DIMACS Technical Report 93-33
June 1993

INVERSE PROBLEMS IN ADDITIVE NUMBER THEORY

by

Melvyn B. Nathanson [1]
Department of Mathematics
Lehman College (CUNY)
Bronx, New York 10468

[1]This book was written while I was a long-term visitor at DIMACS. I am grateful to DIMACS for its hospitality. This work was also supported in part by grants from the PSC-CUNY Research Award Program.

DIMACS is a cooperative project of Rutgers University, Princeton University, AT&T Bell Laboratories and Bellcore.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

Warning to the reader

This is a preliminary and partial version of a book on inverse problems in additive number theory. Some chapters are not yet written, or are incomplete. Other chapters have not been checked, and contain typographical and mathematical errors. You read this book "at your own risk." Please be patient with the manuscript. I would very much appreciate receiving your comments and corrections, substantive or typographical, no matter how trivial. My e-mail address is: nathansn@dimacs.rutgers.edu.

The major inverse theorems are due to Kneser, Vosper, and Freiman. The principal goal of the chapters in this preprint is to give a self-contained account of Ruzsa's proof of Freiman's inverse theorem. The only prerequisite is elementary algebra and number theory. I include proofs of such "standard" results as Menger's theorem in graph theory and Minkowski's theorem on successive minima in the geometry of numbers.

I wish to thank Imre Z. Ruzsa for providing me with preprints of his papers on Freiman's theorem, Endre Szemerédi for describing to me his joint result with Antal Balog, and Gregory Freiman for useful discussions about his work.

How to obtain this report

This is DIMACS Technical Report 93-33. It is stored on-line as a Postscript file, and is available by anonymous ftp to everyone. The following steps will assist you in obtaining DIMACS Technical Reports:

1. ftp to dimacs.rutgers.edu (IP address = 128.6.75.16)
2. Log in as user "anonymous", and leave your email address as password.
3. Change to the directory "pub/dimacs/TechnicalReports/TechReports". Within this directory there are two subdirectories of Technical Reports for the years 1992 and 1993.
4. Change to the directory "1993".
5. Type "get 93-33"
6. The "quit" command will end your ftp session.

For more information about the DIMACS INFO SERVICE, type "telnet dimacs.rutgers.edu" and at the login prompt type "info".

Contents

Warning to the reader i

1 What is an inverse theorem? 1

2 Simple inverse theorems 5
2.1 Finite arithmetic progressions 5
2.2 A special case 10
2.3 Small sumsets: The case $|2A| \leq 3k-4$ 12
2.4 Application: Sumsets and powers of 2 21
2.5 Notes 23
2.6 Exercises 23

3 Sums of sets of congruence classes 25
3.1 Addition in groups 25
3.2 The Cauchy-Davenport-Chowla theorem 26
3.3 Vosper's inverse theorem 28
3.4 Kneser's theorem for finite abelian groups 30
3.5 Notes 31
3.6 Exercises 31

4 The asymptotic density of sumsets 33
4.1 Asymptotic density 33
4.2 Kneser's theorem 35
4.3 Notes 35
4.4 Exercises 35

5 Applications of Kneser's theorem 37
5.1 The maximum density of minimal bases 37
5.2 Sumsets and powers of 2 37
5.3 On the order of co-finite subsets of bases 37
5.4 Notes 37
5.5 Exercises 37

6 Geometry of numbers **39**
6.1 Lattices and determinants 39
6.2 Bases for sublattices 43
6.3 Convex bodies and Minkowski's first theorem 48
6.4 Applications to sums of squares 48
6.5 Successive minima and Minkowski's second theorem 48
6.6 Torsion-free abelian groups 48
6.7 Notes 50
6.8 Exercises 50

7 Plünnecke's inequality **51**
7.1 Plünnecke graphs 51
7.2 Examples of Plünnecke graphs 52
7.3 Multiplicativity of magnification ratios 54
7.4 Menger's theorem 57
7.5 Plünnecke's inequality 60
7.6 Applications to sums of finite sets 64
7.7 Essential components 67
7.8 Notes 71
7.9 Exercises 71

8 Freiman's theorem **73**
8.1 Introduction 73
8.2 Multi-dimensional progressions 73
8.3 Freiman isomorphisms 75
8.4 Ruzsa's proof 79
8.5 Notes 89
8.6 Exercises 90

9 Large subsets of $A \times B$ **91**
9.1 Introduction 91
9.2 The Regularity Lemma 91
9.3 The Balog-Szemerédi theorem 103
9.4 Notes 109
9.5 Exercises 109

10 The geometry of sumsets **111**
10.1 Introduction 111
10.2 Linearly independent hyperplanes 113
10.3 Parallelepipeds 120
10.4 Proof of the theorem 129
10.5 Notes 138
10.6 Exercises 138

References **139**

Chapter 1

What is an inverse theorem?

Let $h \geq 2$, and let $A_1, A_2, \ldots, A_h$ be sets of integers. The *sumset*

$$A_1 + A_2 + \cdots + A_h$$

is the set of all integers of the form $a_1 + a_2 + \cdots + a_h$, where $a_i \in A_i$ for $i = 1, 2, \ldots, h$. If A is a set of integers and if $A_i = A$ for $i = 1, 2, \ldots, h$, then we denote the sumset $A_1 + A_2 + \cdots + A_h$ by hA. Thus, the h-fold sumset hA is the set of all sums of h elements of A, with repetitions allowed. If $0 \in A$, then $A \subseteq 2A \subseteq \ldots \subseteq hA \subseteq \ldots$. Sumsets can also be defined in any abelian group, and, indeed, in any set in which there is a binary operation. In particular, it is sometimes important to consider sumsets in the cyclic group $\mathbf{Z}/m\mathbf{Z}$ of congruence classes modulo m.

Let $\mathbf{N}$ denote the set of nonnegative integers. The set A of integers is a *basis of order* h if $hA = \mathbf{N}$, that is, if every nonnegative integer is the sum of h elements of A, with repetitions allowed. We write $A \sim B$ if the sets A and B coincide for all sufficiently large integers. This means that there exists a number n_0 such that, if $n \geq n_0$, then $n \in A$ if and only if $n \in B$. If A is a set of integers such that $hA \sim \mathbf{N}$, that is, if the sumset hA contains all sufficiently large integers, then A is called an *asymptotic basis of order* h.

A *direct problem* in additive number theory is a problem in which we try to determine the structure and properties of the h-fold sumset hA when the set A is known. For example, if $A^{(2)}$ is the set of squares, then well known theorems of Fermat, Gauss, and Lagrange describe the arithmetic structure of the sumsets $2A^{(2)}, 3A^{(2)}$, and $4A^{(2)}$, respectively. In particular, Lagrange proved that $4A^{(2)} = \mathbf{N}$. Cauchy proved that for every $k \geq 3$ the polygonal numbers of order k form a basis of order k.

Let $A^{(k)}$ denote the set of nonnegative k-th powers. Waring's problem is the theorem that $hA^{(k)} = \mathbf{N}$ for h sufficiently large, that is, the set of k-th powers forms a basis. Let $g(k)$ denote the smallest number such that $A^{(k)}$ is a basis of order $g(k)$. Lagrange's theorem states that $g(2) = 4$, and Wieferich proved that $g(3) = 9$. Landau discovered that every sufficiently large integer is the sum of at most eight cubes, and, indeed, Dickson found that 23 and 239 were the only numbers that could not be written as the sum of eight cubes. For any k, let $G(k)$ denote the smallest number such that $A^{(k)}$ is an asymptotic basis of order

$G(k)$. Thus, Landau proved that $G(3) \leq 8$, and Linnik subsequently showed that $G(3) \leq 7$. The best estimate of $G(k)$ for large k is due to Wooley.

If $\mathcal{P}$ is the set of odd primes, then Goldbach's conjecture, still unsolved, asserts that $2\mathcal{P}$ is the set of all even integers greater than four. Vinogradov has proved that $3\mathcal{P}$ contains all sufficiently large odd integers. Chen proved that every sufficiently large even integer can be written in the form $p + q$, where p is a prime, and q is the product of at most two primes. These are the classical direct problems in additive number theory.

An *inverse problem* in additive number theory is a problem in which we attempt to deduce properties of the set A from properties of the sumset hA. For example, if A is a finite set of integers and if the cardinality of the h-fold sumset hA is small, what can we conclude about the structure of the set A? The simplest inverse theorem states: If A is a set of k integers, and if $|2A| = 2k - 1$, then A is an arithmetic progression. The proof is easy. Let

$$A = \{a_1, a_2, \ldots, a_k\},$$

where

$$a_1 < a_2 < \cdots < a_k.$$

Then the sumset $2A$ contains the $2k - 1$ integers

$$2a_1 < a_1 + a_2 < 2a_2 < \cdots < a_{i-1} + a_i < 2a_i < a_i + a_{i+1} < \cdots < 2a_k,$$

and so $|2A| \geq 2k - 1$. Suppose that $|2A| = 2k - 1$. Since

$$a_{i-1} + a_i < a_{i-1} + a_{i+1} < a_i + a_{i+1}$$

and

$$a_{i-1} + a_i < 2a_i < a_i + a_{i+1}$$

for $i = 2, \ldots, k - 1$, it follows that

$$2a_i = a_{i-1} + a_{i+1}$$

or, equivalently,

$$a_i - a_{i-1} = a_{i+1} - a_i$$

for $i = 2, \ldots, k - 1$. Thus, A is an arithmetic progression. This completes the proof.

The principal results in this book are due to Vosper, Kneser, and Freiman. In 1956, Vosper [75] obtained an inverse to the Cauchy-Davenport theorem. Cauchy had proved (and Davenport had rediscovered) that if A and B are sets of congruence classes modulo a prime p, then

$$|A + B| \geq \min(p, |A| + |B| - 1).$$

Vosper determined the structure of the "critical pairs" A and B such that $A + B \neq \mathbf{Z}/p\mathbf{Z}$ and $|A + B| = |A| + |B| - 1$. In particular, if $|A| \neq 1$ and $|2A| < 2|A| < p$, then A is a arithmetic progression in the group $\mathbf{Z}/p\mathbf{Z}$.

The *lower asymptotic density* $d_L(A)$ of a set A of integers is defined by

$$d_L(A) = \liminf_{n\to\infty} A(n)/n,$$

where $A(n)$, the *counting function* of the set A, denotes the number of positive elements of A that do not exceed n. In 1953, Kneser[38] proved that if A is a set of integers such that $d_L(2A) < 2d_L(A)$, then A is a "large" subset of a union of arithmetic progressions with the same common difference m, that is, A is contained in a set of congruence classes modulo m. If the modulus is a prime p, then it follows from Vosper's theorem that these congruence classes form an arithmetic progression in $\mathbf{Z}/p\mathbf{Z}$ with difference d, and so A is a subset of a set of the form

$$\{a_0 + x_1 d + x_2 m \mid 0 \le x_1 < l_1 \text{ and } x_2 \in \mathbf{Z}\}.$$

A set with this structure is called an infinite 2-dimensional arithmetic progression.

A finite n-dimensional arithmetic progression is a set of the form

$$\{a_0 + a_1 x_1 + \cdots + a_n x_n \mid 0 \le x_i < l_i \text{ for } i = 1, \ldots, n\}.$$

In 1964, Freiman [28] discovered a deep and beautiful theorem about the structure of finite sets of integers with small sumsets. If A is a finite set of integers and if

$$|2A| \le c|A|,$$

then A is a subset of an n-dimensional arithmetic progression Q, where $|Q| \le c'|A|$, and n and c' are constants that depend only on c. Ruzsa has extended this result to sumsets of the form $A + B$, where A and B are finite subsets of a torsion-free abelian group.

Nothing is known, however, about the structure of a finite set A if, for example,

$$|2A| \le |A|^{2-\delta}$$

for some $\delta > 0$, or even if

$$|2A| \le c|A| \log |A|.$$

Nor is anything known about the structure of A if, for some $h \ge 3$,

$$|hA| \le c|A|^{h-1}$$

or even

$$|hA| \le c|A|^2.$$

Closely related to these "structural" inverse problems is another class of inverse problems in additive number theory that we can call "recognition problems." If we are given a finite or infinite set $\mathcal{B}$ of integers, can we determine if $\mathcal{B}$ is a sumset? This means the following: Let $h \ge 2$. Does there exist a set A such that $hA = \mathcal{B}$? Or do there exist sets $A_1, \ldots, A_h$ such that $|A_i| \ge 2$ for $i = 1, 2, \ldots, h$ and $A_1 + \cdots + A_h = \mathcal{B}$? A variant of this problem is: Does $\mathcal{B}$ contain a sumset? The twin prime conjecture is a special case of this inverse problem: Let $\mathcal{P}$ be the set of odd prime numbers. Does there exist an infinite set A such that

$$A + \{0, 2\} \subseteq \mathcal{P}?$$

Practically nothing is known about these problems.

Chapter 2

Simple inverse theorems

2.1 Finite arithmetic progressions

In this chapter we shall completely describe the structure of a finite set of integers when its 2-fold sumset is very small.

Notation. Let $|A|$ denote the cardinality of the set A. Let $\mathbf{Z}$ denote the set of all integers, and let $\mathbf{N}$ denote the set of all nonnegative integers. Let $[a, b]$ denote the *interval of integers* n such that $a \leq n \leq b$. Denote by $(a_1, \ldots, a_k)$ the greatest common divisor of the integers $a_1, \ldots, a_k$.

If $A = \{a_0, a_1, \ldots, a_{k-1}\}$ is a finite set of integers such that $a_0 < a_1 < \cdots < a_{k-1}$, we define

$$d(A) = (a_1 - a_0, a_2 - a_0, \ldots, a_{k-1} - a_0).$$

Let $b_i = (a_i - a_0)/d(A)$ for $i = 0, 1, \ldots, k-1$, and let

$$A^N = \{b_0, b_1, \ldots, b_{k-1}\}.$$

The set A^N is called the *normal form* of A. Clearly, $0 = b_0 < b_1 < \cdots < b_{k-1}$ and

$$d(A^N) = (b_1, \ldots, b_{k-1}) = 1.$$

Moreover,

$$hA = \{ha_0\} + \{da' \mid a' \in A^N\},$$

and so $|hA| = |hA^N|$.

For example, let $A = \{8, 29, 71, 92\}$ and $h = 2$. Then $d(A) = 21$, $A^N = \{0, 1, 3, 4\}$, $2A^N = [0, 8]$, and $2A = \{16 + 21n | n \in [0, 8]\}$.

There are the following trivial lower and upper bounds for the cardinality of sums of finite sets.

Theorem 2.1 *Let $h \geq 2$. Let A be a finite set of integers with $|A| = k$. Then*

$$hk - (h-1) \leq |hA| \leq k(k+1)(k+2)\cdots(k+h-1)/h!.$$

Proof. Let $A = \{a_0, a_1, \ldots, a_{k-1}\}$, where $a_0 < a_1 < \cdots < a_{k-1}$. Then

$$hA \supseteq \{ha_0\} \cup \bigcup_{r=1}^{k-1} \{(h-x)a_{r-1} + ia_r | i \in [1,h]\}.$$

Since

$$ha_{r-1} < (h-1)a_{r-1} + a_r < \cdots < a_{r-1} + (h-1)a_r < ha_r,$$

for $r = 1, \ldots, k-1$, it follows that

$$|hA| \geq 1 + (k-1)h = hk - (h-1).$$

This gives the lower bound.

The upper bound follows from the fact that the number of expressions of the form $a_1 + \cdots + a_h$ with $a_i \in A$ for $i = 1, 2, \ldots, h$ and $a_1 \leq \cdots \leq a_h$ is exactly $k(k+1)(k+2)\cdots(k+h-1)/h!$. This completes the proof.

Theorem 2.2 *Let $h \geq 2$, and let $A_1, A_2, \ldots, A_h$ be finite sets of integers. Then*

$$|A_1| + \cdots + |A_h| - (h-1) \leq |A_1 + \cdots + A_h| \leq |A_1| \cdots |A_h|.$$

Proof. We shall prove the lower bound by induction on h. Let $h = 2$, and let $A_1 = A = \{a_1, \ldots, a_k\}$ and $A_2 = B = \{b_1, \ldots, b_l\}$, where $a_1 < a_2 < \cdots < a_k$ and $b_1 < \cdots < b_l$. Suppose that $|A| = k \leq l = |B|$. The sumset $A + B$ contains the $k + l - 1$ distinct elements

$$\begin{aligned} a_1 + b_1 &< a_1 + b_2 < a_2 + b_2 < a_2 + b_3 < a_3 + b_3 < \cdots < \\ &< a_{i-1} + b_i < a_i + b_i < a_i + b_{i+1} < a_{i+1} + b_{i+1} \\ &< \cdots < a_k + b_k < a_k + b_{k+1} < \cdots < a_k + b_l, \end{aligned}$$

and so $|A_1| + |A_2| \geq |A_1| + |A_2| - 1$. Let $h \geq 3$, and suppose that the lower bound holds for the sum of any $h-1$ finite sets of integers. Then

$$\begin{aligned} |A_1 + \cdots + A_{h-1} + A_h| &= |(A_1 + \cdots + A_{h-1}) + A_h| \\ &\leq |A_1 + \cdots + A_{h-1}| + |A_h| - 1 \\ &\leq |A_1| + \cdots + |A_{h-1}| - (h-2) + |A_h| - 1 \\ &= |A_1| + \cdots + |A_{h-1}| + |A_h| - (h-1), \end{aligned}$$

and so the lower bound holds for all $h \geq 2$.

The upper bound follows from the fact that the number of expressions of the form $a_1 + \cdots + a_h$ with $a_i \in A_i$ for $i = 1, 2, \ldots, h$ is exactly $|A_1| \cdots |A_h|$. This completes the proof.

Lemma 2.1 *Let A and B be finite sets of integers with $|A| = |B| > 1$. If $|A + B| = |A| + |B| - 1$, then A and B are arithmetic progressions with the same common difference.*

Proof. Let $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_k\}$, where $a_1 < a_2 < \cdots < a_k$ and $b_1 < \cdots < b_k$. The sumset $A + B$ contains the $2k - 1$ distinct elements

$$\begin{aligned} a_1 + b_1 &< a_1 + b_2 < a_2 + b_2 < \cdots < \\ &< a_{i-1} + b_i < a_i + b_i < a_i + b_{i+1} < a_{i+1} + b_{i+1} < \\ &< a_{i+1} + b_{i+2} < \cdots < a_k + b_k, \end{aligned}$$

and the condition $|A + B| = 2k - 1$ implies that the preceding sequence of integers is a complete list of the integers in $A + B$. Since

$$a_{i-1} + b_i < a_i + b_i < a_i + b_{i+1}$$

and

$$a_{i-1} + b_i < a_{i-1} + b_{i+1} < a_i + b_{i+1}$$

for $i = 2, \ldots, k - 1$, it follows that

$$a_{i-1} + b_{i+1} = a_i + b_i,$$

or, equivalently,

$$a_i - a_{i-1} = b_{i+1} - b_i \tag{2.1}$$

for $i = 2, \ldots, k - 1$. Similarly, the inequalities

$$a_i + b_i < a_i + b_{i+1} < a_{i+1} + b_{i+1}$$

and

$$a_i + b_i < a_{i+1} + b_i < a_{i+1} + b_{i+1}$$

imply that

$$a_i + b_{i+1} = a_{i+1} + b_i$$

or, equivalently,

$$a_{i+1} - a_i = b_{i+1} - b_i \tag{2.2}$$

for $i = 1, \ldots, k - 1$. It follows from equations (2.1) and (2.2) that there is a positive integer m such that

$$a_{i+1} - a_i = b_{i+1} - b_i = m$$

for $i = 1, \ldots, k - 1$. This completes the proof.

