abelian group G. Using the notation from Definition 4.10, we denote the set of all characters of G as $\widehat{G}$. We could easily see that all $\forall \chi \in \widehat{G}$:

$$\chi(g)^{|G|} = \chi(g^{|G|}) = \chi(1) = 1,$$

that is the range of each character must be the $\left|G\right|$-th root of unit in $\mathbb{C}^*$. It is helpful to study the property of $\widehat{G}$, which we state in the following lemma.

Lemma 4.11. $\widehat{G}$ is a group with multiplication defined as the composition of characters, the neutral element be the characters which maps all elements to 1, (we denote it as 1) and the inverse as the conjugation

Proof. We first prove that $\widehat{G}$ is closed under multiplication. Because it is easily seen that given χ_1, χ_2, $\chi := \chi_1 \circ \chi_2$ satisfies the definition of a character. Let $\chi \in \widehat{G}$, we could also easily see that χ^* is also a character. Furthermore, we have

$$\chi(g) \cdot \chi^*(g) = 1, \quad \forall g \in G$$

clearly $\chi^* = \chi^{-1}$. □

An interesting property of the character is given in the following lemma.

Lemma 4.12. Let $\chi \in \widehat{G}$ which is not identitcal to 1-mapping, i.e., maps every element to 1. Then

$$\sum_{g \in G} \chi(g) = 0$$

Proof. Since χ is not identical to 1, $\exists h \in G$, with $\chi(h) \neq 1$. We obtain

$$\chi(h) \cdot \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(g + h) = \sum_{g \in G} \chi(g)$$

Reforming the equation we obtain:

$$(\chi(h) - 1) \cdot \sum_{g \in G} \chi(g).$$

Since $\chi(h) \neq 1$, it follows that $\sum_{g \in G} \chi(g) = 0$. □

We define an operator

$$\begin{aligned} \langle \cdot, \cdot \rangle : \widehat{G} \times \widehat{G} &\to \mathbb{C}^* \\ \langle f, h \rangle &\mapsto \sum_{g \in G} f(g) h^*(g), \end{aligned}$$

where $*$ denotes the conjugation. The characters have the following property:

Lemma 4.13. Let χ_1, χ_2 be two characters of G, which are not equal. Then

$$\langle \chi_1, \chi_2 \rangle = 0$$

Proof. We know from Lemma 4.11 that $\chi_1 \circ \chi_2^*$ is also a character. Since $\chi_1 \neq \chi_2$, it follows that $\exists g \in G$, s.t. $\chi_1(g) \neq \chi_2(g)$

$$\chi_1 \circ \chi_2^*(g) \neq 1,$$

which means that $\chi_1 \circ \chi_2^*$ is not 1-mapping. Using Lemma 4.12, we obtain

$$\begin{aligned}\langle \chi_1, \chi_2 \rangle &= \sum_{g \in G} \chi_1(g) \cdot \chi_2^*(g) \\ &= \sum_{g \in G} \chi_1 \circ \chi_2^*(g) \\ &= 0.\end{aligned}$$

□

Since all the functions of $\mathbb{C}^G$ can be treated as a $\mathbb{C}$-vector space, and the characters can be seen as a special case of $\mathbb{C}^G$ by extending the range to $\mathbb{C}$, we could find a natural bound of the cardinality of $\widehat{G}$.

Proposition 4.14. $\left|\widehat{G}\right| \leq \left|G\right|$

Proof. From the above discussion, we see that $\mathbb{C}^G$ has dimension $\left|G\right|$. We extend the bilinear form $\langle \cdot, \cdot \rangle$ to $\mathbb{C}^G \times \mathbb{C}^G$, we see that this is a scalar product for the vector space. From Lemma 4.13, we see that different characters are orthogonal to each other. In particular linear independent, the bound follows from the fact that the number of linear independent vectors is bounded by the dimension of the vector space. □

Next we would like to prove that $\left|\widehat{G}\right| = \left|G\right|$. We recall the structure theorem of finite abelian groups that

$$G \cong \mathbb{Z}/n_1\mathbb{Z} \bigoplus \cdots \bigoplus \mathbb{Z}/n_s\mathbb{Z}, \quad \text{for some } n_1, \ldots, n_s \in \mathbb{N}$$

It can be easily seen that if $\chi_1 \in \widehat{G_1}$ and $\chi_2 \in \widehat{G_2}$, then the function χ defined on $G_1 \times G_2$. which has the value $\chi(g_1, g_2) = \chi_1(g_1) \cdot \chi_2(g_2))$ is clearly a character. By setting one of the $g_i, i \in \{1, 2\}$ to 1, we could see that this mapping is injective. We state the result in a proposition.