Lemma 2.2 *Let A and B be finite sets of integers with $|A| = k \geq 2$ and $|B| = l \geq 2$. If $|A + B| = k + l - 1$, then A and B are arithmetic progressions with the same common difference.*

Proof. Let $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_l\}$, where $a_1 < \cdots < a_k$ and $b_1 < \cdots < b_l$. Suppose that $k \leq l$, and let $0 \leq t \leq l-k$. Let $B = B_0^t \cup B_1^t \cup B_2^t$, where

$$\begin{aligned} B_0^t &= \{b_1, b_2, \ldots, b_t\} \\ B_1^t &= \{b_{t+1}, b_{t+2}, \ldots, b_{t+k}\} \\ B_2^t &= \{b_{t+k+1}, b_{t+k+2}, \ldots, b_l\}. \end{aligned}$$

Then

$$A + B \supseteq (a_1 + B_0^t) \cup (A + B_1^t) \cup (a_k + B_2^t). \tag{2.3}$$

Since

$$\begin{aligned} a_1 + B_0^t &\subseteq [a_1 + b_1, a_1 + b_t] \\ A + B_1^t &\subseteq [a_1 + b_{t+1}, a_k + b_{t+k}] \\ a_k + B_2^t &\subseteq [a_k + b_{t+k+1}, a_k + b_l], \end{aligned}$$

it follows that the three sumsets in (2.3) are pairwise disjoint. Since

$$\begin{aligned} |a_1 + B_0^t| &= t \\ |A + B_1^t| &\geq |A| + |B_1^t| - 1 = 2k - 1 \\ |a_k + B_2^t| &= l - t - k, \end{aligned}$$

it follows that

$$\begin{aligned} k + l - 1 &= |A + B| \\ &\geq |a_1 + B_0^t| + |A + B_1^t| + |a_k + B_2^t| \\ &\geq t + (2k - 1) + (l - t - k) \\ &= k + l - 1. \end{aligned}$$

Thus,

$$|A + B_1^t| = 2k - 1$$

for $t = 0, 1, \ldots, l-k$. The preceding lemma implies that there exists a positive integer m such that A is an arithmetic progression with difference m and B_1^t is an arithmetic progression with difference m for all $t = 0, 1, \ldots, l-k$. This means that B is an arithmetic progression with difference m, and the proof is complete.

Theorem 2.3 *Let $h \geq 2$, and let $A_1, A_2, \ldots, A_h$ be $h+1$ finite sets of integers. Then*

$$|A_1 + \cdots + A_h| = |A_1| + \cdots + |A_h| - (h-1) \tag{2.4}$$

if and only if the sets $A_1, \ldots, A_h$ are arithmetic progressions with the same common difference.

Proof. It is easy to check that equation (2.4) holds if the sets A_i are arithmetic progressions with the same common difference.

We shall prove that (2.4) implies that the sets A_i are arithmetic progressions with the same common difference. The proof is by induction on h. The case $h = 2$ is simply the preceding lemma with $A = A_1$ and $B = A_2$. Now let $h \geq 3$, and assume that the theorem holds in the case of $h - 1$ sets. For $j = 1, \ldots, h$, consider the $(h-1)$-fold sumset

$$B_j = A_1 + \cdots + A_{j-1} + A_{j+1} + \cdots + A_h.$$

Then

$$|B_j| \geq \sum_{\substack{i=1 \\ i \neq j}}^{h} |A_i| - (h-2).$$

Since

$$\sum_{i=1}^{h} |A_i| - (h-1) = |A_j + B_j| \geq |A_j| + |B_j| - 1 \geq \sum_{i=1}^{h} |A_i| - (h-1),$$

it follows that

$$|B_j| = \sum_{\substack{i=1 \\ i \neq j}}^{h} |A_i| - (h-2)$$

for $j = 0, 1, \ldots, h$, and so the h sets $A_1, \ldots, A_h$ are arithmetic progressions with the same common difference. This completes the proof.

Corollary 2.1 *Let $h \geq 2$. Let A be a finite set of integers with $|A| = k$. Then $|hA| = hk - (h-1)$ if and only if A is a k-term arithmetic progression.*

Proof. This follows form the preceding theorem by letting $A_i = A$ for $i = 1, \ldots, h$.

Corollary 2.2 *Let A be a finite set of integers with $|A| = k$. Then $|hA| = hk - (h-1)$ if and only if $A^N = [0, k-1]$.*

Corollary 2.1 gives a simple example of an inverse theorem in additive number theory: If A is a finite set of integers, and if, for some $h \geq 2$, the cardinality of the sumset hA is as small as possible, then A must be an arithmetic progression. Thus, the normal form A^N of the set A is an interval of consecutive integers. Geometrically, A^N is the set of lattice points contained inside some convex subset of the real line.

A general inverse problem in additive number theory can be stated as follows: Let A be a finite set of integers. Suppose that the sumset hA is "small." What does this imply about the arithmetic or geometric structure of A? In section 2.3 we shall prove that if $|A| = k$ and $|2A| \leq 3k - 4$, then A is a subset of an arithmetic progression of length at most $2k - 3$.

2.2 A special case

Theorem 2.4 *Let $h \geq 2$ and $k \geq 3$. Let*

$$A = [0, k-2] \cup \{k-1+b\},$$

where $b \geq 0$. Define integers q and r by

$$b = q(k-2) + r,$$

where $q \geq 0$ and $0 \leq r \leq k-3$. If $b \leq (h-1)(k-2)$, then

$$|hA| = hk - (h-1) + q(2h-q-1)(k-2)/2 + (h-q-1)r.$$

If $b \geq (h-1)(k-2)$, then

$$|hA| = hk - (h-1) + h(h-1)(k-2)/2.$$

Proof. Let $A = [0, k-2] \cup \{k-1+b\}$. If $b = 0$, then $q = r = 0$ and $A = [0, k-1]$. The result follows immediately.

The sumset hA is a union of $h+1$ not necessarily disjoint intervals:

$$\begin{aligned} hA &= \bigcup_{m=0}^{h} \{(h-m)[0,k-2] + \{k-1+b\}\} \\ &= \bigcup_{m=0}^{h} \{[0,(h-m)(k-2)] + \{m(k-1+b)\}\} \\ &= \bigcup_{m=0}^{h} [m(k-1+b), m(k-1+b) + (h-m)(k-2)] \\ &= \bigcup_{m=0}^{h} [m(k-1+b), h(k-2) + m(b+1)] \\ &= \bigcup_{m=0}^{h} I_m, \end{aligned}$$

where

$$\begin{aligned} I_m &= [m(k-1+b), h(k-2) + m(b+1)] \\ &= [m(k-1+b), x(k-1+b) + (h-m)(k-2)] \end{aligned}$$

and

$$|I_m| = (h-m)(k-2) + 1.$$

Notice that the intervals I_m "move to the right" in the sense that, as m increases from 0 to h, the sequence of right-hand endpoints of the intervals I_m is strictly increasing, and the sequence of left-hand endpoints of the intervals I_m is strictly increasing.

For $m = 1, \ldots, h$, the set $I_{m-1} \cup I_m$ will be the interval

$$[(m-1)(k-1+b), m(k-1+b) + (h-m)(k-2)],$$

if and only if

$$m(k-1+b) \leq (m-1)(k-1+b) + (h-m+1)(k-2) + 1,$$

which is equivalent to

$$b \leq (h-m)(k-2).$$

Let $0 < b \leq (h-1)(k-2)$. There exists a unique $t \in [1, h-1]$ such that

$$(h-t-1)(k-2) < b \leq (h-t)(k-2).$$

If $1 \leq m \leq t$, then $h - t \leq h - m$ and

$$b \leq (h-t)(k-2) \leq (h-m)(k-2).$$

It follows that $I_{m-1} \cup I_m$ is an interval for $m = 1, \ldots, t$, and so

$$J = \bigcup_{m=0}^{t} I_m = [0, h(k-2) + t(b+1)]$$

is an interval. If $t < m \leq h$, then $h - m \leq h - t - 1$ and

$$(h-m)(k-2) \leq (h-t-1)(k-2) < b.$$

It follows that the intervals $J, I_{t+1}, I_{t+2}, \ldots, I_h$ are pairwise disjoint. Therefore,

$$\begin{aligned}
|hA| &= |\bigcup_{m=0}^{h} I_m| \\
&= |J \cup \bigcup_{m=t+1}^{h} I_m| \\
&= |J| + \bigcup_{m=t+1}^{h} |I_m| \\
&= h(k-2) + t(b+1) + 1 + \sum_{m=t+1}^{h} ((h-m)(k-2)+1) \\
&= h(k-2) + t(b+1) + (h-t+1) + (k-2)\sum_{m=0}^{h-t-1} m \\
&= h(k-2) + tb + h + 1 + (k-2)(h-t)(h-t-1)/2 \\
&= hk - (h-1) + tb + (h-t)(h-t-1)(k-2)/2.
\end{aligned}$$

If $r = 0$, then $q = h - t$ and

$$\begin{aligned} |hA| &= hk - (h-1) + (h-q)q(k-2) + q(q-1)(k-2)/2 \\ &= hk - (h-1) + q(2h-q-1)(k-2)/2. \end{aligned}$$

If $r > 0$, then $q = h - t - 1$ and

$$\begin{aligned} |hA| &= hk - (h-1) + (h-q-1)(q(k-2)+r) + q(q+1)(k-2)/2 \\ &= hk - (h-1) + (h-q-1)r + q(2h-q-1)(k-2)/2. \end{aligned}$$

Let $b > (h-1)(k-2)$. Then

$$(h-m)(k-2) \leq (h-1)(k-2) < b$$

for $m = 1, 2, \ldots, h$, and so the intervals $I_0, I_1, \ldots, I_h$ are pairwise disjoint. Therefore,

$$\begin{aligned} |hA| &= |\bigcup_{m=0}^{h} I_m| = \\ &= \bigcup_{m=0}^{h} |I_m| \\ &= \sum_{m=0}^{h} ((h-m)(k-2)+1) \\ &= (h+1) + h(h+1)(k-2)/2 \\ &= hk - (h-1) + h(h-1)(k-2)/2. \end{aligned}$$

This completes the proof.

Corollary 2.3 *Let $k \geq 3$, and let*

$$A = [0, k-2] \cup \{k-1+b\}.$$

If $0 \leq b \leq k-3$, then

$$|2A| = 2k - 1 + b \leq 3k - 4.$$

If $b \geq k-2$, then

$$|2A| = 3k - 3.$$

2.3 Small sumsets: The case $|2A| \leq 3k - 4$

We proved that the set A must be an arithmetic progression if A is a finite set of integers whose 2-fold sumset is as small as possible. That is, if $|A| = k$ and $|2A| = 2k - 1$, then $A^N = [0, k-1]$. In this section we shall show that if $|A| = k$ and $|2A| \leq 3k - 4$, then A is a subset of a short arithmetic progression. More precisely, we shall prove that if $|A| = k$ and $|2A| = 2k - 1 + b \leq 3k - 4$, then $A^N \subseteq [0, k-1+b]$. The special case analyzed in Theorem 2.4 shows that the lower bounds obtained in this section are best possible.

Theorem 2.5 *Let $h \geq 2$ and $k \geq 3$. Let $A = \{a_0, a_1, \ldots, a_{k-1}\}$ be a set of integers such that $0 = a_0 < a_1 < \cdots < a_{k-1}$. If*

$$r = a_{k-1} - (k-1) \leq k-2,$$

then

$$|hA| \geq hk - (h-1) + (h-1)r.$$

Proof. Consider the set

$$\begin{aligned} S &= \{0\} \cup \bigcup_{m=0}^{h-1} \{ma_{k-1} + a_i \mid i = 1, \ldots, k-1\} \\ &= \{0\} \cup \bigcup_{i=1}^{k-1} \{ma_{k-1} + a_i \mid m = 0, 1, \ldots, h-1\} \end{aligned}$$

Clearly,

$$S \subseteq hA \subseteq [0, ha_{k-1}]$$

since $0 \in A$. Moreover,

$$\begin{aligned} 0 &< a_1 < a_2 < \cdots < a_{k-2} < a_{k-1} \\ &< a_{k-1} + a_1 < a_{k-1} + a_2 < \cdots < a_{k-1} + a_{k-2} < 2a_{k-1} \\ &< 2a_{k-1} + a_1 < 2a_{k-1} + a_2 < \cdots < 2a_{k-1} + a_{k-2} < 3a_{k-1} \\ &\vdots \\ &< (h-1)a_{k-1} + a_1 < \cdots < (h-1)a_{k-1} + a_{k-2} < ha_{k-1}, \end{aligned}$$

and so

$$|S| = h(k-1) + 1 = hk - (h-1).$$

Let $W = [1, a_{k-1}] \backslash A$. Then

$$|W| = a_{k-1} - (k-1) = r.$$

For $w \in W$, define the set $S(w)$ by

$$S(w) = \{ma_{k-1} + w | m = 0, 1, \ldots, h-1\}.$$

Then $S(w) \subseteq [1, ha_{k-1}]$ and $|S(w)| = h$. Moreover, the $r+1$ sets S and $\{S(w)\}_{w \in W}$ are pairwise disjoint, and

$$[0, ha_{k-1}] = S \cup \bigcup_{w \in W} S(w).$$

This implies that

$$hA = S \cup \bigcup_{w \in W} (S(w) \cap hA).$$

We shall prove that

$$|S(w) \cap hA| \geq h - 1$$

for all $w \in W$. This implies that

$$\begin{aligned} |hA| &= |S| + \sum_{w \in W} |S(w) \cap hA| \\ &\geq hk - (h-1) + |W|(h-1) \\ &= hk - (h-1) + r(h-1), \end{aligned}$$

which is the conclusion of the Theorem.

To show that $|S(w) \cap hA| \geq h-1$, we choose $w \in W = [1, a_{k-1}] \backslash A$. There exists a unique $t \in [0, k-2]$ such that $a_t < w < a_{t+1}$. Define the sets I, Y, and Z by

$$\begin{aligned} I &= [w+1, a_{k-1} + w - 1] \\ Y &= I \cap S \\ &= \{a_{t+1}, a_{t+2}, \ldots, a_{k-1}, a_1 + a_{k-1}, \ldots, a_t + a_{k-1}\} \\ Z &= \{w + a_{k-1} - a_j | j = 1, 2, \ldots, k-2\}. \end{aligned}$$

Then $Y \subseteq I$ and $Z \subseteq I$. Since

$$|I| = a_{k-1} - 1 \leq 2k - 4 < 2k - 3 = (k-1) + (k-2) = |Y| + |Z|,$$

it follows that $Y \cap Z \neq \emptyset$. Therefore, there exist $i \in [1, k-1]$ and $j \in [1, k-2]$ such that either

$$a_i = w + a_{k-1} - a_j$$

or

$$a_i + a_{k-1} = w + a_{k-1} - a_j.$$

In the first case, we have

$$w + a_{k-1} = a_i + a_j,$$

and so

$$w + (r+1)a_{k-1} = a_i + a_j + ra_{k-1} \in (r+2)A \subseteq hA$$

for $r = 0, 1, 2, \ldots, h-2$. In the second case, we have

$$w = a_i + a_j$$

and so

$$w + ra_{k-1} = a_i + a_j + ra_{k-1} \in (r+2)A \subseteq hA$$

for $r = 0, 1, \ldots, h-2$. In both cases, $|S(w) \cap hA| \geq h-1$ for all $w \in W$. This completes the proof.

Corollary 2.4 *Let $k \geq 3$. Let $A = \{a_0, a_1, \ldots, a_{k-1}\}$ be a set of integers such that*

$$0 = a_0 < a_1 < \cdots < a_{k-1} \leq 2k - 3.$$

Define r by $a_{k-1} = k - 1 + r$. Then $0 \leq r \leq k - 2$ and

$$|2A| \geq 2k - 1 + r.$$

Theorem 2.6 *Let $k \geq 3$. Let $A = \{a_0, a_1, \ldots, a_{k-1}\}$ be a set of integers such that*

$$0 = a_0 < a_1 < \cdots < a_{k-1}$$

and

$$d(A) = (a_1, \ldots, a_{k-1}) = 1.$$

If $a_{k-1} \geq 2k - 3$, then

$$|2A| \geq 3k - 3.$$

Proof. If $a_{k-1} = 2k - 3$, then Corollary 4.1 gives the result.
Let $a_{k-1} \geq 2k - 2$. The proof will be by induction on $k = |A|$.
Let $k = 3$. Then $A = \{0, a_1, a_2\}$, and $a_2 \geq 4$. We must show that $|2A| \geq 6$.
Clearly, $|2A| \leq 6$, since

$$2A = \{0, a_1, a_2, 2a_1, a_1 + a_2, 2a_2\},$$

and $|2A| \geq 5$, since

$$0 < a_1 < a_2 < a_1 + a_2 < 2a_2.$$

Observe that

$$a_1 < 2a_1 < a_1 + a_2.$$

It follows that $|2A| = 5$ only if $a_2 = 2a_1$, which implies that

$$a_1 = (a_1, a_2) = 1$$

and $a_2 = 2a_1 = 2$. This contradicts the requirement that $a_2 \geq 4$. Thus, $a_2 \neq 2a_1$, and so $|2A| = 6$.

Let $k \geq 4$. Assume that the Theorem holds for $|A| = 3, 4, \ldots, k - 1$. Define

$$A' = A \setminus \{a_{k-1}\} = \{0, a_1, \ldots, a_{k-2}\}.$$

Then

$$d(A') = (a_1, a_2, \ldots, a_{k-2}).$$

There are several cases.

Case 1. Suppose that $d(A') = d > 1$. The hypothesis $d(A) = 1$ implies that $(a_{k-1}, d) = 1$. By Theorem 2.1,

$$|2A'| \geq 2(k - 1) - 1 = 2k - 3.$$

Also, d divides n for all $n \in 2A'$. Define

$$B = \{a_{k-1} + a_i | i = 0, 1, \ldots, k-2\}.$$

Then $|B| = k-1$. Since $(n, d) = 1$ for all $n \in B$, it follows that $B \cap 2A' = \emptyset$. Moreover, $n < 2a_{k-1}$ for all $n \in B \cup 2A'$. Since

$$2A' \cup B \cup \{2a_{k-1}\} \subseteq 2A,$$

we conclude that

$$|2A| \geq |2A'| + |B| + 1 \geq (2k-3) + (k-1) + 1 = 3k-3,$$

which is what we had to prove.

In the remaining cases we can assume that $d(A') = 1$.

Case 2. Suppose that $a_{k-2} < 2(k-2) = 2k-4$.

Case 2a. Suppose that $a_i < 2i$ for $i = 1, 2, \ldots, k-2$. Then $0 < a_1 < 2$ implies that $a_1 = 1$. Define

$$\mathcal{C} = [1, 2k-4] \setminus A'.$$

Then

$$|\mathcal{C}| = (2k-4) - (k-2) = k-2.$$

Let $c \in \mathcal{C}$. Then $c > a_1 = 1$. There exists a unique $t \in [1, k-2]$ such that

$$a_t < c < a_{t+1}.$$

Consider the sets

$$\mathcal{D}_1 = \{a_i | i \in [1, t]\}$$

and

$$\mathcal{D}_2 = \{c - a_j | j \in [1, t]\}.$$

Then $|\mathcal{D}_1| = |\mathcal{D}_2| = t$ and

$$\mathcal{D}_1 \cup \mathcal{D}_2 \subseteq [1, c-1].$$

Suppose that $t < k-2$. Since

$$c < a_{t+1} < 2(t+1) = 2t+2,$$

it follows that $c \leq 2t$ and so $c - 1 < 2t$.

Suppose that $t = k-2$. Then the condition $c \leq 2k-4$ implies that

$$c - 1 \leq 2k-5 < 2k-4 = 2t.$$

In both cases we conclude that

$$\mathcal{D}_1 \cap \mathcal{D}_2 \neq \emptyset$$

and so

$$a_i = c - a_j$$

for some $i, j \in [1, t] \subseteq [1, k-2]$. Thus, $c = a_i + a_j \in 2A'$.

If $n \in A' \cup \mathcal{C} = [0, 2k-4]$, then $n \leq 2k-4 < 2k-2 \leq a_{k-1}$, and so the sets $A', \mathcal{C}$, and

$$\{a_{k-1}\} + A = \{a_{k-1} + a_i | i \in [0, k-1]\}$$

are pairwise disjoint. Since

$$A' \cup \mathcal{C} \cup (\{a_{k-1}\} + A) \subseteq 2A,$$

it follows that

$$|2A| \geq (k-1) + (k-2) + k = 3k-3.$$

Case 2b. Suppose that $a_{i-1} \geq 2(i-1)$ for some $i \in [2, k-2]$. Choose $s \in [2, k-2]$ such that $a_j < 2j$ for $j = s, s+1, \ldots, k-2$, but $a_{s-1} \geq 2(s-1)$. Then

$$2s-2 \leq a_{s-1} < a_s < 2s,$$

and so $a_s = 2s-1$ and $a_{s-1} = 2s-2$. Define sets A_1 and A_2 by

$$A_1 = \{a_0, a_1, \ldots, a_{s-1}, a_s\}$$

and

$$A_2 = \{a_{s-1}, a_s, a_{s+1}, \ldots, a_{k-2}, a_{k-1}\}.$$

Since $a_s - a_{s-1} = 1$, it follows that $d(A_1) = d(A_2) = 1$. Moreover,

$$3 \leq k_1 = |A_1| = s+1 \leq k-1$$

since $s \in [2, k-2]$, and

$$a_{k_1-1} = a_s = 2s-1 = 2k_1 - 3 = (k_1-1) + (k_1-2).$$

It follows from Corollary 4.1 that

$$|2A_1| \geq (2k_1-1) + (k_1-2) = 3k_1 - 3 = 3s.$$

Define the set A_2^* by

$$A_2^* = A_2 - \{a_{s-1}\} = \{0, 1, a_{s+1} - a_{s-1}, \ldots, a_{k-1} - a_{s-1}\}.$$

Let $k_2 = |A_2| = |A_2^*|$. Then

$$3 \leq k_2 = (k-1) - (s-1) + 1 = k - s + 1 \leq k-1$$

since $s \in [2, k-2]$. Moreover, the integer $a_{k-1} - a_{s-1}$, which is the largest element of A_2^*, satisfies the inequality

$$a_{k-1} - a_{s-1} \geq (2k-2) - (2s-2) = 2(k-s) = 2k_2 - 2.$$

Since $k_2 \leq k-1$, it follows from the induction hypothesis that

$$|2A_2| = |2A_2^*| \geq 3k_2 - 3 = 3k - 3s.$$

Since

$$2A_1 \cup 2A_2 \subseteq 2A$$

and

$$2A_1 \cap 2A_2 = \{2a_{s-1}, a_{s-1} + a_s, 2a_s\},$$

it follows that

$$|2A| \geq |2A_1| + |2A_2| - 3 \geq 3s + (3k - 3s) - 3 = 3k - 3.$$

Case 3. Suppose that $a_{k-2} \geq 2(k-2) = 2k-4$. It follows from the induction hypothesis that $|2A'| \geq 3(k-1)-3 = 3k-6$. To prove the Theorem, it suffices to show that $|2A \backslash 2A'| \geq 3$. Since $n \leq 2a_{k-2}$ for $n \in 2A'$, we see immediately that

$$\{a_{k-1} + a_{k-2}, 2a_{k-1}\} \subseteq 2A \setminus 2A'.$$

Case 3a. Suppose that $m = a_{k-1} - a_{k-2} \geq 2$. Then

$$a_{k-1} \equiv a_i \pmod{m}$$

for $i = k-2, k-1$. If

$$a_{k-1} \equiv a_i \pmod{m}$$

for $i = 0, 1, 2, \ldots, k-1$, then m divides each a_i and so

$$(a_1, a_2, \ldots, a_{k-1}) \geq m > 1,$$

which is false. Therefore, there exists $t \leq k-3$ such that

$$a_{k-1} \equiv a_i \pmod{m}$$

for $i = t+1, t+2, \ldots, k-1$, but m does not divide $a_{k-1} - a_t$.

We shall show that $a_{k-1} + a_t \in 2A \backslash 2A'$. If $a_{k-1} + a_t \in 2A'$, then there exist $r, s \in [t+1, k-2]$ such that $a_{k-1} + a_t = a_r + a_s$. This implies that

$$a_{k-1} - a_t = (a_{k-1} - a_r) + (a_{k-1} - a_s) \equiv 0 \pmod{m},$$

which contradicts our hypothesis. Therefore,

$$\{a_{k-1} + a_t, a_{k-1} + a_{k-2}, 2a_{k-1}\} \subseteq 2A \setminus 2A',$$

and so

$$|2A| = |2A'| + |2A \setminus 2A'| \geq (3k-6) + 3 = 3k-3.$$

Case 3b. Suppose that $a_{k-1} - a_{k-2} = 1$ and $a_{k-2} - a_{k-3} \geq 2$.Consider the integer $a_{k-1} + a_{k-3} \in 2A$. Observe that

$$a_{k-2} + a_{k-3} < a_{k-1} + a_{k-3}$$

and

$$a_{k-2} + a_{k-3} < 2a_{k-2}.$$

It follows that $a_{k-1} + a_{k-3} \in 2A'$ if and only if $a_{k-1} + a_{k-3} = 2a_{k-2}$, that is, if and only if

$$1 = a_{k-1} - a_{k-2} = a_{k-2} - a_{k-3},$$

which is false. Therefore,

$$\{a_{k-1} + a_{k-3}, a_{k-1} + a_{k-2}, 2a_{k-1}\} \subseteq 2A \setminus 2A',$$

and so

$$|2A| \geq 3k-3.$$

Case 3c. Suppose that $a_{k-1} - a_{k-2} = a_{k-2} - a_{k-3} = 1$. Then $a_{k-1} - a_{k-3} = 2$. Consider the set

$$\begin{aligned} A^* &= \{a_{k-1} - a_i | i \in [0, k-1]\} \\ &= \{0, 1, 2, a_{k-1} - a_{k-4}, \ldots, a_{k-1} - a_1, a_{k-1}\}. \end{aligned}$$

Then $2A^* = \{2a_{k-1} - n | n \in 2A\}$ and $|2A^*| = |2A|$. It follows from the preceding cases that if $|2A^*| < 3k-3$, then $a_1 = 1$ and $a_2 = 2$. Therefore, we can assume that

$$A = \{a_0, a_1, a_2, a_3, \ldots, a_{k-3}, a_{k-2}, a_{k-1}\},$$

where $a_0 = 0, a_1 = 1, a_2 = 2, a_{k-3} = a_{k-1} - 2$, and $a_{k-2} = a_{k-1} - 1$. Choose the smallest $s \geq 1$ such that $a_s \geq 2s$. Then $a_i < 2i$ for $i = 1, 2, \ldots, s-1$, and

$$3 \leq s \leq k-3.$$

Define sets A_1 and A_2 by

$$A_1 = \{a_0, a_1, \ldots, a_{s-1}, a_s\}$$

and

$$A_2 = \{a_{s-1}, a_s, \ldots, a_{k-2}, a_{k-1}\}.$$

Then $d(A_1) = 1$. Moreover,

$$4 \leq k_1 = |A_1| = s + 1 \leq k-2$$

since $s \in [3, k-3]$, and

$$a_{k_1-1} = a_s \geq 2s = 2k_1 - 2.$$

The induction hypothesis implies that

$$|2A_1| \geq 3k_1 - 3 = 3s.$$

Define the set A_2^* by

$$A_2^* = A_2 - \{a_{s-1}\} = \{0, a_s - a_{s-1}, \ldots, a_{k-2} - a_{s-1}, a_{k-1} - a_{s-1}\}.$$

Then $d(A_2^*) = 1$. Also, $s \in [3, k-3]$ implies that

$$4 \leq k_2 = |A_2| = |A_2^*| = k - s + 1 \leq k - 2$$

and

$$a_{k-1} - a_{s-1} > (2k-2) - (2s-2) = 2(k-s) = 2k_2 - 2.$$

Again, the induction hypothesis implies that

$$|2A_2| = |2A_2^*| \geq 3k_2 - 3 = 3k - 3s.$$

Since

$$2A_1 \cup 2A_2 \subseteq 2A$$

and

$$2A_1 \cap 2A_2 = \{2a_{s-1}, a_{s-1} + a_s, 2a_s\},$$

it follows that

$$|2A| \geq |2A_1| + |2A_2| - 3 \geq 3s + (3k - 3s) - 3 = 3k - 3.$$

This completes the proof of the Theorem.

Theorem 2.7 *Let $k \geq 3$, and let A be a set of k integers such that*

$$|2A| \leq 3k - 4.$$

Define the integer r by $|2A| = 2k-1+r$. Then $r \leq k-3$, and A is contained in an arithmetic progression of length $k + r \leq 2k - 3$.

Proof. Let $A^N = \{a_0, a_1, \ldots, a_{k-1}\}$, where $0 = a_0 < a_1 < \cdots < a_{k-1}$ and $d(A^N) = 1$. Since

$$|2A^N| = |2A| \leq 3k - 4,$$

it follows from Theorem 4.2 that $a_{k-1} \leq 2k - 3$. Define the integer r' by

$$a_{k-1} = k - 1 + r'.$$

Then $0 \leq r' \leq k - 2$, and Corollary 4.1 implies that

$$2k - 1 + r' \leq |2A^N| = 2k - 1 + r.$$

Therefore, $r' \leq r$ and $a_{k-1} \leq k - 1 + r$. It follows that

$$A^N \subseteq [0, k - 1 + r],$$

and so A is a subset of an arithmetic progression with k+r terms. This completes the proof.

Corollary 2.5 *Let A be a set of k nonnegative integers such that $0 \in A$ and $d(A) = 1$. Let $n(A)$ denote the largest element of A. If $n(A) \geq 2k-3$, then $|2A| \geq 3k-3$.*

Proof. Suppose that $|2A| \leq 3k-4$. Let $|2A| = 2k-1+r$. Then $0 \leq r \leq k-3$, and Theorem 2.7 implies that A is a subset of an arithmetic progression of length $k+r$. Let m be the difference between consecutive terms of this progression. Since $0 \in A$, it follows that

$$A \subseteq \{0, m, 2m, \ldots, (k+r-1)m\}.$$

Since the greatest common divisor $d(A)$ of the elements of A is 1, it follows that $m = 1$, and so

$$A \subseteq [1, k+r-1] \subseteq [1, 2k-4],$$

which contradicts the fact that the greatest element of A is $n(A) \geq 2k-3$. Therefore, $|2A| \geq 3k-3$.

Corollary 2.6 *Let A be a set of k nonnegative integers with $0 \in A$. Let $n(A)$ denote the largest element of A. If $n(A) \leq 2k-3$, then $|2A| \geq k + n(A)$.*

Proof. If $d(A) \geq 2$, then $n(A) \geq 2(k-1) > 2k-3$, which is false. Therefore, $d(A) = 1$. If $|2A| \geq 3k-3$, then

$$|2A| \geq 3k-3 = k + (2k-3) \geq k + n(A),$$

and the result holds. If $|2A| \leq 3k-4$, then $|2A| = 2k-1+r$, where $0 \leq r \leq k-3$, and Theorem 2.7 implies that $A \subseteq [0, k-1+r]$. This means that $n(A) \leq k-1+r$, and so

$$|2A| = 2k-1+r = k + (k-1+r) \geq k + n(A).$$

This completes the proof.

2.4 Application: Sumsets and powers of 2

Let $n \geq 1$, and let B be the set of all multiples of 3 contained in the interval $[1, n]$. Then $|B| \leq n/3$, and every sum of elements of B is divisible by 3. This implies that no such sum is a power of 2. We shall prove that the set of $[n/3]$ multiples of 3 is the extreme case: If $B \subseteq [1, n]$ and $|B| > n/3$, then some power of 2 can be written as the sum of a bounded number of elements of B.

Theorem 2.8 *Let $n \geq 1$, and let B be a set of integers contained in the interval $[1, n]$. If $|B| > n/3$, then there is a power of 2 that can be written as the sum of at most 6 elements of B.*

Proof. Let $n(B)$ denote the alrgest element of B. Since $B \subseteq [1,n]$ implies that $n \geq n(B)$ and since $|B| > n/3 \geq n(B)/3$, we can assume that $n(B) = n$. Let d be the greatest common divisor of the elements of B. If $d \geq 3$, then $n \geq d|B| \geq 3|B| > n$, which is impossible. If $d = 2$, let $B' = \{b/2 \mid b \in B\} \subseteq [1, n/2]$. Then the greatest common divisor of the elements of B' is 1. Moreover, $n(B') = n/2$ and $|B'| = |B| > n/3 > n(B')/3$. Thus, the set B' satisfies the conditions of the Theorem. If the Theorem holds in the case $d = 1$, then there exist integers $b_1, \ldots, b_h \in B$ such that $2^s = b_1/2 + \cdots + b_h/2 \in hB'$, and so $2^{s+1} = b_1 + \cdots + b_h \in hB$. Therefore, we can assume that $d = 1$.

Let $A = \{0\} \cup B$. Then $d(A) = 1$ and $n(A) = n(B) = n$. Let

$$k = |A| = |B| + 1 > n/3 + 1.$$

If $k \leq (n+3)/2$, then $n \geq 2k-3$. Since $d(A) = 1$, Corollary 2.5 implies that $|2A| \geq 3k-3 \geq n+1$. If $k > (n+3)/2$, then $|2A| \geq 2k - 1 > n + 2$. In both cases,

$$|2A| \geq n + 1.$$

Observe that $4A = 2A+2A = 2(2A)$, and that $n(2A) = 2n$ and $d(2A) = 1$. If $2n \geq 2|2A|-3$, then Corollary 2.5 implies that $|4A| \geq 3|2A| - 3 \geq 3n$. If $2n < 2|2A| - 3$, then Corollary 2.6 implies that

$$|4A| \geq |2A| + n(2A) = |2A| + 2n \geq 3n + 1.$$

In both cases,

$$|4A| \geq 3n.$$

If n is a power of 2, we are done, since $n \in B$. If n is not a power of 2, then neither $2n$ nor $4n$ is a power of 2, and there exists a unique integer s such that

$$2n < 2^s < 4n.$$

If $2^{s-1} \in 2^s - 2A$, then there exist $a_1, a_2 \in A$ such that

$$2^{s-1} = 2^s - (a_1 + a_2),$$

and so

$$2^{s-1} = 2^s - 2^{s-1} = a_1 + a_2 \in 2A = B \cup 2B,$$

and we are done. Therefore, we can assume that $2^{s-1} \notin 2^s - 2A$, and so

$$2^s - 2A \subseteq [2^s - 2n, 2^s] \setminus \{2^{s-1}\} \subseteq [0, 4n] \setminus \{2^{s-1}\}.$$

Similarly, if $2^{s-1} \in 4A$, then 2^{s-1} is the sum of at most 4 elements of B. Thus, we can also assume that

$$4A \subseteq [0, 4n] \setminus \{2^{s-1}\}.$$

Since

$$|2^s - 2A| = |2A| \geq n + 1$$

and

$$|4A| \geq 3n$$

and

$$|[0, 4n] \setminus \{2^{s-1}\}| = 4n,$$

it follows that

$$(2^s - 2A) \cap 4A \neq \emptyset,$$

and so $2^s \in 6A = 6(B \cup \{0\}$, that is, 2^s can be written as the sum of at most 6 elements of B. This completes the proof.

2.5 Notes

The principal result in this chapter is Theorem 2.7, which was first proved by Freiman [26] and which also appears in Chapter 1 of his book [29]. Freiman [31] also proved Theorem 2.8 on the represention of powers of 2 as the sum of a small number of integers in certain subsets of $[1, n]$. A weaker result on representing powers of 2 as the sum of a bounded number of terms had been obtained earlier by Nathanson and Sárkőzy [56]. Erdős and Freud conjectured that if $A \subseteq [1, n]$ and $|A| > n/3$, then some power of 2 can be written as the sum of distinct elements of A. This was proved by Erdős and Freiman [23], Nathanson and Sárkőzy [56], and Freiman [31]. An early version of this chapter appeared in Nathanson [53].

2.6 Exercises

1. Determine the structure of all sets A such that $|A| = k$ and $|2A| = 2k$.

Chapter 3

Sums of sets of congruence classes

3.1 Addition in groups

Let G be an abelian group, and let A and B be subsets of G. The sumset $A+B$ is the set of all elements of G that can be written in the form $a+b$, where $a \in A$ and $b \in B$. The direct problem for addition in groups is to find a lower bound for $|A+B|$ in terms of $|A|$ and $|B|$. The inverse problem is to determine the structure of the pairs A, B of sets for which $|A+B|$ is minimal. The following simple theorem shows that the result is trivial if G is finite and $|A|+|B| > |G|$.

Theorem 3.1 *Let G be a finite abelian group, and let A and B be subsets of G such that $|A|+|B| > |G|$. Then $A+B = G$.*

Proof. Let $g \in G$, and let $g - B = \{g - b \mid b \in B\}$. Then

$$|A| + |g - B| = |A| + |B| > |G|$$

implies that $(g-B) \cap A \neq \emptyset$, and so there exist $a, b \in G$ such that $a = g - b$. Thus, $g \in A+B$ for all $g \in G$. This completes the proof.

It follows from Theorem 3.1 that it is enough to examine only the pairs A, B for which $|A|+|B| \leq |G|$.

The inverse problem for addition in groups is to describe the structure of the "critical pairs" of subsets A, B for which the cardinality of the sumset $A+B$ is as small as possible. We shall solve both the direct and the inverse problems for cyclic groups $\mathbf{Z}/p\mathbf{Z}$, where p is a prime number. In this case, Cauchy and Davenport proved that $|A+B| \geq \min(p, |A|+|B|-1)$, and Vosper completely classified the pairs A, B for which $|A+B| = |A|+|B|-1$.

We shall also prove an important theorem of Kneser on the size of sumsets in arbitrary abelian groups.

3.2 The Cauchy-Davenport-Chowla theorem

In this chapter we shall study sumsets in the group of congruence classes modulo a prime p. The basic result, known as the Cauchy-Davenport theorem, is a lower bound on the cardinality of the sum of two sets of congruence classes mod p.

Lemma 3.1 *Let A and B be finite subsets of the abelian group G, and let e be any element of G. Let*

$$\begin{aligned} A(e) &= A \cup (B + e) \\ B(e) &= B \cap (A - e). \end{aligned}$$

Then

$$A(e) + B(e) \subset A + B \tag{3.1}$$

and

$$|A(e)| + |B(e)| = |A| + |B|. \tag{3.2}$$

If $e \in A$ and $0 \in B$, then $e \in A(e)$ and $0 \in B(e)$.

Remark. The pair of sets $A(e), B(e)$ defined in the Lemma is called the *e-transform* of the sets A, B.

Proof. The set inclusion (3.1) follows immediately from the definition of the e-transform.

Since $A \subseteq A(e)$ and $B(e) \subseteq B$, to prove equation (3.2) it suffices to show that

$$|A(e) \setminus A| = |A(e)| - |A| = |B| - |B(e)| = |B \setminus B(e)|.$$

Observe that

$$\begin{aligned} A(e) \setminus A &= (B + e) \setminus A \\ &= \{g \in G \mid g - e \in B, g \notin A\} \\ &= e + \{b \in B \mid b + e \notin A\} \\ &= e + \{b \in B \mid b \notin A - e\} \\ &= e + \{b \in B \mid b \notin B(e)\} \\ &= e + B \setminus B(e). \end{aligned}$$

Therefore,

$$A(e) \setminus A = e + (B \setminus B(e))$$

and so

$$|A(e) \setminus A| = |e + (B \setminus B(e)|) = |B \setminus B(e)|.$$

If $e \in A$, then $e \in A(e)$ and $0 \in A - e$. If $0 \in B$ and $e \in A$, then $0 \in B \cap A - e = B(e)$. This completes the proof.