Proposition 4.15. Let G_1, G_2 be two finite abelian groups and $\widehat{G}G_1, \widehat{G}G_2$ the corresponding character groups. The map:

$$\begin{aligned}&\widehat{G_1} \times \widehat{G_2} \to \widehat{G_1 \times G_2}\\ &(\chi_1, \chi_2) \mapsto \chi : (g_1, g_2) \mapsto \chi_1(g_1) \cdot \chi_2(g_2)\end{aligned}$$

is injective.

Once we have proved that $\left|\widehat{G_i}\right| = \left|G_i\right|, i \in \left\{1, 2\right\}$, we can obtain from Proposition 4.15 that $\left|\widehat{G}\right| \geq \left|G\right|$, combining the result with Proposition 4.14, we could obtain that

$$\left|\widehat{G}\right| = \left|G\right|$$

So we only need to focus on a single decomposition.

Lemma 4.16. The character group of $\mathbb{Z}/n\mathbb{Z}$ has cardinality n.

Proof. Each character χ can be uniquely defined by $\chi(1)$ since $\mathbb{Z}/n\mathbb{Z}$ is cycle. The map

$$\chi(1) \mapsto \exp^{\frac{2\pi k}{n}}, \quad k = 0, \ldots, n-1$$

uniquely determines n characters. □

Combining Proposition 4.14 and Lemma 4.16, we can find the number character groups of a finite abelian group which we state in the following lemma.

Lemma 4.17. For every finite abelian group G, it holds that

$$\left|\widehat{G}\right| = \left|G\right|$$

Now we can examine the eigenvalues of the adjacency matrix of the cayley graph. With the help of characters, the eigenvalues can be easily computed:

Proposition 4.18. Let $A_{G,S}$ be the adjacency matrix of a cayley graph $\Gamma_{G,S}$. Then all the eigenvalues of $A_{G,S}$ is given by

$$\sum_{s \in S} \chi(s) \tag{4.18.1}$$

where $\chi \in \widehat{G}$. The corresponding eigenvector is given as

$$\begin{pmatrix} \chi(g_1) \\ \vdots \\ \chi(g_n) \end{pmatrix},$$

where we have enumerated the elements of G as $g_1, \ldots, g_n$.

Proof. Since G is an finite abelian group. From Lemma 4.17, we know that $\big|characterGroup\big| = \big|G\big|$. In particular the vector

$$v_\chi := \begin{pmatrix} \chi(g_1) \\ \vdots \\ \chi(g_n) \end{pmatrix},$$

is dinstict for different $\chi \in \widehat{G}$. Since $A_{G,S}$ is symmetric, it has exactly n eigenvectors. We are done if we prove that the eigenvalue of v_χ is given by (4.18.1). We define $A := A_{G,S}$ and $\chi_g := \chi(g)$ for convenience of notation.

$$\begin{aligned} (Av_\chi)_g &= \sum_h A_{g,h} v_{\chi,h} \\ &= \sum_{s \in S} \chi(g+s) \\ &= \chi_g \cdot \sum_{s \in S} \chi(s) \end{aligned}$$

□

Since $\Gamma_{G,S}$ is $\big|S\big|$-regular. It has an eigenvalue of $\big|S\big|$. Moreover, based on Proposition 4.18, we deduce the following proposition.

Proposition 4.19. The largest eigenvalue of $A_{G,S}$ is $\big|S\big|$.