Theorem 3.2 (I. Chowla) *Let $m \geq 2$, and let A and B be subsets of $\mathbf{Z}/m\mathbf{Z}$. If $0 \in B$ and if $(b, m) = 1$ for all $b \in B \setminus \{0\}$, then*

$$|A + B| \geq \min(m, |A| + |B| - 1).$$

Proof. The Theorem is true if $|A| + |B| > m$. The Theorem is also true if $|A| = 1$ or $|B| = 1$, since $|A + B| = |A| + |B| - 1$ in this case. Suppose that the Theorem is false. Then there exists a pair of sets $A, B \subseteq \mathbf{Z}/m\mathbf{Z}$ such that $|A| > 1$, $|B| > 1$, and

$$|A + B| < |A| + |B| - 1 < m.$$

In particular, $A \neq G$. Choose the pair A, B so that the cardinality of B is minimal. Since $|B| > 1$, there exists $b \in B, b \neq 0$. If $a + b \in A$ for all $a \in A$, then $a + ib \in A$ for all $i = 0, 1, 2, \ldots$, which is impossible, since $(b, m) = 1$ implies that $\{ib \mid i = 0, 1, 2, \ldots\} = G$, and so $A = G$. Therefore, there exists an element $e \in A$ such that $e + b \notin A$. Apply the e-transform to the pair A, B. Since $A(e) + B(e) \subseteq A + B$, it follows that

$$|A(e) + B(e)| \leq |A + B| < |A| + |B| - 1 = |A(e)| + |B(e)| - 1 < m.$$

Since $e \in A$ and $0 \in B$, it follows that $0 \in B(e) \subseteq B$. Since $e + b \notin A$, it follows that $b \notin A - e$, and so $b \notin B \cap (A - e) = B(e)$. Therefore, $B(e)$ is a proper subset of B, and $|B(e)| < |B|$, which contradicts the minimality of $|B|$. This completes the proof.

Theorem 3.3 (Cauchy-Davenport) *Let p be a prime, and let A and B be nonempty subsets of $\mathbf{Z}/p\mathbf{Z}$. Then*

$$|A + B| \geq \min(p, |A| + |B| - 1).$$

Proof. Let $b_0 \in B$, and let $B' = B - b_0$. Then $|B'| = |B|$ and $|A + B'| = |A + B - b_0| = |A + B|$. Since $0 \in B'$ and since $(b, p) = 1$ for all $b \in B', b \neq 0$, it follows from Theorem 3.2 applied to the sets A, B' that

$$|A + B| = |A + B'| \geq \min(p, |A| + |B'| - 1) = \min(p, |A| + |B| - 1).$$

This completes the proof.

Corollary 3.1 *Let $h \geq 2$. Let p be a prime, and let $A_1, A_2, \ldots, A_h$ be nonempty subsets of $\mathbf{Z}/p\mathbf{Z}$. Then*

$$|A_1 + A_2 + \cdots + A_h| \geq \min(p, \sum_{i=1}^{h} |A_i| - (h - 1)).$$

Proof. By induction on h. The case $h = 2$ is the Cauchy-Davenport theorem. Let $h \geq 3$, and suppose that the result holds for any $h - 1$ subsets of $\mathbf{Z}/p\mathbf{Z}$. Let $A_1, A_2, \ldots, A_h$ be nonempty subsets of $\mathbf{Z}/p\mathbf{Z}$, and let $B = A_1 + \cdots + A_{h-1}$. Then

$$|B| = |A_1 + \cdots + A_{h-1}| \geq \min(p, \sum_{i=1}^{h-1} |A_i| - (h - 2))$$

and so

$$\begin{aligned}
|A_1 + A_2 + \cdots + A_h| &= |(A_1 + \cdots + A_{h-1}) + A_h| \\
&= |B + A_h| \\
&\geq \min(p, |B| + |A_h| - 1) \\
&\geq \min\left(p, \left(\min\left(p, \sum_{i=1}^{h-1} |A_i| - (h-2)\right)\right) + |A_h| - 1\right) \\
&\geq \min\left(p, \sum_{i=1}^{h} |A_i| - (h-1)\right).
\end{aligned}$$

This completes the proof.

3.3 Vosper's inverse theorem

An *arithmetic progression* in the group G is a set of the form $\{a + id \mid i = 0, 1, \ldots, k-1\}$, where d is called the *common difference* d of the progression and k is called the *length* of the progression. The order of the group element d in G must be at least k.

Let p be a prime number, and let A and B be nonempty subsets of $G = \mathbf{Z}/p\mathbf{Z}$. If $|A| + |B| > p$, then $A + B = G$. If $|A| + |B| \leq p$, then the Cauchy-Davenport theorem states that $|A + B| \geq |A| + |B| - 1$. The pair of sets A, B is called a *critical pair* if $|A| + |B| \leq p$ and $|A + B| = |A| + |B| - 1$. Vosper completely characterized the structure of all critical pairs of subsets of the group of congruence classes modulo a prime. He proved that, with two simple exceptions, the sets A and B form a critical pair if and only if they are arithmetic progressions with the same common difference.

We denote the complement of the set S in G by $\overline{S}$.

Theorem 3.4 (Vosper) *Let p be prime, and let A and B be subsets of the group $G = \mathbf{Z}/p\mathbf{Z}$ such that $|A| + |B| \leq p$. Then*

$$|A + B| = |A| + |B| - 1$$

if and only if one of the follows three conditions holds:

(i) $\min(|A|, |B|) = 1$

(ii) $1 \leq |A| \leq p - 1$ *and* $B = \overline{c - A}$ *for some* $c \in G$

(iii) *A and B are arithmetic progressions with the same common difference.*

Proof. Let $|A| = k$ and $|B| = l$. If $\min(k, l) = 1$, say, $k = 1$, then $|A + B| = |B| = l = k + l - 1$, and so A, B is a critical pair.

Let $c \in G$, and let A be any subset of G such that $1 \leq |A| = k \leq p - 1$. Let $B = \overline{c - A}$. Then $c \notin A + B$, and so $|A + B| \leq p - 1$. Since $l = |B| = p - |c - A| = p - k$, the Cauchy-Davenport theorem implies that

$$p - 1 = k + l - 1 \leq |A + B| \leq p - 1,$$

and so $|A+B| = p-1 = k+l-1$. Thus, A, B is a critical pair.

If A and B are arithmetic progressions with the same common difference d, then

$$A = \{a + id \mid i = 0, 1, \ldots, k-1\}$$

and

$$B = \{b + id \mid \ = 0, 1, \ldots, k-1\}.$$

Therefore,

$$A + B = \{a + b + id \mid \ = 0, 1, \ldots, k+l-2\},$$

and so $|A+B| \leq k+l-1$. By the Cauchy-Davenport theorem, $|A+B| \geq k+l-1$, and so $|A+B| = k+l-1$. Thus, A and B form a critical pair.

Conversely, let A, B be a critical pair. If $|A| = 1$ or $|B| = 1$, the pair is of the form (i).

If $A+B| = p-1$, then $\overline{A+B} = \{c\}$ for some $c \in G$. Since $c \notin A+B$, it follows that

$$B \subseteq \overline{c-A},$$

and so

$$l = |B| \leq |\overline{c-A}| = p - |c-A| = p-k.$$

Since

$$p-1 = |A+B| = k+l-1 \leq p-1,$$

it follows that $l = p-k$ and so $B = \overline{c-A}$. Thus, the pair A, B is of the form (ii).

For the remainder of the proof, we can assume that A, B is a critical pair such that $\min(|A|, |B|) > 1$ and $|A+B| < p-1$. We shall show that A and B are arithmetic progressions with the same common difference. This will require several lemmas.

Lemma 3.2 *Let If A is an arithmetic progression, then B is an arithmetic progression with the same common difference.*

Lemma 3.3 *If $\min(|A|, |B|) = 2$, then A and B are arithmetic progressions with the same common difference.*

Proof. This follows immediately from the Lemma 3.2, since every set with two elements is an arithmetic progression.

Lemma 3.4 *Let $D = \overline{A+B}$. Then the sets $-A, D$ form a critical pair.*

Lemma 3.5 *If $A+B$ is an arithmetic progression, then A and B are arithmetic progressions with the same common difference.*

Proof. If $A+B$ is an arithmetic progression, then $D = \overline{A+B}$ is an arithmetic progression. By Lemma 3.4, the pair $-A, D$ is critical, and so, by Lemma 3.2, the set $-A$ is an arithmetic progression. It follows that A is an arithmetic progression, and, since the pair A, B is critical, the sets A and B are arithmetic progressions with the same common difference.

Lemma 3.6 *If $0 \in B$ and $|B| = l \geq 3$, then there exists $e \in A$ such that $A(e), B(e)$ is a critical pair and $2 \leq |B(e)| < |B|$, where $A(e)$ and $B(e)$ are the e-transforms of A and B.*

Proof of Vosper's theorem, concluded. Let A, B be a critical pair with $|B| = l \geq 2$. The proof is by induction on l. If $l = 2$, the result follows from Lemma 3.3. Let $l \geq 3$, and assume that the theorem holds for all critical pairs A, B with $|B| < l$. By Lemma 3.6, there exists $e \in A$ such that $A(e), B(e)$ is a critical pair with $2 \leq |B(e)| < l$. It follows by the induction assumption that $A(e)$ and $B(e)$ are arithmetic progressions with the same common difference. Therefore, $A(e) + B(e)$ is an arithmetic progression. By Lemma 3.1,

$$|A(e)| + |B(e)| = |A| + |B| = k + l$$

and

$$A(e) + B(e) \subseteq A + B.$$

By the Cauchy-Davenport theorem,

$$k + l - 1 = |A(e)| + |B(e)| + l \leq |A(e) + B(e)| \leq |A + B| = k + l - 1,$$

and so $|A(e)+B(e)| = k+l-1$. Therefore, $A+B = A(e)+B(e)$ is an arithmetic progression, and Lemma 3.5 implies that A and B are arithmetic progressions with the same common difference.

3.4 Kneser's theorem for finite abelian groups

For any subset nonempty S of the abelian group G, let

$$H(S) = \{g \in G \mid g + S = S\}.$$

Then $H(S)$ is a subgroup of G. Clearly, $0 \in H(S)$ for every set S, and $H(S) = G$ if and only if $S = G$. The set S is called *periodic* if $S \neq G$ and $H(S) \neq \{0\}$.

Theorem 3.5 *Let G be an abelian group, and let A and B be finite, nonempty subsets of G. Let $H = H(A + B)$. Then either*

$$|A + B| \geq |A| + |B|$$

or

$$|A + B| = |A + H| + |B + H| - |H| \geq |A| + |B| - |H|.$$

Corollary 3.2 *Let G be an abelian group, and let A and B be finite, nonempty subsets of G. Let $H = H(A + B)$. For any set $S \subseteq G$, let $\overline{S}$ denote the image of S in the quotient group G/H. Then either*

$$|A + B| \geq |A| + |B|$$

or

$$|\overline{A + B}| = |\overline{A}| + |\overline{B}| - 1.$$

3.5 Notes

The e-transform was first used by Dyson [20] and van der Corput [13]. The Cauchy-Davenport theorem was proved by Cauchy [8] in 1813. Davenport [15] rediscovered the result in 1935, and discovered in 1947 that Cauchy had proved it first (see Davenport [17]). I. Chowla [9] published his generalization of the Cauchy-Davenport theorem in 1935. All of these results can be found in Mann [46] and Halberstam and Roth [33].

3.6 Exercises

1. Extend I. Chowla's theorem 3.2 to sums of $h \geq 3$ subsets of $\mathbf{Z}/m\mathbf{Z}$.

2. Prove that Kneser's theorem 3.5 implies the Cauchy-Davenport theorem.

3. Prove that Kneser's theorem 3.5 implies the theorem of I. Chowla.

4. Let p be a prime, and let $(a,p) = (b,p) = 1$. Let $f(x,y) = ax^2 + by^2$. Use the Cauchy-Davenport Theorem to prove that $f(x,y) \equiv n \pmod{p}$ is solvable for all n.

5. Let p be a prime, and let $k \geq 3$. Let $a_1, a_2, \ldots, a_k$ be non-zero elements of $G = \mathbf{Z}/p\mathbf{Z}$. Let
$$f(x_1, x_2, \ldots, x_k) = a_1x_1^k + a_2x_2^k + \cdots + a_kx_k^k.$$
Prove that the congruence $f(x_1, x_2, \ldots, x_k) \equiv n \pmod{p}$ is solvable for all n.

Chapter 4

The asymptotic density of sumsets

4.1 Asymptotic density

Let A be a set of integers, and let $A(n)$ denote the number of elements $a \in A$ such that $1 \leq a \leq n$. The function $A(n)$ is called the *counting function* of A. Note that the set A can contain 0 as well as negative integers, but that these numbers are ignored by the counting function.

The counting function is used to define three different asymptotic densities. Let A be a set of integers. The *lower asymptotic density* of A, denoted $d_L(A)$, is defined by

$$d_L(A) = \liminf_{n\to\infty} A(n)/n.$$

The *upper asymptotic density* of A, denoted $d_U(A)$, is defined by

$$d_U(A) = \limsup_{n\to\infty} A(n)/n.$$

If the limit

$$d(A) = \lim_{n\to\infty} A(n)/n,$$

then $d(A)$ is called the *asymptotic density* of A. In this chapter we consider only $d_L(A)$ and $d(A)$. The sets A and B are *asymptotically equivalent*, denoted $A \sim B$, if A and B coincide for all sufficiently large numbers, that is, there exists N_0 such that, for all $n \geq N_0$, the integer $n \in A$ if and only if $n \in B$. If $A \sim B$, then $d_L(A) = d_L(B)$.

Let $m \geq 2$, and let k and l be positive integers such that $k + l \leq m$. Let

$$A = \{a \in \mathbf{Z} \mid a \equiv r \pmod{m} \text{ for } r = 0, 1, \ldots, k-1\}$$

and

$$B = \{b \in \mathbf{Z} \mid b \equiv s \pmod{m} \text{ for } s = 0, 1, \ldots, l-1\}.$$

Then

$$A + B = \{c \in \mathbf{Z} \mid c \equiv t \pmod{m} \text{ for } t = 0, 1, \ldots, k+l-2\}$$

Since $d_L(A) = k/m, d_L(B) = l/m$, and $d_L(A+B) = (k+l-1)/m$, it follows that

$$d_L(A+B) = d_L(A) + d_L(B) - 1/m < d_L(A) + d_L(B).$$

Kneser discovered the remarkable theorem that if A and B are sets of integers, then $d_L(A+B) \geq d_L(A)+d_L(B)$ except when A and B are "essentially" unions of congruence classes with respect to the same modulus. Kneser's theorem is the first and perhaps the most profound inverse theorem in additive number theory. It can be stated as follows: If A and B are sets of integers such that

$$d_L(A+B) < d_L(A) + d_L(B),$$

then A and B must have a precise structure, namely, there exists a modulus $m \geq 1$ and sets $A^{(m)}$ and $B^{(m)}$ of congruence classes modulo m such that $A \subseteq A^{(m)}$, $B \subseteq B^{(m)}$, and the sumsets $A+B \sim A^{(m)}+B^{(m)}$. Let $G = \mathbf{Z}/m\mathbf{Z}$, and let A' (resp. B') be the set of congruence classes in G whose union is $A^{(m)}$ (resp. $B^{(m)}$. Then

$$\begin{aligned} |A'+B'|/m &= d(A^{(m)}+B^{(m)}) \\ &= d_L(A+B) \\ &< d_L(A) + d_L(B) \\ &\leq d(A^{(m)}) + d(A^{(m)}) \\ &= |A'|/m + |B'|/m \end{aligned}$$

and so

$$|A'+B'| < |A'| + |B'|$$

in the group G. If $m = p$ is prime, and if $|A'| > 1$, $|B'| > 1$, and $|A'| + |B'| < p$, then A' and B' are arithmetic progressions with the same common difference d. This means that there exist congruence classes $a' \pmod p$ and $b' \pmod p$ such that

$$A' = \{a' + id \mid i = 0, 1, \ldots, |A'| - 1\}$$

and

$$B' = \{b' + jd \mid j = 0, 1, \ldots, |B'| - 1\},$$

and so

$$A \subseteq A^{(m)} = \{a' + id + np \mid i = 0, 1, \ldots, |A'| - 1 \text{ and } n \in \mathbf{Z}\}$$

and

$$B \subseteq B^{(m)} = \{b' + jd + np \mid j = 0, 1, \ldots, |B'| - 1 \text{ and } n \in \mathbf{Z}\}.$$

The sets $A^{(m)}$ and $B^{(m)}$ are infinite two-dimensional arithmetic progressions. Finite n-dimensional arithmetic progressions will reappear in Freiman's theorem.

4.2 Kneser's theorem

4.3 Notes

4.4 Exercises

1. Let A be a set of integers. Prove that $d(A)$ exists if and only if $d_L(A) = d_U(A)$.

2. Let A be a set of integers such that $d_L(A) > 1/2$. Prove that A is an asymptotic basis of order 2.

Chapter 5

Applications of Kneser's theorem

5.1 The maximum density of minimal bases

5.2 Sumsets and powers of 2

5.3 On the order of co-finite subsets of bases

5.4 Notes

5.5 Exercises

Chapter 6

Geometry of numbers

6.1 Lattices and determinants

The abelian group generated by a set S of vectors in the n-dimensional Euclidean space $\mathbf{R}^n$ is the set of all finite sums of the form

$$x_1 a_1 + \cdots + x_m a_m,$$

where $a_1, \ldots, a_m \in S$ and $x_1, \ldots, x_m \in \mathbf{Z}$. A *lattice* in $\mathbf{R}^n$ is an abelian group Λ generated by a set of n linearly independent vectors. A set of n generators for a lattice in $\mathbf{R}^n$ is called a *basis* for the lattice. For example, the *standard basis* for $\mathbf{R}^n$ is the set of vectors $e_1, \ldots, e_n$, where

$$\begin{aligned} e_1 &= (1,0,0,0,\ldots,0) \\ e_2 &= (0,1,0,0,\ldots,0) \\ e_3 &= (0,0,1,0,\ldots,0) \\ &\vdots \\ e_n &= (0,0,0,\ldots,0,1). \end{aligned}$$

The integer lattice $\mathbf{Z}^n$ is the lattice with basis $e_1, \ldots, e_n$.

The basis of a lattice is not uniquely determined by the lattice. For example, let Λ be the lattice in $\mathbf{Z}^2$ generated by the vectors $a_1 = (7,5)$ and $a_2 = (4,3)$. Since $a_1 \in \mathbf{Z}^2$ and $a_2 \in \mathbf{Z}^2$, it follows that $\Lambda \subseteq \mathbf{Z}^2$. Conversely, since

$$e_1 = 3a_1 - 5a_2 \in \Lambda$$

and

$$e_2 = -4a_1 + 7a_2 \in \Lambda,$$

it follows that $\mathbf{Z}^2 \subseteq \Lambda$. Thus, $\mathbf{Z}^2 = \Lambda$, and the sets $\{e_1, e_2\}$ and $\{a_1, a_2\}$ are distinct bases for $\mathbf{Z}^2$. Observe that the determinant of the vectors a_1, a_2 is

$$\det(a_1, a_2) = \begin{vmatrix} 7 & 4 \\ 5 & 3 \end{vmatrix} = 1.$$

This example can be generalized. Let $U = (u_{ij})$ be an $n \times n$ matrix with integer entries and determinant $\det(U) = \pm 1$. Then there exists an inverse matrix $V = (v_{ij})$ with integer entries and $\det(V) = \det(U) = \pm 1$ such that

$$\sum_{k=1}^{n} v_{ik} u_{kj} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Let $a_1, \ldots, a_n$ be a basis for the lattice Λ in $\mathbf{R}^n$. We shall use the matrix U to construct another basis for Λ. For $i = 1, \ldots, n$, let

$$a'_i = \sum_{j=1}^{n} u_{ij} a_j.$$

Let Λ' be the group in $\mathbf{R}^n$ generated by the n vectors $a'_1, \ldots, a'_n$. Since $a'_i \in \Lambda$ for $i = 1, \ldots, n$, it follows that $\Lambda' \subseteq \Lambda$. Since V is the inverse of the matrix U,

$$\begin{aligned} \sum_{k=1}^{n} v_{ik} a'_k &= \sum_{k=1}^{n} v_{ik} \sum_{j=1}^{n} u_{kj} a_j \\ &= \sum_{j=1}^{n} \sum_{k=1}^{n} v_{ik} u_{kj} a_j \\ &= \sum_{j=1}^{n} \delta_{ij} a_j \\ &= a_i \in \Lambda', \end{aligned}$$

and so $\Lambda \subseteq \Lambda'$. Thus, $\Lambda = \Lambda'$, and the vectors $a_1, \ldots, a_n$ and $a'_1, \ldots, a'_n$ are both bases for the lattice Λ.

Conversely, let $a_1, \ldots, a_n$ and $a'_1, \ldots, a'_n$ be two bases for the lattice Λ. For $i, j, k = 1, \ldots, n$, there exist integers u_{ik} and v_{kj} such that

$$a'_i = \sum_{k=1}^{n} u_{ik} a_k$$

and

$$a_k = \sum_{j=1}^{n} v_{kj} a'_j.$$

This implies that

$$\begin{aligned} a'_i &= \sum_{k=1}^{n} u_{ik} a_k \\ &= \sum_{k=1}^{n} u_{ik} \sum_{j=1}^{n} v_{kj} a'_j \\ &= \sum_{k=1}^{n} \left(\sum_{j=1}^{n} u_{ik} v_{kj} \right) a'_j. \end{aligned}$$

Since the n vectors $a'_1, \ldots, a'_n$ are linearly independent, it follows that

$$\sum_{k=1}^{n} u_{ik} v_{kj} = \delta_{ij}$$

for $i, j = 1, \ldots, n$. Similarly,

$$\sum_{k=1}^{n} v_{ik} u_{kj} = \delta_{ij}$$

for $i, j = 1, \ldots, n$. Let U and V be the matrices $U = (u_{ij})$ and $V = (v_{ij})$. Then $V = U^{-1}$, and, since the matrix elements u_{ij} and v_{ij} are integers, $\det(U) = \det(V) = \pm 1$. Thus, any two bases for a lattice Λ in $\mathbf{R}^n$ differ by a unimodular transformation.

Let $a_1, \ldots, a_n$ and $a'_1, \ldots, a'_n$ be two bases for the lattice Λ, and let $U = (u_{ij})$ be the unimodular integer matrix such that

$$a'_i = \sum_{j=1}^{n} u_{ij} a_j.$$

Let

$$a_i = \sum_{j=1}^{n} a_{ij} e_j$$

and

$$a'_i = \sum_{j=1}^{n} a'_{ij} e_j,$$

where a_{ij} and a'_{ij} are the coordinates of a_i and a'_i with respect to the standard basis vectors $e_1, \ldots, e_n$. Then

$$\begin{aligned} \sum_{j=1}^{n} a'_{ij} e_j &= a'_i \\ &= \sum_{k=1}^{n} u_{ik} a_k \\ &= \sum_{k=1}^{n} u_{ik} \sum_{j=1}^{n} a_{kj} e_j \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} u_{ik} a_{kj} e_j, \end{aligned}$$

and so

$$a'_{ij} = \sum_{k=1}^{n} u_{ik} a_{kj}$$

for $i, j = 1, \ldots, n$. In terms of the matrices $A = (a_{ij})$ and $A' = (a'_{ij})$, the preceding identity is equivalent to the matrix equation

$$A' = UA$$

and so

$$|\det(A')| = |\det(AU)| = |\det(U)||\det(A)| = |\det(A)|. \tag{6.1}$$

The *determinant of the lattice* Λ, denoted $d(\Lambda)$, plays a fundamental role in the geometry of numbers. This determinant is defined by

$$d(\Lambda) = |\det(A)| = |\det(a_1, \ldots, a_n)|,$$

where $a_1, \ldots, a_n$ is a basis for Λ. It follows from 6.1 that $d(\Lambda)$ is independent of the choice of basis for the lattice Λ. Moreover, $(\Lambda) \neq 0$, since the basis for a lattice is a set of n linearly independent vectors in $\mathbf{R}^n$.

The *fundamental domain* or *fundamental parallelepiped* of the lattice Λ with respect to the basis $a_1, \ldots, a_n$ is the set

$$\begin{aligned} F(\Lambda) &= F(\Lambda; a_1, \ldots, a_n) \\ &= \{\sum_{i=1}^{n} x_i a_i : 0 \leq x_i < 1 \text{ for } i = 1, \ldots, n\} \subseteq \mathbf{R}^n. \end{aligned}$$

If $a_j = \sum_{i=1}^{n} a_{ij} e_i$, then the volume of the fundamental domain is

$$\begin{aligned} \mathrm{vol}(F(\Lambda; a_1, \ldots, a_n)) &= \int \cdots \int_{F(\Lambda; a_1, \ldots, a_n)} dV \\ &= \int_0^1 \cdots \int_0^1 |\det(a_{ij})| dx_1 \cdots dx_n \\ &= d(\Lambda). \end{aligned}$$

Thus, while the fundamental domain of a lattice depends on the choice of basis for the lattice, the volume of a fundamental domain is independent of the choice of basis.

Let v be any vector in $\mathbf{R}^n$. Then

$$v = \sum_{i=1}^{n} v_i a_i,$$

where $v_i \in \mathbf{R}$. Let $v_i = m_i + x_i$, where $m_i \in \mathbf{Z}$ and $x_i \in [0, 1)$. Then

$$\begin{aligned} \sum_{i=1}^{n} m_i a_i &\in \Lambda, \\ \sum_{i=1}^{n} x_i a_i &\in F(\Lambda), \end{aligned}$$

and so

$$v = \sum_{i=1}^{n} v_i a_i = \sum_{i=1}^{n} m_i a_i + \sum_{i=1}^{n} x_i a_i \in \Lambda + F(\Lambda).$$

Thus, $\mathbf{R}^n = \Lambda + F(\Lambda)$. If

$$v = \sum_{i=1}^{n} m_i a_i + \sum_{i=1}^{n} x_i a_i = \sum_{i=1}^{n} m_i' a_i + \sum_{i=1}^{n} x_i' a_i$$

with $m_i, m_i' \in \mathbf{Z}$ and $x_i, x_i' \in [0,1)$, then the linear independence of the vectors $a_1, \ldots, a_n$ implies that $m_i + x_i = m_i' + x_i'$, and so $m_i = m_i'$ and $x_i = x_i'$. Thus, the representation of a vector as the sum of an element in the lattice and an element in the fundamental domain is unique.

6.2 Bases for sublattices

Let Λ and M be lattices in $\mathbf{R}^n$ with bases $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, respectively. If $M \subseteq \Lambda$, then M is called a *sublattice* of Λ. In this case, $b_j \in \Lambda$ for $j = 1, \ldots, n$, and there exist integers v_{ij} such that

$$b_j = \sum_{i=1}^{n} v_{ij} a_i.$$

Let

$$a_j = \sum_{i=1}^{n} a_{ij} e_i$$

and

$$b_j = \sum_{i=1}^{n} b_{ij} e_i.$$

Then

$$b_{ij} = \sum_{k=1}^{n} v_{ik} a_{kj}.$$

This implies that

$$d(M) = |\det(b_{ij}| = |\det(v_{ik})||\det(a_{ij})| = |\det(v_{ik})| d(\Lambda).$$

Let

$$d = |\det(v_{ik})| = d(M)/d(\Lambda).$$

Then $d \neq 0$, and d is independent of the choice of bases for the lattices Λ and M. Also, d is an integer, since $v_{ik} \in \mathbf{Z}$ for all $i, k = 1, \ldots, n$. This number d is called the *geometrical index* of the lattice M in Λ.

Since $\det(v_{ij}) = \pm d$, it follows that the matrix $V = (v_{ij}) \in M_n(\mathbf{Z})$ has an inverse V^{-1} of the form

$$V^{-1} = (v_{ij}^{-1}) = (w_{ij}/d),$$

where $w_{ij} \in \mathbf{Z}$. This means that

$$\sum_{k=1}^{n} v_{ik} w_{kj} = d\delta_{ij}$$

for all $i, j = 1, \ldots, n$. Then

$$\begin{aligned}
\sum_{k=1}^{n} w_{kj} b_k &= \sum_{k=1}^{n} w_{kj} \sum_{i=1}^{n} v_{ik} a_i \\
&= \sum_{i=1}^{n} \sum_{k=1}^{n} v_{ik} w_{kj} a_i \\
&= \sum_{i=1}^{n} d\delta_{ij} a_i \\
&= da_j \in dM
\end{aligned}$$

for $j = 1, \ldots, n$, and so

$$d\Lambda \subseteq M \subseteq \Lambda.$$

Theorem 6.1 *Let Λ be a lattice in $\mathbf{R}^n$ with basis $a_1, \ldots, a_n$, and let M be a sublattice of Λ. There exists a basis $b_1, \ldots, b_n$ of M of the form*

$$\begin{aligned}
b_1 &= v_{11} a_1 \\
b_2 &= v_{12} a_1 + v_{22} b_2 \\
b_3 &= v_{13} a_1 + v_{23} b_2 + v_{33} b_3 \\
&\vdots \\
b_n &= v_{1n} a_1 + v_{2n} a_2 + \ldots + v_{nn} a_n
\end{aligned}$$

where $v_{ij} \in \mathbf{Z}$ for $i = 1, \ldots, n$ and $j = i, \ldots, n$, and $v_{ii} \geq 1$ for $i = 1, \ldots, n$.

Proof. Let d be the geometric index of M in Λ. Fix $i \in \{1, 2, \ldots, n\}$. Consider the set of all vectors in M of the form

$$x_1 a_1 + x_2 a_2 + \cdots + x_i a_i,$$

where $x_1, x_2, \ldots, x_i \in \mathbf{Z}$ and $x_i \neq 0$. This set is nonempty, since $d\Lambda \subseteq M$, and so $da_i \in M$. Let b_i be the vector in this set for which $|x_i|$ is minimal. Since M is a lattice, if $v \in M$, then also $-v \in M$, and so we can assume that $x_i \geq 1$. Let

$$b_i = v_{i1} a_1 + v_{i2} + \cdots + v_{ii} a_i.$$

We shall prove that the n vectors $b_1, \ldots, b_n$ form a basis for M. Since $v_{ii} \geq 1$ for $i = 1, \ldots, n$, it follows that these vectors are linearly independent. Let $c \in M$. Then

$$c = w_{1n} a_1 + w_{2n} + \cdots + w_{nn} a_n,$$

where $w_{1n}, \ldots, w_{nn} \in \mathbf{R}$. Let

$$w_{nn} = q_n v_{nn} + r_n,$$

where $q_n \in \mathbf{Z}$ and $0 \leq r_n < v_{nn}$. Since $q_n v_{nn} \in M$, it follows that

$$c - q_n b_n \in M$$

and

$$\begin{aligned} c - q_n b_n &= (w_{1n} - q_n v_{n1})a_1 + \cdots + (w_{n-1,n} - q_n v_{n,n-1} a_{n-1} + (w_{nn} - q_n v_{nn} a_n \\ &= (w_1 - q_n v_{n1})a_1 + \cdots + (w_{n-1} - q_n v_{n,n-1})a_{n-1} + r_n a_n. \end{aligned}$$

Since $0 \leq r_n < v_{nn}$, it follows from the minimality in the choice of the vector b_n that $r_n = 0$, and so

$$c - q_n b_n = \sum_{i=1}^{n-1} w_{i,n-1} a_i.$$

Let

$$w_{n-1,n-1} = q_{n-1} v_{n-1,n-1} + r_{n-1},$$

where $q_{n-1} \in \mathbf{Z}$ and

$$0 \leq r_{n-1} < v_{n-1,n-1}.$$

By the previous argument, it follows that

$$c - q_n b_n - q_{n-1} b_{n-1} = \sum_{i=1}^{n-2} w_{i,n-2} a_i.$$

Repeating this argument inductively, we conclude that there exist integers $q_1, q_2, \ldots, q_n$ such that

$$c = q_1 b_1 + q_2 b_2 + \cdots + q_n b_n,$$

and so $b_1, \ldots, b_n$ is a basis for the sublattice M of the desired type. This completes the proof.

Corollary 6.1 *Let M be a sublattice of the lattice Λ, and let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be bases for Λ and M, respectively, such that $b_i = \sum_{j=1}^{i} v_{ij} a_j$. Let d be the geometric index of M in Λ. Then*

$$d = v_{11} v_{22} \cdots v_{nn}.$$

Proof. This follows immediately from the fact that $d = \det(v_{ij})$.

Corollary 6.2 *Let Λ be a sublattice of the integer lattice $\mathbf{Z}^n$. Then there exists a basis $a_1, \ldots, a_n$ for Λ such that*

$$\begin{aligned} a_1 &= a_{11} e_1 \\ a_2 &= a_{21} e_1 + a_{22} e_2 \\ &\vdots \\ a_n &= a_{n1} e_1 + a_{n2} e_2 \cdots + a_{nn} en, \end{aligned}$$

where $a_{ij} \in \mathbf{Z}$ for $i = 1, \ldots, n$ and $j = i, \ldots, n$, and $a_{ii} \geq 1$ for $i = 1, \ldots, n$. Moreover, $d(\Lambda) = a_{11} a_{22} \cdots a_{nn}$.

Proof. The standard basis $e_1, \ldots, e_n$ for $\mathbf{R}^n$ is a basis for the integer lattice $\mathbf{Z}^n$, and Λ is a sublattice of $\mathbf{Z}^n$. The result follows immediately from the Theorem.

Theorem 6.2 *Let M be a sublattice of the lattice Λ in $\mathbf{R}^n$. Let $b_1, \ldots, b_n$ be a basis for M. There exists a basis $a_1, \ldots, a_n$ of Λ of the form*

$$\begin{aligned} b_1 &= v_{11}a_1 \\ b_2 &= v_{12}a_1 + v_{22}b_2 \\ b_3 &= v_{13}a_1 + v_{23}b_2 + v_{33}b_3 \\ &\vdots \\ b_n &= v_{1n}a_1 + v_{2n}a_2 + \ldots + v_{nn}a_n \end{aligned}$$

where $v_{ij} \in \mathbf{Z}$ for $i = 1, \ldots, n$ and $j = i, \ldots, n$, and $v_{ii} \geq 1$ for $i = 1, \ldots, n$.

Proof. Let d be the geometric index of M in Λ. Then $d\Lambda \subseteq M$. It follows from the previous Theorem that there exists a basis $a'_1, \ldots, a'_n$ for $d\Lambda$ which is "upper triangular with respect to $b_1, \ldots, b_n$Since $a'_i \in d\Lambda$ for $i = 1, \ldots, n$, then $a'_i = da_i$ for some $a_i \in \Lambda$, and $a_1, \ldots, a_n$ is a basis for Λ. Solving the system of linear equations

$$\begin{aligned} da_1 &= w_{11}b_1 \\ da_2 &= w_{12}b_1 + w_{22}b_2 \\ da_3 &= w_{13}b_1 + w_{23}b_2 + w_{33}b_3 \\ &\vdots \\ da_n &= w_{1n}b_1 + w_{2n}b_2 + \ldots + w_{nn}b_n \end{aligned}$$

successively for the vectors $b_1, \ldots, b_n$ in terms of the vectors $a_1, \ldots, a_n$ we obtain another upper triangular system with rational coefficients v_{ij}:

$$\begin{aligned} b_1 &= (d/w_{11})a_1 = v_{11}a_1 \\ b_2 &= -(dw_{12}/w_{11}w_{22})a_1 + (d/w_{22})a_2 \\ &= v_{12}a_1 + v_{22}a_2 \\ b_3 &= v_{13}a_1 + v_{23}a_2 + v_{33}a_3 \\ &\vdots \\ b_n &= v_{1n}a_1 + v_{2n}a_2 + \ldots + v_{nn}a_n. \end{aligned}$$

The coefficients $v_{ii} = d/w_{ii}$ are positive for $i = 1, \ldots, n$. Since $a_1, \ldots, a_n$ is a basis for Λ, every element in Λ is an integral linear combination of the vectors $a_1, \ldots, a_n$. In particular, since $b_i \in M \subseteq \Lambda$, it follows that each coefficient w_{ij} is an integer. Thus, the basis $a_1, \ldots, a_n$ for Λ satisfies the requirements of the Theorem. This completes the proof.

Let M be a sublattice of the lattice Λ. Then M is a subgroup of the abelian group Λ. The quotient group Λ/M is the set of all cosets

$$a + M = \{a + b : b \in M\},$$

where $a \in \Lambda$, and the addition of two cosets is defined by $(a_1+M)+(a_2+M) = (a_1+a_2)+M$. The *algebraic index* of M in Λ is the order of the quotient group Λ/M, and is denoted $[\Lambda : M]$. Note that $a_1 + M = a_2 + M$ if and only if $a_1 - a_2 \in M$.

Theorem 6.3 *Let M be a sublattice of the lattice Λ in $\mathbf{R}^n$. Then*

$$[\Lambda : M] = d(M)/d(\Lambda),$$

that is, the geometric index of M in Λ is equal to the algebraic index of M in Λ.

Proof. Let $a_1, \ldots, a_n$ be a basis for Λ, and let $b_1, \ldots, b_n$ be a basis for the sublattice M of the form $b_j = \sum_{i=1}^{j} v_{ij}a_i$, where v_{jj} is a positive integer with the property that, if $\sum_{i=1}^{j} u_i a_i \in M$ and $u_i \neq 0$, then $|u_j| \leq v_{jj}$. We shall prove that the set

$$S = \{\sum_{i=1}^{n} w_i a_i : w_i \in \mathbf{Z}, 0 \leq w_i < v_{ii}\}$$

is a complete set of representatives of the cosets in Λ/M.

Let $s = \sum_{i=1}^{n} w_i a_i \in S$ and $s' = \sum_{i=1}^{n} w_i' a_i \in S$. If $s + M = s' + M$, then

$$s - s' = \sum_{i=1}^{n} (w_i - w_i')a_i \in M.$$

If $s \neq s'$, let j be the largest integer such that $w_j \neq w_j'$. Then $0 \leq w_j, w_j' < v_{jj}$ implies that $|w_j - w_j'| < v_{jj}$, which contradicts the minimality of v_{jj}. Therefore, $s + M \neq s' + M$, and so the elements of S represent distinct cosets in Λ/M.