Proof. From Proposition 4.18, all the eigenvalue is given by

$$\sum_{s \in S} \chi(s)$$

Since the range of the character is constrained in the unit circle in the complex plane. It follows that for an arbitrary eigenvalue α of $A_{G,S}$

$$\begin{aligned}\big|\alpha\big| &= \big|\sum_{s\in S}\chi(s)\big| \\ &\leq \sum_{s\in S}\big|\chi(s)\big| \\ &\leq \sum_{s\in S} 1 \\ &= \big|S\big|\end{aligned}$$

□

We state the following important lemma which gives a lowerbound between number of edges that spans two disjoint set of the nodes. A proof of the Lemma can be found in [6].

Lemma 4.20. Let λ be the largest eigenvalue of $A_{G,S}$, and μ the second largest eigenvalue. Let $A \subset G$ and $B = G \setminus A$. We define

$$e(A,B) = \big|\big\{\ (u,v)\ \big|\ u \in A, v \in B\ \big\}\big|,$$

i.e., the number of edges that have one end in A and another end in B. It holds that $e(A,B)$ has a lower bound given by the following equaiton

$$e(A,B) \geq (\lambda - \mu)\cdot\frac{\big|A\big|\big|B\big|}{\big|G\big|}$$

We now state a Lemma which is useful for the proof of the *Alon-Dubiner Theorem*.

Lemma 4.21. Suppose $W \geq 1$, let A be a sequence of elements in $\mathbb{Z}_p^d$ (possible to have equal elements) and suppose no hyperplanes in $\mathbb{Z}_p^d$ contains more than $\frac{|A|}{4W}$ members of A. Then $\forall Y \subset \mathbb{Z}_p^d$, with $\big|Y\big| \leq \frac{p^d}{2}$, $\exists a \in A$, s.t.

$$\big|(a+Y)\setminus Y\big| \geq \frac{W}{16p}\big|Y\big| \tag{4.21.1}$$

Remark 4.22. It follows from the condition that

$$\begin{aligned} p \cdot \frac{\big|A\big|}{4W} &\geq \big|A\big| \\ \frac{p}{4} &\geq W. \end{aligned}$$

Proof. We put $G = \mathbb{Z}_p^d$. The idea is that we construct $Z := G \setminus Y$ and use Lemma 4.21 to find an element which satisfies (4.21.1) We define the multiset S_a for every $a \in A$.

$$S_a := \big\{a, \ldots, 2\lceil \frac{p}{W} \rceil a\big\} \cup \big\{-a, \ldots, -2\lceil \frac{p}{W} \rceil a\big\}.$$

We define the multiset S as

$$S := \bigcup_{a \in A} S_a.$$

We construct the cayley graph $\Gamma_{G,S}$ and from Proposition 4.19, we see that the largest eigenvalue $\lambda = \big|S\big| = 4\lceil \frac{p}{W} \rceil \cdot \big|A\big|$. We would like to find out the upper bound of the second largest eigenvalue. Recall from Proposition 4.15 that for each character χ, it has a from of:

$$\chi((g_1, \ldots, V_d)) = \prod_{l=1}^{d} \chi_l(g_l),$$

where we have decomposed each element of $g \in G$ in to the component in the direct sum, and χ_l is the character of the group $\mathbb{Z}_p$. From Lemma 4.16, we see that each χ_l has the form of

$$\chi_l(g_l) = \exp(\frac{2\pi i \cdot g_l v_l}{p}),$$

for some $v_l \in \mathbb{Z}_p$. As a consequence we can write the χ as the following:

$$\begin{aligned} \chi(g) &= \prod_{l=1}^{d} \exp(\frac{2\pi i \cdot v_l}{p}) \\ &= \exp(\frac{2\pi i \sum_l g_l v_l}{p}) \end{aligned}$$

For some $v := (v_1, \ldots, v_d)$ in $\mathbb{Z}_p^d$. From Lemma 4.17 we could see that each such v determines uniquely a character, thus we can specify every character with a vector $v \in \mathbb{Z}_p^d$. From Proposition 4.18, the eigenvalue is given by

$$\sum_{s \in S} \omega^{v \cdot s} = \sum_{a \in S_a} \sum_{s \in S_a} \omega^{v \cdot s},$$

where $\omega = \exp(\frac{2\pi i}{p})$, and $\cdot$ denotes the standard scalar product. Since we want to find out the bound of second largest eigenvalue, we set $v \neq 0$.($v = 0$ leads to the largest eigenvalue). We can choose the range of the dot product s.t. for $a \in A$,