Let $a = \sum_{i=1}^{n} x_i a_i \in \Lambda$, and let $x_n = q_n v_{nn} + r_n$, where $0 \leq r_n < v_{nn}$. Then

$$a - q_n b_n = \sum_{i=1}^{n-1} x_i' a_i + r_n a_n,$$

and

$$(a - q_n b_n) + M = a + M$$

in Λ/M. Similarly, let $x_{n-1} = q_{n-1} v_{n-1,n-1} + r_{n-1}$, where $0 \leq r_{n-1} < v_{n-1,n-1}$. Then

$$a - q_n b_n - q_{n-1} b_{n-1} = \sum_{i=1}^{n-2} x_i'' a_i + r_{n-1} a_{n-1} + r_n a_n,$$

and

$$(a - q_n b_n) - q_{n-1} b_{n-1} + M = a + M$$

in Λ/M. Continuing inductively, we obtain integers $q_1, \ldots, q_n$ and $r_1, \ldots, r_n$ such that

$$a - q_1 b_1 - q_2 b_2 - \cdots - q_n b_n = r_1 a_1 + r_2 a_2 + \cdots + r_n a_n = s \in S,$$

and so

$$a = s + q_1 b_1 + q_2 b_2 + \cdots + q_n b_n \in M.$$

Thus, the elements of S form a complete set of representatives for the cosets in the quotient group Λ/M. This proves that

$$\Lambda : M] = v_{11} v_{22} \cdots v_{nn} = d(M)/d(\Lambda).$$

6.3 Convex bodies and Minkowski's first theorem

Theorem 6.4 *Let Λ be a lattice in $\mathbf{R}^n$ and let K be an open, convex body in $\mathbf{R}^n$ symmetric with respect to the origin. If $vol(K) > 2^n det(\Lambda)$, then K contains a nonzero element of the lattice Λ.*

6.4 Applications to sums of squares

6.5 Successive minima and Minkowski's second theorem

Theorem 6.5 *Let Λ be a lattice in $\mathbf{R}^n$ and let K be a compact, convex body in $\mathbf{R}^n$ symmetric with respect to the origin. Let $\lambda_1, \ldots, \lambda_n$ be the successive minima of K with respect to the lattice Λ. Then*

$$\lambda_1 \cdots \lambda_n \leq 2^n det(\Lambda)/vol(K).$$

6.6 Torsion-free abelian groups

The abelian group G is *torsion-free* if $mg = 0$ for $m \in \mathbf{Z}$ and $g \in G$ only if either $m = 0$ or $g = 0$. The set $\{g_i\}_{i \in I} \subseteq G$ is a *set of generators* for G if every element g of G can be represented in the form

$$g = \sum_{i \in I} m_i g_i,$$

where $m_i \in \mathbf{Z}$ and $m_i = 0$ for all but finitely many $i \in I$. The group G is *finitely generated* if it contains a finite set of generators. The abelian group G is *free* if it contains a subset $\{g_i\}_{i \in I}$ such that every element $g \in G$ has a *unique* representation in the form

$$g = \sum_{i \in I} m_i g_i,$$

where $m_i \in \mathbf{Z}$ and $m_i = 0$ for all but finitely many $i \in I$. In this case, the set $\{g_i\}_{i \in I}$ is called a *basis* for G. Every free abelian group is torsion-free. A finitely generated free abelian group is a free abelian group with a finite basis. The group $G = \{0\}$ is the finitely generated free abelian group whose basis is the empty set.

Lemma 6.1 *Let G be a finitely generated free abelian group, $G \neq \{0\}$. Then $G \cong \mathbf{Z}^n$ for some $n \geq 1$.*

Proof. Let $\{g_1, \ldots, g_n\}$ be a basis for G. It suffices to observe that the map $\phi : G \to \mathbf{Z}^n$ defined by $\phi(\sum_{i=1}^n m_i g_i) = (m_1, \ldots, m_n)$ is an isomorphism. This completes the proof.

Lemma 6.2 *Let G be an abelian group, and let A be a free abelian group. Let $\phi : G \to A$ be a surjective homomorphism. Then $G = K \bigoplus H$, where K is the kernel of ϕ, and H is a subgroup G such that $\phi : H \to A$ is an isomorphism.*

Proof. Let $\{a_i\}_{i \in I}$ be a basis for the free abelian group A. Since the map ϕ is onto, there exist elements $h_i \in G$ such that $\phi(h_i) = a_i$. Let H be the subgroup of G generated by $\{h_i\}_{i \in I}$. The homomorphism ϕ restricted to H maps H onto A. Let $h = \sum_{i \in I} m_i h_i \in H$. Then

$$\begin{aligned} \phi(h) &= \phi(\sum_{i \in I} m_i h_i) \\ &= \sum_{i \in I} m_i \phi(h_i) \\ &= \sum_{i \in I} m_i a_i \\ &= 0 \end{aligned}$$

if and only if $m_i = 0$ for all $i \in I$, or, equivalently, $\phi(h) = 0$ if and only if $h = 0$. Thus, $\phi : H \to A$ is an isomorphism. If $h \in K \cap H$, then $h \in K$ implies that $\phi(h) = 0$, and $h \in H$ implies that $h = 0$, hence $K \cap H = \{0\}$. Since $\phi(H) \to A$ is onto, for every $g \in G$ there exists $h \in H$ such that $\phi(g) = \phi(h)$, and so $g - h \in K$. Therefore, $G = K + H$, and so $G = K \bigoplus H$. This completes the proof.

Lemma 6.3 *A subgroup of a finitely generated free abelian group is a finitely generated free abelian group.*

Proof. Let G be a finitely generated free abelian group with basis $\{g_1, \ldots, g_r\}$, and let H be a subgroup of G. If $H = \{0\}$, the lemma holds. Suppose that $H \neq \{0\}$. The proof will be by induction on r.

If $r = 1$, then $G = \mathbf{Z}g_1$ for some $g_1 \in G$. Since every subgroup of $\mathbf{Z}$ is of the form $\mathbf{Z}d$ for some $d \in \mathbf{Z}$, it follows that H is the free abelian group with basis $\{dg_1\}$.

Let $r \geq 2$, and suppose that the lemma holds for any free abelian group with a basis of at most $r - 1$ elements. Let G be a free abelian group with basis $\{g_1, \ldots, g_r\}$, and let G' be the subgroup of G with basis $\{g_1, \ldots, g_{r-1}\}$. Then G' is a finitely generated free abelian group with a basis of cardinality $r - 1$. Define $\phi : G \to \mathbf{Z}g_r$ by

$$\phi(\sum_{i=1}^{r} m_i g_i) = n_r g_r.$$

Then G' is the kernel of ϕ.

Let $\phi' : H \to \mathbf{Z}g_r$ be restriction of the homomorphism ϕ to H. Then the image of ϕ' is $\mathbf{Z}dg_r \cong \mathbf{Z}$ for some integer $d \in \mathbf{Z}$. Let H' be the kernel of ϕ'. Since

$$H' = H' \cap G' \subseteq G',$$

and since G' is free with a basis of cardinality $r-1$, it follows from the induction hypothesis that H' is a finitely generated free abelian group. The map ϕ' maps H onto the free abelian group $\mathbf{Z}dg_r$, and so, by the previous lemma, $H \cong H' \bigoplus H''$, where $\phi' : H'' \to \mathbf{Z}dg_r$ is an isomorphism. Therefore, H'' is free. Since the direct sum of two finitely generated free abelian groups is a finitely generated free abelian group, the proof of the lemma is complete.

Theorem 6.6 *Let G be a finitely generated torsion-free abelian group, $G \neq \{0\}$. Then G is isomorphic to the integer lattice $\mathbf{Z}^n$ for some $n \geq 1$.*

Proof. Let $B = \{g_j\}_{j \in J}$ be a finite set of generators for G, and let $B' = \{g'_1, \ldots, g'_r\}$ be a maximal subset of B such that $\sum_{i=1}^{r} m_i g'_i = 0$ with $m_i \in \mathbf{Z}$ if and only if $m_i = 0$ for all $i = 1, \ldots, r$. Let G' be the subgroup of G generated by $\{g'_1, \ldots, g'_r\}$. Then G' is a finitely generated free abelian group. Let $g_j \in B$. By the maximality of B', there exist integers $u_j, m_1, \ldots, m_r$ not all zero such that

$$u_j g_j + m_1 g'_1 + \cdots + m_r g'_r = 0.$$

If $u_j = 0$, then the condition on the set B' implies that $m_i = 0$ for $i = 1, \ldots, r$, which is impossible. Thus, $u_j \neq 0$, and so $u_j g_j \in G'$. Let $m = \max\{|u_j| : j \in J\}$. Then $mg_j \in G'$ for all $g_j \in B$. Since B generates G, it follows that

$$G^{(m)} = \{mg : g \in G\} =\subseteq G'.$$

Since G' is a finitely generated free abelian group, it follows from the previous lemma that the subgroup G^m is also a finitely generated free abelian group. Since G is torsion-free, the map $\psi : G \to G^m$ defined by $\psi(g) = mg$ is an isomorphism, and so G is also a finitely generated free abelian group. This completes the proof of the theorem.

6.7 Notes

The proof of Theorem 6.6 follows Lang [42].

6.8 Exercises

1. Let G be a free abelian group with finite basis $\{g_1, \ldots, g_r\}$. Let $m \geq 2$. Define the homomorphism $\phi : G \to \bigoplus_{i=1}^{r} \mathbf{Z}/m\mathbf{Z}$ by

$$\phi(\sum_{i=1}^{r} m_i g_i) = (n_1 \pmod m, \ldots, n_r \pmod m).$$

   Prove that $G/mG \cong \bigoplus_{i=1}^{r} \mathbf{Z}/m\mathbf{Z}$, and so $|G/mG| = m^r$. Conclude that every basis for G has cardinality r. This number is called the *rank* of G.

Chapter 7

Plünnecke's inequality

7.1 Plünnecke graphs

A *directed graph* $G = (V(G), E(G))$ consists of a finite set $V(G)$ of *vertices* and a set $E(G)$ of *edges*, where each edge $e \in E(G)$ is an ordered pair (v, v') of distinct elements of $V(G)$. Let $h \geq 1$. A directed graph $G = (V(G), E(G))$ is a *graph of level* h if the vertex set $V(G)$ is the union of $h+1$ pairwise disjoint nonempty sets $V_0, V_1, \ldots, V_h$, and if every edge of G is of the form (v, v'), where $v \in V_{i-1}$ and $v' \in V_i$ for some $i = 1, \ldots, h$; that is,

$$E(G) \subseteq \bigcup_{i=1}^{h} (V_{i-1} \times V_i) .$$

A directed graph $G = (V(G), E(G))$ of level h is a *Plünnecke graph of level* h if it satisfies the following two conditions:

1. Let $1 \leq i \leq h-1$ and $k \geq 2$, and let $u \in V_{i-1}$, $v \in V_i$, and $w_1, \ldots, w_k \in V_{i+1}$ be $k+2$ distinct vertices of G such that $(u, v) \in E(G)$ and $(v, w_j) \in E(G)$ for $j = 1, \ldots, k$. Then there exist distinct vertices $v_1, \ldots, v_k \in V_i$ such that $(u, v_j) \in E(G)$ and $(v_j, w_j) \in E(G)$ for $j = 1, \ldots, k$.

2. Let $1 \leq i \leq h-1$ and $k \geq 2$, and let $u_1, \ldots, u_k \in V_{i-1}$, $v \in V_i$, and $w \in V_{i+1}$ be $k+2$ distinct vertices of G such that $(u_j, v) \in E(G)$ for $j = 1, \ldots, k$, and $(v, w) \in E(G)$. Then there exist distinct vertices $v_1, \ldots, v_k \in V_i$ such that $(u_j, v_j) \in E(G)$ and $(v_j, w) \in E(G)$ for $j = 1, \ldots, k$.

Let $G = (V(G), E(G))$ be a directed graph. A *path in* G *from vertex* a *to vertex* b is a finite sequence of vertices $a = v_0, v_1, \ldots, v_{k-1}, v_k = b$ such that $(v_{i-1}, v_i) \in E(G)$ for $i = 1, \ldots, k$, and $v_i \neq a, b$ for $i = 1, 2, \ldots, k-1$. The path can also be identified with the sequence of edges $(v_0, v_1), (v_1, v_2), \ldots, (v_{k-1}, v_k)$. The vertices $v_1, \ldots, v_{k-1}$ are called the *intermediate vertices* of the path.

Let X and Y be nonempty subsets of $V(G)$. The *image* of X in Y, denoted $\mathrm{im}(X, Y)$, is the set of all $y \in Y$ such that there exists a path from x to y for some $x \in X$. The

magnification ratio of X in Y is

$$D(X,Y) = \min\left\{\frac{|\mathrm{im}(Z,Y)|}{|Z|} : \emptyset \neq Z \subseteq X\right\}.$$

Let G be a graph of level h with vertex set $V(G) = \cup_{i=0}^{h} V_i$. For $i = 1, \ldots, h$, the *i-th magnification ratio of G* is

$$D_i = D_i(G) = D(V_0, V_i).$$

Plünnecke proved that if G is a Plünnecke graph of level h, then

$$D_h^{1/h} \leq D_{h-1}^{1/(h-1)} \leq \cdots \leq D_2^{1/2} \leq D_1.$$

These simple inequalities have powerful consequences in additive number theory.

7.2 Examples of Plünnecke graphs

Addition graphs. Let A and B be subsets of an abelian group. We want to construct a Plünnecke graph of level h whose i-th vertex set is the sumset $A + iB$ and whose edges are the ordered pairs of group elements of the form $(v, v+b)$, where $b \in B$ and $v \in A + (i-1)B$ for some $i = 1, \ldots, h$. Since the sumsets $\{A + iB\}_{i=0}^{h}$ are not necessarily pairwise disjoint, we have to be careful about the definition of the addition graph.

Let $h \geq 1$, and let A and B be finite, nonempty subsets of an abelian group. The *addition graph G of level h* constructed from the pair A, B is the graph whose vertex set $V(G)$ and edge set $E(G)$ are defined as follows:

$$V(G) = \cup_{i=0}^{h} V_i,$$

where

$$V_i = (A + iB) \times \{i\},$$

and

$$E(G) = \bigcup_{i=1}^{h} \{((v, i-1), (v+b, i)) : v \in V_{i-1}, b \in B\}.$$

It is clear that the sets $V_0, V_1, \ldots, V_h$ are pairwise disjoint, and that the pair $((v,j),(v',j')) \in E(G)$ if and only if $v' - v \in B$ and $j' - j = 1$. Thus, the addition graph is a directed graph of level h.

We shall prove that the addition graph is a Plünnecke graph. Let $1 \leq i \leq h-1$, and let $(u, i-1) \in V_{i-1}$, $(v,i) \in V_i$, and $(w_1, i+1), \ldots, (w_k, i+1) \in V_{i+1}$ be $k+2$ distinct vertices in $V(G)$ such that

$$((u, i-1), (v,i)) \in E(G)$$

and

$$((v,i), (w_j, i+1)) \in E(G)$$

for $j = 1, \ldots, k$. Then $v - u = b \in B$ and $w_j - v = b_j \in B$ for $j = 1, \ldots, k$. Let $v_j = u + b_j$. Since the k elements w_j are distinct, it follows that the k elements b_j are distinct and so the k elements v_j are distinct. Since $u \in A + (i-1)B$, it follows that $v_j \in A + iB$ and so $((u, i-1), (v_j, i)) \in E(G)$ for $j = 1, \ldots, k$. Since A and B are subsets of an abelian group, it follows that

$$w_j = b_j + v = b_j + (b + u) = (u + b_j) + b = v_j + b$$

and so $((v_j, i), (w_j, i+1)) \in E(G)$ for $j = 1, \ldots, k$. This shows that the addition graph satisfies the first property of Plünnecke graphs. It can be shown in the same way that the second property is satisfied.

Addition graphs are an important class of Plünnecke graphs. To simplify notation, we shall henceforth denote the i-th vertex set in the addition graph by $A + iB$.

Truncated addition graphs Let $n \geq 1$ and let A and B be finite, nonempty sets of nonnegative integers such that $A \cap [1, n] \neq \emptyset$ and $0 \in B$. Then

$$V_i = A + iB \cap [1, n] \neq \emptyset$$

for $i = 0, 1, 2, \ldots$. The *truncated addition graph* G of level h constructed from the pair of sets A and B is the graph with vertex set $V(G) = \cup_{i=0}^{h} V_i$ and edge set

$$E(G) = \{(v, v') \in V_{i-1} \times V_i : v' - v \in B\}.$$

This is a Plünnecke graph of level h.

Independent addition graphs. Let $B = \{b_1, \ldots, b_n\}$ be a set of n elements of an abelian group such that the $\binom{n+h-1}{h}$ h-fold sums of the form $b_{j_1} + \cdots + b_{j_h}$ with $j_1 \leq \cdots \leq j_h$ are distinct. Let $I_{n,h}$ denote the addition graph of level h constructed from the sets $A = \{0\}$ and B. Then $V_0 = \{0\}$, and $V_i = iB$ and

$$\frac{n^i}{i!} \leq |V_i| = \binom{n+i-1}{i} \leq n^i.$$

for $i = 1, \ldots, h$. The graph $I_{n,h}$ is called the *independent addition graph* of level h on n elements.

It is easy to construct examples of independent addition graphs. Let $A = \{0\}$ and let $B = \{b_1, \ldots, b_n\}$ be a set of n positive integers such that $b_i > hb_{i-1}$ for $i = 2, \ldots, h$. Then the h-fold sums of elements of $B \cup \{0\}$ are distinct, and the $h + 1$ sets $\{V_i\}_{i=0}^{h}$ are pairwise disjoint.

Contracted graphs Let $G = (V(G), E(G))$ be a Plünnecke graph of level h, where $V(G) = \cup_{i=0}^{h} V_i$. Let $a, b \in V(G)$. A *path in G from vertex a to vertex b* is a finite sequence of vertices $a = v_0, v_1, \ldots, v_{k-1}, v_k = b$ such that $(v_{i-1}, v_i) \in E(G)$ for $i = 1, \ldots, k$, and $v_i \neq a, b$ for $i = 1, \ldots, k-1$. Let $0 \leq j < k \leq h$, and let X and Y be nonempty subsets of V_j and V_k, respectively, such that for some $a \in X$ and $b \in Y$ there is a path in G from a to b. Let $V(X, Y)$ be the set of all vertices $v \in V(G)$ that lie on some path from X to Y. Let

$$V_i(X, Y) = V(X, Y) \cap V_{j+i}$$

for $i = 0, 1, \ldots, k-j$. Then

$$V(X,Y) = \cup_{i=0}^{k-j} V_i(X,Y).$$

Since there exists a path from the vertex $a \in X$ to the vertex $b \in Y$, it follows that $V_i(X,Y) \neq \emptyset$ for $i = 0, 1, \ldots, k-j$. Let $G(X,Y)$ be the graph with vertex set $V(X,Y)$ and edge set $E(X,Y)$ consisting of all edges $(v,v') \in E(G)$ with $v, v' \in V(X,Y)$. Then the *contracted graph* $G(X,Y)$ is a Plünnecke graph of level $k-j$.

Product graphs. Let $h \geq 1$, $r \geq 2$, and let $G_1, \ldots, G_r$ be Plünnecke graphs of level h. We construct the *product graph* $G = G_1 \times \cdots \times G_r$ as follows: Let $V(G_j) = \cup_{i=0}^{h} V_{i,j}$ for $j = 1, \ldots, r$, and define

$$V_i = V_{i,1} \times \cdots \times V_{i,r}$$

for $i = 0, 1, \ldots, h$. Let $1 \leq i \leq h$, and let $v_{i-1,j} \in V_{i-1,j}$ and $v_{i,j} \in V_{i,j}$ for $j = 1, \ldots, r$. Let $E(G)$ consist of all ordered pairs of r-tuples

$$((v_{i-1,1}, \ldots, v_{i-1,r}), (v_{i,1}, \ldots, v_{i,r})) \in V_{i-1} \times V_i$$

such that $(v_{i-1,j}, v_{i,j}) \in E(G_j)$ for $j = 1, \ldots, r$. It is easy to check that the product graph G is a Plünnecke graph of level h.

Inverse graphs. Let G be a Plünnecke graph of level h with vertex set $V(G) = \cup_{i=0}^{h} V_i$ and edge set $E(G)$. Let $V_i^{-1} = V_{h-i}$ for $i = 0, 1, \ldots, h$. The *inverse graph* G^{-1} is the graph whose vertex set is $V(G^{-1}) = \cup_{i=0}^{h} V_i^{-1}$ and whose edge set $E(G^{-1})$ is determined by the condition that $(v,v') \in E(G^{-1})$ if and only if $(v',v) \in E(G)$. Then G^{-1} is a Plünnecke graph of level h.

7.3 Multiplicativity of magnification ratios

In this section we shall prove that the magnification ratios of graphs of level h are multiplicative.

Theorem 7.1 *Let G' and G'' be directed graphs of level h. Then*

$$D_i(G' \times G'') = D_i(G')D_i(G'')$$

for $i = 1, \ldots, h$.