$$\frac{-p}{2} < v \cdot a \leq \frac{p}{2}$$

By assumption for each $c \in \mathbb{Z}_p$, there are at most $\frac{|A|}{4W}$ elements $a \in A$ s.t. $v \cdot a = c$. In particular:

$$\begin{aligned} \left|\left\{ a \in A \;\middle|\; v \cdot < W \right\}\right| &\leq \frac{|A|}{4W} \cdot (W - 1 - (-W + 1) + 1) \\ &= \frac{|A|}{4W} \cdot (2W - 1) < \frac{|A|}{2}, \end{aligned}$$

which means that there are at least $\frac{|A|}{2}$ elements $a \in A$, such that $v \cdot a \geq |W|$. For each such a, if $k = |v \cdot a|$ and $l = 2\lceil \frac{p}{W} \rceil$, then

$$\begin{aligned} \sum_{s \in S_a} \omega^{v \cdot s} &= \sum_{j=1}^{l} \omega^{(}k \cdot j) + \sum_{j=1}^{l} \omega^{(} - k \cdot j) \\ &= \frac{\omega^k(1 - \omega^{kl})}{1 - \omega^k} + \frac{\omega^{-k}(1 - \omega^{kl})}{1 - \omega^{-k}}. \end{aligned}$$

The amplitude is thus bounded by

$$\begin{aligned} \left|\sum_{s \in S_a} \omega^{v \cdot s}\right| &\leq \frac{\left|\omega^k\right| 2 \cdot \left|1 - \omega^{kl}\right|}{\left|1 - \omega^k\right|} \\ &\leq \frac{4}{\left|1 - \omega^k\right|} \\ &\overset{(*)}{\leq} \frac{4}{\frac{\pi W}{p}} = \frac{4}{\pi} \cdot \frac{p}{W} \\ &< \frac{2p}{W} \leq l. \end{aligned} \tag{4.22.1}$$

Where we have used the fact at $(*)$ that

$$\left|1-\omega^k\right|^2 \geq 2-2\cos(2\theta), \quad \theta = \frac{\pi W}{p}$$

By analysing the difference of the squared amplitude, we could see that see that

$$2-2\cos(2\theta)-\theta^2 \geq 0, \quad \theta \in (0, \frac{\pi}{2})$$

which means that $\left|1-\omega^k\right| \geq \frac{\pi W}{p}$.
Now since (4.22.1) holds for at least $\frac{|A|}{2}$ elements, we set the multiset

$$B := \left\{ a \in A \;\middle|\; v \cdot a \geq W \right\},$$

and the multiset C which is a submultiset of B which has cardinality $\frac{\left|A\right|}{2}$. we obtain that

$$\begin{aligned}
\Big|\sum_{a\in S_a}\sum_{s\in S_a}\omega^{v\cdot s}\Big| &\leq \Big|\sum_{a\in C}\sum_{s\in S_a}\omega^{v\cdot s}\Big| + \Big|\sum_{a\in A\setminus C}\sum_{s\in S_a}\omega^{v\cdot s}\Big| \\
&\leq \sum_{a\in A\setminus C}\Big|\sum_{s\in S_a}\omega^{v\cdot s}\Big| + \sum_{a\in A\setminus C}\sum_{s\in S_a}\big|\omega^{v\cdot s}\big| \\
&\leq \sum_{a\in C} l + \sum_{a\in A\setminus C}\sum_{s\in S_a} 1 \\
&\leq \frac{\left|A\right|}{2}\cdot l + \frac{\left|A\right|}{2}\cdot 2l \\
&\leq \frac{3\left|A\right|}{2} l = \frac{3}{4}\cdot\lambda
\end{aligned}$$

Let $Y \subset G$ with $\left|Y\right| \leq \frac{|G|}{2}$ and define $Z := G \setminus Y$. From Lemma 4.20, we know that $e(Y,Z)$, i.e., the number of edges which has one end in Y and the other in Z has a lower bound:

$$\begin{aligned}
e(Y,Z) &\geq (\lambda-\mu)\frac{\left|Y\right|\left|Z\right|}{\left|G\right|} \\
&\geq \frac{\lambda}{4}\frac{\left|Y\right|\frac{|G|}{2}}{\left|G\right|} \\
&\geq \lambda\cdot\frac{\left|Y\right|}{8}.
\end{aligned}$$

This implies that the set $M := \left\{ (y,s) \;\middle|\; y \in Y, s \in S, y+s \notin Y \right\}$ has at least $\lambda \cdot \frac{|Y|}{8}$. elements. We could observe that by averaging, there is at least one element $s \in S$, s.t. $M_s := \left\{ y \in Y \;\middle|\; y+s \notin Y \right\}$ has more than $\frac{|Y|}{8}$ elements, because otherwise $\forall s \in S, |M_s| < \frac{|Y|}{8}$. It follows that

$$\begin{aligned} |M| &= |\bigcup_{s\in S} M_s| \\ &\leq \sum_{s\in S} |M_s| \\ &< \sum_{s\in S} \frac{|Y|}{8} \\ &= \lambda \cdot \frac{|Y|}{8} \;\lightning \end{aligned}$$

Choose such s, we know from the construction of the set S that $s \in S_a$ for some $a \in A$, that is $\exists \epsilon, i \in \mathbb{N}, i \leq 2\lceil \frac{p}{W} \rceil$, s.t. $s = \epsilon i a$. From Lemma 4.7, we know that

$$\begin{aligned} |(s+Y) \setminus Y| &= |(\epsilon i a + Y) \setminus Y| \\ &= |(ia+Y) \setminus Y| \\ &\leq i \cdot (a+Y) \setminus Y, \end{aligned}$$

which follows that

$$|(a+Y) \setminus Y| \geq \frac{|Y|}{8 \cdot i} \geq \frac{|Y|}{8 \cdot 2\lceil \frac{p}{W} \rceil} \geq \frac{W}{16p}|Y|$$

□

4.3 Proof of *Alon-Dubiner Theorem*

We can now prove the *Alon-Dubiner Theorem*.

Theorem 4.23 (*Alon-Dubiner Theorem*)**.** $\exists c > 0$, s.t. $\forall n \in \mathbb{N}$,

$$f(n,d) < (cd \log_2 d)^d n$$

We first note that it suffices to prove the theorem for the prime number. Because if $f(p,d) \leq cp \leq 2cp-(2c-1)$, then it follows from (2.3.1) that for $n = pq$:

$$\begin{aligned} f(pq,d) \leq f(p,d) + p(f(q,d)-1) &\leq 2cp - 2c + 1 + p(2cq - 2cp) \\ &= 2cn - (2c-1) \leq 2cn \end{aligned}$$

As a consequence, we only prove the case when $n = p$ for p a prime number. It is worth noting that the constant c is fixed given a dimension d, but may vary once the dimesion d changes.
That is :

$$f(p,d) \leq c(d)p \tag{4.23.1}$$

We prove (4.23.1) by induction on d, with the constant $c(d)$ are defined by the following scheme:

$$c(1) = 2,\ c(d) = 256(d\log_2 d + 5)c(d-1) + (d+1), \text{for } d \geq 2 \tag{4.23.2}$$

We first check that the $c(d)$ satisfy the condition that $c(d) \leq (cd\log_2 d)^d$ for some absolute constant c. Since by induction:

$$\begin{aligned} c(d) &= 256(d\log_2 d + 5)c(d-1) + (d+1) \\ &\leq 256(d\log_2 d + 5)(c(d-1)\log_2(d-1))^{d-1} + (d+1) \\ &\leq 256(d\log_2 d + 5)(cd\log_2 d)^{d-1} + (d+1) \\ &\leq (256^{1/d}c^{\frac{d-1}{d}}d\log_2 d)^d + \mathcal{O}((d\log_2 d)^d) \\ &\leq (\tilde{c}d\log_2 d)^d \end{aligned}$$