Proof. Let $V(G') = \cup_{i=0}^{h} V_i'$ and $V(G'') = \cup_{i=0}^{h} V_i''$. Let Z' and Z'' be nonempty subsets of V_0' and V_0'', respectively, such that

$$D_i(G') = |\mathrm{im}(Z', V_i')|/|Z'|$$

and

$$D_i(G'') = |\mathrm{im}(Z'', V_i'')|/|Z''|.$$

Then

$$Z' \times Z'' \subseteq V_0' \times V_0'' = V_0(G' \times G'').$$

Since $|Z' \times Z''| = |Z'||Z''| \neq 0$, and

$$|\mathrm{im}(Z' \times Z'', V_i' \times V_i'')| = |\mathrm{im}(Z', V_i')||\mathrm{im}(Z'', V_i'')|,$$

it follows that

$$D_i(G' \times G'') \leq \frac{|\mathrm{im}(Z' \times Z'', V_i' \times V_i'')|}{|Z' \times Z''|} = D_i(G')D_i(G'').$$

To complete the proof of the Theorem, we must prove the reverse inequality

$$D_i(G' \times G'') \geq D_i(G')D_i(G'').$$

We do this first in the special case where G' and G'' are graphs of level 1, and G'' satisfies the additional conditions that its vertex sets V_0'' and V_1'' are both copies of a nonempty set T, and $E(G'')$ consists of the $|T|$ edges (t,t) for $t \in T$. Then $D_1(G'') = 1$.

Remark: Since the vertex sets V_0'' and V_1'' of a graph of level 1 must be disjoint, we can (just as in the construction of the addition graph) formally define $V_0'' = \{(t,0) : t \in T\}$ and $V_1'' = \{(t,1) : t \in T\}$. Then $E(G'') = \{((t,0),(t,1)) : t \in T\}$.

Let $G = G' \times G''$, and let Z be a nonempty subset of $V_0(G) = V_0' \times V_0'' = V_0' \times T$. Then

$$Z = \cup_{t \in T}(Z_t' \times \{t\}),$$

where

$$Z_t' = \{v' \in V_0' : (v',t) \in Z\}.$$

It follows that

$$\begin{aligned} |\mathrm{im}(Z, V_1' \times V_1'')| &= \sum_{t \in T} |\mathrm{im}(Z_t', V_1')| \\ &\geq D_1(G') \sum_{t \in T} |Z_t'| \\ &= D_1(G')|Z|, \end{aligned}$$

and so

$$D_1(G')D_1(G'') = D_1(G') \leq D_1(G' \times G'').$$

Therefore,

$$D_1(G')D_1(G'') = D_1(G' \times G'')$$

for graphs of this special type.

We now consider the general case. Let G' and G'' be graphs of level h, and let $1 \leq i \leq h$. Consider the graph $H = (V(H), E(H))$ of level 3 such that $V(H) = W_0 \cup W_1 \cup W_2$, where

$$\begin{aligned} W_0 &= V_0' \times V_0'' = V_0(G' \times G'') \\ W_1 &= V_i' \times V_0'' \\ W_2 &= V_i' \times V_i'' = V_i(G' \times G''), \end{aligned}$$

and $((a,c),(b,d))$ is an edge in $E(H)$ if and only if either $c = d \in V_0''$ and there is a path in G' from $a \in V_0'$ to $b \in V_i'$, or $a = b \in V_i'$ and there is a path in G'' from $c \in V_0''$ to $d \in V_i''$. It follows that there is a path in H from $(v_0', v_0'') \in W_0$ to $(v_i', v_i'') \in W_2$ if and only if there is a path in $G' \times G''$ from $(v_0', v_0'') \in V_0(G' \times G'')$ to $(v_i', v_i'') \in V_i(G' \times G'')$. Therefore,

$$D_2(H) = D_i(G' \times G'').$$

Let H_1 be the subgraph of H with vertices $W_0 \cup W_1$, and let H_2 be the subgraph of H with vertices $W_1 \cup W_2$. Then H_1 and H_2 are both graphs of level 1 of the special type considered above. Moreover,

$$D_1(H_1) = D_i(G')$$

and

$$D_1(H_2) = D_i(G'').$$

It follows that

$$\begin{aligned}
D_i(G' \times G'') &= D_2(H) \\
&= D(W_0, W_2) \\
&= \min\{\frac{|\mathrm{im}(Z, W_2)|}{|Z|} : \emptyset \neq Z \subseteq W_0\} \\
&= \min\{\frac{|\mathrm{im}(Z, W_1)|}{|Z|} \frac{|\mathrm{im}(Z, W_2)|}{|\mathrm{im}(Z, W_1)|} : \emptyset \neq Z \subseteq W_0\} \\
&= \min\{\frac{|\mathrm{im}(Z, W_1)|}{|Z|} \frac{|\mathrm{im}(\mathrm{im}(Z, W_1), W_2)|}{|\mathrm{im}(Z, W_1)|} : \emptyset \neq Z \subseteq W_0\} \\
&\geq \min\{\frac{|\mathrm{im}(Z, W_1)|}{|Z|} : \emptyset \neq Z \subseteq W_0\} \times \\
&\quad \times \min\{\frac{|\mathrm{im}(\mathrm{im}(Z, W_1), W_2)|}{|\mathrm{im}(Z, W_1)|} : \emptyset \neq Z \subseteq W_0\} \\
&\geq \min\{\frac{|\mathrm{im}(Z, W_1)|}{|Z|} : \emptyset \neq Z \subseteq W_0\} \times \\
&\quad \times \min\{\frac{|\mathrm{im}(Z', W_2)|}{|Z'|} : \emptyset \neq Z' \subseteq W_1\} \\
&= D_1(H_1) D_1(H_2) \\
&= D_i(G') D_i(G'').
\end{aligned}$$

This completes the proof.

Corollary 7.1 *Let $h \geq 1$ and $r \geq 2$, and let $G_1, \ldots, G_r$ be graphs of level h. Let G be the product graph $G_1 \times \cdots \times G_r$. Then*

$$D_i(G) = D_i(G_1) \cdots D_i(G_r)$$

for $i = 1, \ldots, h$.

7.4 Menger's theorem

Let $G = (V(G), E(G))$ be a directed graph. Let $a = v_0, v_1, \ldots, v_k = b$ and $a = w_0, w_1, \ldots, w_l = b$ be two paths in G from vertex a to vertex b. These paths are *disjoint* if $v_i \neq w_j$ for $i = 1, \ldots, k-1$ and $j = 1, \ldots, l-1$. A set S of vertices *separates the vertex a from the vertex b* if every path from a to b contains at least one element of S. Let S be a set that separates vertex a from vertex b, but contains neither a nor b. Let m be the maximum number of pairwise disjoint paths from a to b. Since S contains at least one vertex from each of these paths, it follows that $|S| \geq m$. Menger's theorem asserts that there exists a separating set S such that $|S| = m$.

Notation. Let $G = (V(G), E(G))$ be a directed graph, and let $V' \subseteq V(G)$. The *complete subgraph* of G generated by V' is the graph with vertex set V' and edge set

$$E' = \{(v, v') \in E(G) : v, v' \in V'\} = (V' \times V') \cap E(G).$$

Let $W \subseteq V$ and and $F \subseteq E(G)$. Let $G \setminus \{W, F\}$ denote the graph with vertex set $V(G) \setminus W$ and edge set consisting of all edges $(v, v') \in E(G) \setminus F$ such that $v, v' \in V(G) \setminus W$.

Theorem 7.2 (Menger) *Let a and b be vertices of a directed graph $G = (V(G), E(G))$, and suppose that $(a, b) \notin E(V)$. Let m be the maximum number of pairwise disjoint paths from the vertex a to the vertex b, and let S be a set of vertices of minimum cardinality that separates the vertex a from the vertex b and that contains neither a nor b. Then $|S| = m$.*

Proof. Let l be the cardinality of S. The proof will be by induction on l.

If $l = 0$, then $S = \emptyset$ and there are no paths from a to b, hence $m = 0$. If $l = 1$, then $S = \{v\}$, and there exists at least one path from a to b. Moreover, every path from a to b contains the intermediate vertex v. Thus, there cannot exist two pairwise disjoint paths from a to b, and so $m = 1$.

Let $l \geq 2$, and suppose that the theorem is true for any directed graph G and any pair of vertices a, b of G for which there exists a separating set, not containing a or b, that has cardinality at most $l - 1$. If the theorem is false for l, then there exists at least one graph with vertices a, b and separating set S not containing a or b with $|S| = l$, and with at most $l - 1$ pairwise disjoint paths from a to b.

Consider such graphs in which the number of vertices is minimal, and from this set of graphs choose one for which the number of edges is minimal. Call this graph $G = (V(G), E(G))$. Then G contains two vertices a, b with $(a, b) \notin E(G)$ such that there exists a separating set S of cardinality $|S| = l$ that contains neither a nor b, and there exist at most $l - 1$ pairwise disjoint paths from a to b. Let e be any edge in $E(V)$, and let $G' = G \setminus \{e\} = (V(G), E(G) \setminus \{e\})$. By the minimality of G, there exists a set $S(e)$ that separates a from b in G' and that contains $l - 1$ elements. If $e = (v_1, v_2)$ and $v_1 \neq a$, then $S(e) \cup \{v_1\}$ separates a from b in G, since every path from a to b either contains some vertex in $S(e)$, or includes the edge e. Similarly, if $v_2 \neq b$, then $S(e) \cup \{v_2\}$ separates a from b in G.

Let $v \in V(G)$. If a, v, b is a path in G, then $v \in S$ for every separating set S such that $S \cap \{a, b\} = \emptyset$. Moreover, $S' = S \setminus \{v\}$ separates a from b in the graph $G' = G \setminus \{v\}$

Conversely, if S' is a separating set for G', then $S' \cup \{v\}$ is a separating set for G, and so $|S'| \geq l-1$ for any separating set in G'. By the induction hypothesis, the graph G' contains $l-1$ pairwise disjoint paths from$a$ to b, and these paths, together with the path a, v, b, give l pairwise disjoint paths from a to b in G, which is impossible. Therefore, if v is any vertex in $V(G)$, then either $(a, v) \notin E(G)$ or $(v, b) \notin E(G)$.

Let $S = \{s_1, \ldots, s_l\}$ separate a from b in G, and suppose $a \notin S$ and $b \notin S$. We shall prove that either $(a, s_i) \in E(G)$ for $i = 1, \ldots, l$ or $(s_i, b) \in E(G)$ for $i = 1, \ldots, l$.

Let $P(a, S)$ denote the set of all paths in G that start at the vertex a and end at some vertex $s \in S$, with no intermediate vertex in S. Let $P(S, b)$ denote the set of all paths in G that start at some vertex $s \in S$ and end at b, with no intermediate vertex in S. Since S is a separating set, every path in G from a to b contains at least one point that lies in S, and so the path has an initial segment that belongs to $P(a, S)$ and a terminal segment that belongs to $P(S, b)$. Moreover, since the separating set S is minimal, for every $s \in S$ there is a path from a to s and from s to b.

Let $I(a, S)$ be the set of intermediate points of paths in $P(a, S)$, and let $I(S, b)$ be the set of intermediate points of paths in $P(S, b)$. If there exists $c \in I(a, S) \cap I(S, b)$, then there is a path in G from a to c that contains no point of S, and there is a path in G from c to b that contains no point of S. Concatenating these two paths produces a path in G from a to b that contains no element of S, which is impossible. Therefore,

$$I(a, S) \cap I(S, b) = \emptyset.$$

We shall prove that either $I(a, S) = \emptyset$ or $I(S, b) = \emptyset$. Suppose that both sets $I(a, S)$ and $I(S, b)$ are nonempty. Let $H = (V(H), E(H))$ be the graph with vertex set

$$\begin{aligned} V(H) &= \{a, s, b\} \cup I(S, b) \\ &= \{a\} \cup \{v \in V(G) : v \text{ lies on a path in } P(S, b)\}. \end{aligned}$$

and edge set

$$\begin{aligned} E(H) &= \{(a, s) : s \in S\} \cup \\ &\quad \cup \{(v, v') \in E(G) : (v, v') \text{ lies on a path in } P(S, b)\}. \end{aligned}$$

Since $I(a, S) \cap I(S, b) = \emptyset$ and $I(a, S) \neq \emptyset$, it follows that $I(a, S) \cap V(H) = \emptyset$ and so $|V(H)| < |V(G)|$.

Let T be a set that separates a from b in the graph H, with $a, b \notin T$. Let π be a path from a to b in G, and let π' be the terminal segment of this path that belongs to $P(S, b)$. Then π' is a path from some $s \in S$ to b. Since (a, s) is an edge in H, it follows that a, s concatenated with the path π' is a path in H from a to b. Since T separates a from b in H, it follows that T contains an intermediate vertex of this path, and so either $s \in T$ or some intermediate vertex of π' belongs to T. Thus, T also separates a from b in G. Therefore, $|T| \geq l$. Since the graph H contains strictly fewer vertices than the graph G, it follows from the minimality condition that there are l pairwise disjoint paths in H from a to b. In particular, for each

$s \in S$ there exists a path $\pi_2(s) \in P(S, b)$ from s to b such that the l paths $\pi_2(s)$ are pairwise disjoint.

Similarly, for each $s \in S$ there exists a path $\pi_1(s) \in P(a, S)$ from a to s such that the l paths $\pi_1(s)$ are pairwise disjoint. Concatenating the path $\pi_1(s)$ from a to s with the path $\pi_2(s)$ from s to b produces a path $\pi(s)$ from a to b, and these $|S| = l$ paths are pairwise disjoint, which is impossible in the graph G. Therefore, either $I(a, S) = \emptyset$ and $(a, s) \in E(G)$ for all $s \in S$, or $I(S, b) = \emptyset$ and $(s, b) \in E(G)$ for all $s \in S$. Moreover, since G contains no path of the form a, s, b, it follows that these two possibilities are mutually exclusive.

Let a, v, v' be the initial segment of a shortest path in G from a to b. Then $v' \neq b$. and $e = (v, v') \in E(G)$. Let $S(e)$ be a minimal separating set for the graph $G \setminus \{e\}$. Then $|S(e)| = l - 1 \geq 1$, and $S(e) \cup \{v\}$ is a separating set for G. Since $(a, v) \in E(G)$, it follows that $(a, s) \in E(G)$ for all $s \in S(e)$. Similarly, $S(e) \cup \{v'\}$ is a separating set for G. If $(a, v') \in E(G)$, then there is a shorter path from a to b than the one that starts a, v, v', wwhich is impossible. Therefore, $(a, v') \notin E(G)$. It follows that $(v', b) \in E(H)$, and so $(s, b) \in E(G)$ for all $s \in S(e)$. Thus, if $s \in S(e)$, then $(a, s) \in E(G)$ and $(s, b) \in E(G)$, and so a, s, b is a path in G, which is impossible. This completes the proof of Menger's theorem.

Let $G = (V(G), E(G))$ be a directed graph, and let X and Y be nonempty, disjoint sets of vertices of G. A *path in G from the set X to the set Y* is a finite sequence of vertices $v_0, v_1, \ldots, v_{k-1}, v_k$ such that $v_0 \in X$, $v_k \in Y$, and $(v_{i-1}, v_i) \in E(G)$ for $i = 1, \ldots, k$. Let $v_0, v_1, \ldots, v_k$ and $w_0, w_1, \ldots, w_l$ be two paths in the graph G from X to Y. These paths are *totally disjoint* if $v_i \neq w_j$ for $i = 0, 1, \ldots, k$ and $j = 0, 1, \ldots, l$. A set S of vertices *separates the set X from the set Y* if every path from X to Y contains at least one element of S.

Theorem 7.3 *Let X and Y be nonempty, disjoint sets of vertices of a directed graph G. Let m be the maximum number of pairwise totally disjoint paths from X to Y. Let n be the minimum cardinality of a set S of vertices that separates the set X from the set Y. Then $m = n$.*

Proof. Let $G = (V(G), E(G))$ be a directed graph. Let X and Y be nonempty, disjoint subsets of $V(G)$, and let m be the maximum number of pairwise totally disjoint paths from X to Y. Let a and b be elements not belonging to $V(G)$. We shall construct a new graph $G^* = (V(G^*), E(G^*))$ by adjoining two new vertices a and b to $V(G)$. Let

$$V(G^*) = V(G) \cup \{a, b\}$$

and

$$E(G^*) = E(G) \cup \{(a, x) : x \in X\} \cup \{(y, b) : y \in Y\}.$$

Then $a = v_0, v_1, \ldots, v_{k-1}, v_k = b$ is a path in G^* from vertex a to vertex b if and only if $v_1, \ldots, v_{k-1}$ is a path in G from the set X to the set Y. Two paths $a = v_0, v_1, \ldots, v_{k-1}, v_k = b$ and $a = w_0, w_1, \ldots, w_{l-1}, w_l = b$ are disjoint paths from a to b in G^* if and only if $v_1, \ldots, v_{k-1}$ and $w_1, \ldots, w_{l-1}$ are totally disjoint paths from X to Y in G. It follows that m is also the maximum number of pairwise disjoint paths from a to b in the graph G^*.

By Theorem 7.2, there exists a set $S \subseteq V(G^*)$ such that S separates a from b, $a, b \notin S$, and $|S| = m$. Then $S \subseteq V(G)$, and S separates X from Y in G. This completes the proof.

7.5 Plünnecke's inequality

Let $G = (V(G), E(G))$ be a directed graph, and let $v \in V(G)$. Let $d^+(v) = d^+(v, G)$ denote the number of vertices $v' \in V(G)$ such that $(v, v') \in E(G)$. Let $d^-(v) = d^-(v, G)$ denote the number of vertices $v' \in V(G)$ such that $(v', v) \in E(G)$.

Lemma 7.1 *Let G be a Plünnecke graph, and let $(u, v) \in E(G)$. Then*

$$d^+(u) \geq d^+(v)$$

and

$$d^-(u) \leq d^-(v).$$

Proof. These inequalities follow immediately from properties 1 and 2 of the definition of a Plünnecke graph.

Lemma 7.2 *Let G be a Plünnecke graph of level h, and let $V(G) = \cup_{i=0}^{h} V_i$. If $D_h \geq 1$, then there are $|V_0|$ totally disjoint paths from V_0 to V_h.*

Proof. Let m be the maximum number of pairwise totally disjoint paths from V_0 to V_h. By Theorem 7.3, there exists a set S of cardinality m that separates V_0 from V_h. For $v \in V(G)$, let $i(v)$ denote the unique integer i such that $v \in V_i$. Choose a separating set S such that $|S| = m$ and

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

is minimal. We shall prove that

$$S \subseteq V_0 \cup V_h.$$

Suppose not. Then $S \cap V_j \neq \emptyset$ for some $j \in [1, h-1]$. Let

$$S \cap V_j = \{s_1, \ldots, s_q\},$$

where $q = |S \cap V_j| \geq 1$. Let $\pi_1, \ldots, \pi_m$ be m pairwise totally disjoint paths from V_0 to V_h, and suppose that s_i is a vertex on path π_i for $i = 1, \ldots, q$. Since $0 < j < h$, it follows that, for each $i \in [1, q]$, the vertex s_i has a predecessor $r_i \in V_{j-1}$ on the path π_i and a successor $t_i \in V_{j+1}$ on the path π_i. It follows from the minimality of $\sum_{s \in S} i(s)$ that the set

$$S^* = \{r_1, \ldots, r_q, s_{q+1}, \ldots, s_m\}$$

does not separate V_0 from V_h, and so there exists a path π^* from V_0 to V_h that does not intersect S^*. However, since π^* cannot avoid the separating set S, it follows that s_i lies on π^* for some $i \in [1, q]$, say, $i = 1$. Let r^* be the predecessor of s_1 on the path π^*. Then $r^* \notin \{r_1, \ldots, r_q\}$.