Besides, we could see that for the case $p \leq 32d$, (4.23.1) is already true, because:

$$\begin{aligned} f(p,d) \overset{(*)}{\leq} (p-1)\cdot p^d + 1 = p^d \cdot p - p^d + 1 \\ &\leq (32d)^d \cdot p \\ &\leq (32d\log_2 d)^d p, \end{aligned}$$

where we have used (2.2.1) at (*). We have already seen from Theorem 2.7 that $f(p,1) = 2p-1$, so (4.23.1) holds for $d = 1, c(1) = 2$. Assuming the correctedness for $d-1$, we now prove the case d. We denote S be a sequence of $c(d)p$ elements from $\mathbb{Z}_p^d$. We consider two cases:

1. $\exists V \subset S$, with $\left|V\right| \leq c(d-1)p$, s.t all the elements of V lies on a hyperplane of $\mathbb{Z}_p^d$.
2. There are no such Vs.

For the first case, we can use the induction hypothesis to find a 0-sum p-subsequence in V, thus also in S.
For the second case, we want to show that there exists pairwise disjoint subsets $A_1, \ldots, A_s, B_1, \ldots, B_t, A_1^{'}, \ldots, A_{s'}^{'}, B_1^{'}, \ldots, B_{t'}^{'}$ of S, with the following properties:

1. $\left|A_i\right| = \left|A_j^{'}\right| = s+1, \quad \forall i = 1, \ldots, s, j = 1, \ldots, s'$
2. $\left|B_i\right| = \left|B_j^{'}\right| = 2, \quad \forall i = 1, \ldots, t, j = 1, \ldots, t'$
3. $s + t + s' + t' \leq p$
4. $\left|A_1 + \ldots + A_s + B_1 + \ldots + B_t\right| > \frac{p^d}{2}$
5. $\left|A_1^{'} + \ldots + A_{s'}^{'} + B_1^{'} + \ldots + B_{t'}^{'}\right| > \frac{p^d}{2}$

From property 5 and 6, we could see from Lemma 4.3 that

$$A_1 + \ldots + A_s + B_1 + \ldots + B_t + A_1^{'} + \ldots + A_{s'}^{'} + B_1^{'} + \ldots + B_{t'}^{'} = \mathbb{Z}_p^d. \quad (4.23.3)$$

which means, every element from $\mathbb{Z}_p^d$ can be written as a sum of $s+t+s'+t'$ elements in S. By choosing arbitrarily $m := p - s + t + s' + t'$ elements in S $s_1, \ldots, s_m$, that do not lie in $A_i, A_{i'}, B_j, B_{j'}$, we can see that from (4.23.3) that $\exists a_i \; in A_i, a_{i'}^{'} \in A_{i'}, b_j \in B_j, b_{j'}^{'} \in B_{j'}$, s.t

$$-(s_1 + \ldots + s_m) = a_1 + \ldots + a_s + b_1 + \ldots + b_t + a_1^{'} + \ldots + a_{s'}^{'} + b_1^{'} + \ldots + b_{t'}^{'}$$

By reforming the equation above we obtain

$$0 = a_1 + \ldots + a_s + b_1 + \ldots + b_t + a_1^{'} + \ldots + a_{s'}^{'} + b_1^{'} + \ldots + b_{t'}^{'} + s_1 + \ldots + s_m$$

which means, we have found a 0-sum p-subsequence in S. It remains to prove the existence of the sets $A_i, A_{i'}, B_j, B_{j'}$. We will construct them one by one. We first define $W = 64(d \log_2 d + 5d)$ and we see from (4.23.2):

$$\frac{c(d)p - p(d+1)}{4W} = c(d-1)p > \left|V \cap S\right| \quad \forall V \text{ hyperplane in } \mathbb{Z}_p^d. \quad (4.23.4)$$

This equation implies that even after we have deleted at most $p(d+1)$ elements from the set S, we can still make sure that no hyperplanes in $\mathbb{Z}_p^d$ contains $\frac{1}{4W}$ of the elements of the remaining sets. By property 1, 2, 3, we see that

$$\big|A_1 \cup \ldots \cup A_s \cup B_1 \cup \ldots \cup B_t \cup A_1^{'} \cup \ldots \cup A_{s'}^{'} \cup B_1^{'} \cup \ldots \cup B_{t'}^{'}\big| \leq p(d+1)$$

which means that the condition for the Lemma 4.21 will hold throughout the selection of the subsets $A_i, A_{i'}, B_j, B_{j'}$.

We now begin to construct A_i. First, we find x be a power of 2 which satisfies:

$$\frac{p}{32d} \leq x \leq \frac{p}{16d}.$$

Put $s = s' = 4xd$, which leads to $s \leq \frac{p}{4}$. Let $A_1, \ldots, A_s$ and $A_1^{'}, \ldots, A_{s'}^{'}$ be pairwise disjoint affine basis of $\mathbb{Z}_p^d$, each be a subset of S. The existence of those sets is justifed by (4.23.4), because it means that the remaing set of S does not lie on a hyperplane.(since each hyperplane contains less than $\frac{1}{4W}$ elements of the remaing set). By Proposition 4.6

$$\big|A_1 + \ldots + A_s\big| \geq x^d \geq (\frac{p}{32})^d. \tag{4.23.5}$$

and similar inequality holds for $A_1^{'}, \ldots, A_{s'}^{'}$. We next want to find the subset $B_1, \ldots, B_t$. We put $Y = A_1 + \ldots + A_s$. If $\big|Y\big| \geq \frac{p^d}{2}$ we are already done, since we could just set $t = 0$ to satisfy property 4.

$$\big|(a+Y) \cup Y\big| \geq (1 + \frac{W}{16p})\big|Y\big|.$$

Define $B_1 = a + s', s' \subset S'$, and observe that

$$\big|Y + B_1\big| = \big|(a+Y) \cup Y\big| = \big|(a+Y) \setminus Y\big| + \big|Y\big| \geq (1 + \frac{W}{16p})\big|Y\big|.$$

Next, we update $Y = A_1 + \ldots + A_s + B_1$ and update S' to be $S' \setminus B_1$. We stop this iterative process until the set Y has cardinality $\big|Y\big| > \frac{p^d}{2}$. We could check that the total step t will be bounded by $\frac{p}{4}$.

$$\frac{W}{16p}^{\,k} \cdot (\frac{p}{32d})^d \leq (1 + \frac{W}{16p})^k \big|A_1 + \ldots + A_s\big| \leq \frac{p^d}{2}$$

We deduce that

$$\begin{aligned}
\frac{W}{16p}^{k} \cdot (\frac{p}{32d})^{d} &\leq \frac{p^{d}}{2} \\
k &\leq \frac{16p}{W} \log_2(\frac{(32d)^{d}}{2}) \\
&\leq \frac{16p}{W} \log_2(32d)^{d} \\
&= \frac{16p}{64(d \log_2 d + 5d)} \cdot (d \log_2 d + 5d) \\
&= \frac{p}{4}
\end{aligned}$$

The set $B_j^{'}$ can be found similarly. We see that all the four variables s, t, s', t' are bounded by $\frac{p}{4}$, which means that $s + t + s' + t' \leq p$.

5 Conclusion

In this thesis, I have studied the number $f(n, d)$ and stated the proof of the *Kemnitz' Conjecture*. We see the generalization to the higher dimension which states that $f(n, d)$ is bounded linearly with respect to n. Further problems maybe to determine the subsequence in the 2-dimensional case which is guaranteed to be $4n - 3$. It is also conjectured that $f(n, d) \leq c^d n$ for all n and d.

References

[1] C. Reiher, "On kemnitz' conjecture concerning lattice-points in the plane," *The Ramanujan Journal*, vol. 13, p. 333–337, Jun 2007.

[2] M. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*. Graduate Texts in Mathematics, Springer New York, 1996.

[3] H. Harborth, "Ein extremalproblem für gitterpunkte.," *Journal für die reine und angewandte Mathematik*, vol. 0262_0263, pp. 356–360, 1973.

[4] I. Z. Ruzsa, "Sumsets and structure." `http://www.math.cmu.edu/users/af1p/Teaching/AdditiveCombinatorics/Additive-Combinatorics.pdf`. Accessed: 2021-1-4.

[5] N. Alon and M. Dubiner, "A lattice point problem and additive number theory," *Comb.*, vol. 15, no. 3, pp. 301–309, 1995.

[6] N. Alon and J. H. Spencer, *The Probabilistic Method*. Wiley Publishing, 4th ed., 2016.