We shall consider the following sets of vertices:

$$\begin{aligned} S_q^- &= \{r_1, \ldots, r_q\} \subseteq V_{j-1} \\ S_q^* &= \{r^*, r_1, \ldots, r_q\} \subseteq V_{j-1} \\ S_q &= \{s_1, \ldots, s_q\} \subseteq V_j \\ S_q^+ &= \{t_1, \ldots, t_q\} \subseteq V_{j+1}. \end{aligned}$$

The contracted graph

$$G^* = G(S_q^*, S_q^+).$$

is a Plünnecke graph of level 2. Let

$$V(G^*) = V_0^* \cup V_1^* \cup V_2^*.$$

Since r_i, s_i, t_i are successive vertices on the path π_i for $i = 1, \ldots, q$, and r^*, s_1, t_1 are successive vertices on the path π^*, it follows that

$$\begin{aligned} V_0^* &= S_q^* \\ V_1^* &\supseteq S_q \\ V_2^* &= S_q^+. \end{aligned}$$

Suppose that $V_1^* \neq S_q$. Then there exists a vertex $s' \in V_1^* \setminus S_q$, and so $(r', s') \in E(G)$ and $(s', t') \in E(G)$ for some $r' \in S_q^* = \{r^*\} \cup S_q^-$ and $t_l \in S_q^+$. The path π_l has a terminal segment that starts at t_l, ends in V_h, and does not intersect S. If $r' = r_{l'} \in S_q^-$, then the path $\pi_{l'}$ has an initial segment that starts in V_0, ends at r', and does not intersect S. Combining this initial segment with the path r', s', t', and then with the terminal segment of π_l from r_l into V_h, we obtain a path from V_0 into V_h that does not intersect the separating set S, which is impossible. If $r' = r^*$, then the path π^* passes through r', and does not intersect $\{s_{q+1}, \ldots, s_m\}$, which are the elements of the separating set S that do not belong to V_j. The initial segment of the path π^* from V_0 to r', followed by the segment r', s', t', and then the terminal segment of the path π_l, is again a path from V_0 into V_h that does not intersect the separating set S, which is impossible. Therefore,

$$V_1^* = S_q.$$

Since (r_i, s_i) and (s_i, t_i) are edges in the contracted graph G^*, it follows from Lemma 7.1 that

$$d^+(r_i, G^*) \geq d^+(s_i, G^*)$$

and

$$d^-(t_i, G^*) \geq d^-(s_i, G^*).$$

Since the number of edges leaving V_0^* is exactly equal to the number of edges going into V_1^*, and the number of edges leaving V_1^* is exactly equal to the number of edges going into V_2^*,

it follows that

$$\begin{aligned}\sum_{i=1}^{q} d^+(r_i, G^*) &\geq \sum_{i=1}^{q} d^+(s_i, G^*) \\ &= \sum_{i=1}^{q} d^-(t_i, G^*) \\ &\geq \sum_{i=1}^{q} d^-(s_i, G^*) \\ &= d^+(r^*, G^*) + d^+(r_i, G^*) \\ &\geq 1 + \sum_{i=1}^{q} d^+(r_i, G^*).\end{aligned}$$

This contradiction proves that $S \subseteq V_0 \cup V_h$.

Since $|S|$ is the number of pairwise totally disjoint paths from V_0 to V_h, it is clear that

$$|S| \leq |V_0|.$$

If $V_0 \subseteq S$, then $|V_0| \leq |S|$, and so $|V_0| = |S|$, which is what we needed to prove.

If $V_0 \not\subseteq S$, then $V_0 \setminus S \neq \emptyset$, and, since S is a separating set, every path in G from $V_0 \setminus S$ must end in $V_h \cap S$. It follows that

$$1 \leq D_h \leq \frac{|\mathrm{im}(V_0 \setminus S, V_h|}{|V_0 \setminus S|} \leq \frac{|V_h \cap S|}{|V_0 \setminus S|}$$

and so

$$|V_0 \setminus S| \leq |V_h \cap S|.$$

Therefore,

$$\begin{aligned}|S| &= |V_0 \cap S| + |V_h \cap S| \\ &\geq |V_0 \cap S| + |V_0 \setminus S| \\ &= |V_0|,\end{aligned}$$

and so $|V_0| = |S|$. This completes the proof.

Lemma 7.3 *Let G be a Plünnecke graph of level h. If $D_h \geq 1$, then $D_i \geq 1$ for $i = 1, \ldots, h$.*

Proof. Since $D_h \geq 1$, it follows from the previous Lemma that there are $|V_0|$ totally pairwise disjoint paths from V_0 to V_h, and each of these paths contains a vertex from the set V_i. The vertices in V_i corresponding to distinct paths are distinct. Let Z be a nonempty subset of V_0. Since there are $|Z|$ pairwise totally disjoint paths emanating from Z, it follows that

$$|Z| \leq |\mathrm{im}(Z, V_i)|$$

and so

$$D_i = D(V_0, V_i) = \min\{|\mathrm{im}(Z, V_i)|/|Z|\} \geq 1.$$

This completes the proof.

Theorem 7.4 (Plünnecke) *Let G be a Plünnecke graph of level h, and let $D_1, \ldots, D_h$ be the magnification ratios of G. Then*

$$D_h^{1/h} \le D_{h-1}^{1/(h-1)} \le \cdots \le D_2^{1/2} \le D_1$$

Proof. It suffices to prove that $D_i \ge D_h^{i/h}$ for $i = 1, \ldots, h$. If $D_h = 0$, the result is trivial. If $D_h = 1$, then Lemma 7.3 implies that $D_i \ge 1 = D_h^{i/h}$.

Let $A = \{0\}$ and let B be a set of n integers such that the addition graph of level h constructed from A and B is the independent addition graph $I_{n,h}$. Then

$$n^i/i! \le D_i(I_{n,h}) = D(\{0\}, iB) = |iB| \le n^i$$

for $i = 1, \ldots, h$.

Suppose that $0 < D_h < 1$. Let r and n be positive integers such that $D_h^r n^h \ge h!$. Let G^r be the product of r copies of the graph G, and consider the product graph $G^r \times I_{n,h}$. This is a Plünnecke graph of level h, and, because of the multiplication of magnification ratios,

$$D_h(G^r \times I_{n,h}) = D_h(G)^r D_i(I_{n,h}) \ge D_h(G)^r n^h/h! \ge 1.$$

It follows from Lemma 7.3 that

$$1 \le D_i(G^r \times I_{n,h}) = D_i(G)^r D_i(I_{n,h}) \le D_i(G)^r n^i$$

and so

$$D_i(G) \ge n^{-i/r}.$$

For $r \ge 1$, let

$$n = \left\lfloor 1 + (h! D_h(G)^{-r})^{1/h} \right\rfloor.$$

Then

$$(h! D_h(G)^{-r})^{1/h} < n \le 1 + (h! D_h(G)^{-r})^{1/h} \le 2(h! D_h(G)^{-r})^{1/h},$$

where the last inequality follows from the fact that $D_h < 1$. Since $D_h(G)^r n^h > h!$, it follows that

$$\begin{aligned} D_i(G) &\ge n^{-i/r} \\ &\ge \left(2(h! D_h(G)^{-r})^{1/h}\right)^{-i/r} \\ &= \left(2(h!)^{1/h}\right)^{-i/r} D_h(G)^{i/h}. \end{aligned}$$

Since this inequality holds for all $r \ge 1$, and

$$\lim_{r \to \infty} \left(2(h!)^{1/h}\right)^{-i/r} = 1,$$

it follows that $D_i(G) \ge D_h(G)^{i/h}$ for $i = 1, \ldots, h$.

Finally, we consider the case $D_h > 1$. For $r \geq 1$, let $n = [D_h(G)^{r/h}]$. Then

$$2 \leq n \leq D_h(G)^{r/h} < n + 1 < 2n$$

and so $D_h(G)^r n^{-h} \geq 1$.

Let $I_{n,h}^{-1}$ be the inverse of the independent addition graph. Its magnification ratios satisfy the conditions

$$D_h(I_{n,h}^{-1}) = |hB|^{-1} = \binom{n+h-1}{h-1}^{-1} \geq n^{-h}$$

and

$$D_i(I_{n,h}^{-1}) \leq \frac{|(h-i)B|}{|hB|} \leq \frac{n^{h-i}}{n^h/h!} = h!n^{-i}.$$

Consider the $(r+1)$-fold product graph $G^r \times I_{n,h}$. It follows from the multiplicativity of magnification ratios that

$$D_h(G^r \times I_{n,h}^{-1}) = D_h(G)^r D_h(I_{n,h}) \geq D_h(G)^r n^{-h} \geq 1$$

and so, by Lemma 7.3,

$$1 \leq D_i(G^r \times I_{n,h}^{-1}) \leq D_i(G)^r h! n^{-i}.$$

This implies that

$$D_i(G) \geq (h!)^{-1/r} n^{i/r} > (h!)^{-1/r} (D_h(G)^{r/h}/2)^{i/r} = (2^i h!)^{-1/r} D_h(G)^{i/h}.$$

Since this inequality holds for all $r \geq 1$, it follows that $D_i(G) \geq D_h(G)^{i/h}$ for $i = 1, \ldots, h$. This completes the proof of Plünnecke's inequality.

7.6 Applications to sums of finite sets

Theorem 7.5 *Let B be a finite subset of an abelian group. Then*

$$|hB| \leq |jB|^{h/j}$$

for $j = 1, \ldots, h$.

Proof. Let G be the addition graph of level h constructed from the sets $A = \{0\}$ and B. Since $D_j = |jB|$ for $j = 1, \ldots, h$, it follows from Plünnecke's inequality that

$$|hB| = D_h \leq D_j^{h/j} = |jB|^{h/j}.$$

Theorem 7.6 *Let A and B be finite subsets of an abelian group, and let $1 \leq j \leq h$. If $|A| = n$ and $|A + jB| \leq cn$, then there is a nonempty subset $A' \subseteq A$ such that*

$$|A' + hB| \leq c^{h/j} |A'|.$$

Proof. Let G be the addition graph of level h constructed from the sets A and B. There exists a nonempty set $A' \subseteq A$ such that

$$D_h = \frac{|A' + hB|}{|A'|}.$$

Then $D_h \leq D_j^{h/j}$ implies that

$$\begin{aligned} |A' + hB| &= |A'|D_h \\ &\leq |A'|D_j^{h/j} \\ &\leq |A'| \left(\frac{|A + jB|}{|A|} \right)^{h/j} \\ &\leq c^{h/j}|A'|. \end{aligned}$$

Theorem 7.7 *Let B be a finite subset of an abelian group, and let $h \geq 2$. If $|B| = n$ and $|2B| \leq cn$, then*

$$|hB| \leq c^h n.$$

Proof. We apply Theorem 7.6 with $A = B$ and $j = 1$. Then there is a nonempty set $B' \subseteq B$ such that

$$|hB| \leq |B' + hB| \leq c^h|B'| \leq c^h|B| = c^h n.$$

Lemma 7.4 *Let U, V, and W be finite subsets of an abelian group. Then*

$$|U||V - W| \leq |U + V||U + W|.$$

Proof. Let $|V| = k$ and $V = \{v_1, \ldots, v_k\}$. If $d \in V - W$, then $d = v_i - w$ for some $v_i \in V$ and $w \in W$. Let $i(d)$ be the least integer i such that $v_i - d \in W$. Define $v(d) \in V$ and $w(d) \in W$ by $v(d) = v_{i(d)}$ and $w(d) = v(d) - d$. Thus, to each $d \in V - W$ there is a unique representation of the form $d = v(d) - w(d)$.

Define the function

$$\phi : U \times (V - W) \to (U + V) \times (U + W)$$

by

$$\phi(u, d) = (u + v(d), u + w(d)).$$

We shall prove that the map ϕ is one-to-one. If

$$\phi(u_1, d_1) = \phi(u_2, d_2),$$

then

$$(u_1 + v(d_1), u_1 + w(d_1)) = (u_2 + v(d_2), u_2 + w(d_2))$$

and so

$$u_1 + v(d_1) = u_2 + v(d_2)$$

and

$$u_1 + w(d_1) = u_2 + w(d_2).$$

This implies that

$$d_1 = v(d_1) - w(d_1) = v(d_2) - w(d_2) = d_2.$$

Since each $d \in V - W$ uniquely determines $v(d) \in V$, it follows that $v(d_1) = v(d_2)$ and so $u_1 = u_2$. Therefore, ϕ is a one-to-one mapping of finite sets, and so

$$|U||V - W| = |U \times (V - W)| \leq |(U + V) \times (U + W)| = |U + V||U + W|.$$

This completes the proof.

Theorem 7.8 *Let A and B be finite subsets of an abelian group such that $|A| = n$ and $|A + B| \leq cn$. Let $k \geq 1$ and $l \geq 1$. Then*

$$|kB - lB| \leq c^{k+l}n.$$

Proof. Since $|kB - lB| = |lB - kB|$, we can assume without loss of generality that $k \leq l$. By Theorem 7.6 there exists a nonempty set $A' \subseteq A$ such that

$$|A' + kB| \leq c^k|A'| = c'n' \tag{7.1}$$

where $c' = c^k$ and $n' = |A'|$. Applying Theorem 7.6 with A, j, h, n, and c replaced by A', k, l, n', and c', respectively, we obtain from inequality (7.1) that there exists a nonempty subset $A'' \subseteq A'$ such that

$$|A'' + lB| \leq (c')^{l/k}|A''| = c^l|A''|. \tag{7.2}$$

It follows from Lemma 7.4 with $U = A''$, $V = kB$, and $W = lB$, and from inequalities (7.1) and (7.2) that

$$\begin{aligned} |A''||kB - lB| &\leq |A'' + kB||A'' + lB| \\ &\leq |A'' + kB||A' + lB| \\ &\leq c^k|A''|c^l|A'| \\ &\leq c^{k+l}n|A''|. \end{aligned}$$

We complete the proof by dividing this inequality by $|A''|$.

7.7 Essential components

For any set A of integers, let $A(m,n)$ denote the number of elements $a \in A$ such that $m < a \leq n$. In particular, the function $A(n) = A(0,n)$ counts the number of positive elements of A not exceeding n. The *Shnirel'man density* of the set A is defined by

$$\sigma(A) = \inf_{n=1,2,3,\ldots} \frac{A(n)}{n}.$$

Then $0 \leq \sigma(A) \leq 1$ for every set A, and $\sigma(A) > 0$ implies that $1 \in A$.

The set B is called an *essential component* if $\sigma(A+B) > \sigma(A)$ for every set A with $0 < \sigma(A) < 1$. Let A and B be sets of integers with $1 \in A$ and $0 \in B$. Shnirel'man [70] proved the fundamental inequality

$$\sigma(A+B) \geq \sigma(A) + \sigma(B)(1 - \sigma(A)).$$

If $\sigma(A) < 1$ and $\sigma(B) > 0$, then $\sigma(A+B) > \sigma(A)$. Shnirel'man's inequality implies that any set B with $0 \in B$ and positive Shnirel'man density is an essential component.

There exist sets of density zero that are also essential components. Khinchin [37] proved that the set of squares form an essential component. The squares are a basis of order 4. Erdős [22] generalized Khinchin's result by proving that if B is any basis of order h and if A is any set of integers such that $0 < \sigma(A) < 1$, then

$$\sigma(A+B) \geq \sigma(A) + (2h)^{-1}\sigma(A)(1 - \sigma(A)) > \sigma(A).$$

Plünnecke applied his graph-theoretic method to obtain a considerable improvement of Erdős's theorem. Plünnecke's proof relies heavily on properties of the *impact function*, or *Wirkungsfunktion*, defined for $0 \leq \xi \leq 1$ and for any set B of nonnegative integers by

$$\phi(\xi, B) = \inf\{\sigma(A+B) : A \subseteq \mathbf{N}_0, \sigma(A) \geq \xi\}.$$

Lemma 7.5 *Let $0 \leq \xi \leq 1$ and let B be a set of nonnegative integers. Then*

$$\begin{aligned} \phi(\xi, B) &= \inf\{(A+B)(n)/n : A \subseteq \mathbf{N}_0, \sigma(A) \geq \xi \\ &\qquad \text{and } A(n)/n \leq A(m)/m \text{ for } m = 1, \ldots, n\}. \end{aligned}$$

Proof. If $\sigma(A) \geq \xi$ and $n \geq 1$, then

$$\phi(\xi, B) \leq \sigma(A+B) \leq (A+B)(n)/n.$$

It suffices to show that for any $\varepsilon > 0$ there exist a set A and an integer n satisfying the conditions of the lemma such that

$$\phi(\xi, B) \leq (A+B)(n)/n < \phi(\xi, B) + \varepsilon.$$

It follows from the definition of the impact function that there exists a set A of nonnegative integers such that $\sigma(A) \geq \xi$ and

$$\phi(\xi, B) \leq \sigma(A + B) < \phi(\xi, B) + \varepsilon/2.$$

It follows from the definition of Shnirel'man density that there exists an integer $n \geq 1$ such that

$$\sigma(A + B) \leq (A + B)(n)/n < \sigma(A + B) + \varepsilon/2.$$

Therefore,

$$\phi(\xi, B) \leq (A + B)(n)/n < \phi(\xi, B) + \varepsilon. \tag{7.3}$$

Let n be the smallest positive integer such that inequality 7.3 is satisfied for some set A with $\sigma(A) \geq \xi$. For these choices of n and A, we have

$$(A + B)(n)/n < \phi(\xi, B) + \varepsilon \leq (A + B)(m)/m,$$

for $m = 1, \ldots, n - 1$, and so

$$(A + B)(m, n) = (A + B)(n) - (A + B)(m) < (n - m)(\phi(\xi, B) + \varepsilon).$$

We shall show that

$$A(n)/n \leq A(m)/m$$

for $m = 1, \ldots, n$. If not, there exists an integer l such that $1 \leq l \leq n - 1$ and

$$A(l)/l = \min\{A(m)/m : m = 1, \ldots, n - 1\} < A(n)/n.$$

Then

$$\begin{aligned} A(l, l + i) &= A(l + i) - A(l) \\ &\geq (l + i)A(l)/l - A(l) = iA(l)/l \\ &\geq i\sigma(A) \geq i\xi \end{aligned}$$

for $i = 1, \ldots, n - l$. Let

$$A^* = \{a - l : a \in A, l < a \leq n\} \cup \{m \in \mathbf{N}_0 : m > n - l\}.$$

For $i = 1, \ldots, n - l$,

$$A^*(i) = A(l, l + i) \geq i\xi.$$

For $i > n - l$, since A^* contains all integers $m > n - l$,

$$A^*(i) = A^*(n - l) + (i - n + l) \geq (n - l)\xi + i - n + l \geq i\xi.$$

Therefore,

$$A^*(i) \geq i\xi$$

for all $i \geq 1$, and so $\sigma(A^*) \geq \xi$.

Moreover,

$$\begin{aligned}(A^* + B)(n-l) &= |\{a^* \in A^*, b \in B : 0 < a^* + b \leq n - l\}| \\ &= |\{a^* \in A^*, b \in B : l < a^* + l + b \leq n\}| \\ &= |\{a \in A, b \in B : l < a + b \leq n\}| \\ &= (A + B)(l, n) \\ &< (n-l)(\phi(\xi, B) + \varepsilon).\end{aligned}$$

Since $1 \leq n - l < n$, this contradicts the minimality of n, and so $A(n)/n \leq A(m)/m$ for all $m = 1, \ldots, n$. This completes the proof.

Lemma 7.6 *Let B be a set of nonnegative integers with $0 \in B$, let $0 < \xi \leq 1$, and let $h \geq 1$. Then*

$$\phi(\xi, jB) \geq \xi^{1-j/h}\sigma(hB + 1)^{j/h}$$

for $j = 1, \ldots, h$.

Proof. Let A be a set of nonnegative integers with $\sigma(A) \geq \xi > 0$. Then $1 \in A$. Suppose that

$$\gamma = A(n)/n \leq A(m)/m$$

for $m = 1, \ldots, n$. Then $\xi \leq \sigma(A) \leq \gamma$. Let G be the Plünnecke graph obtained by truncating at n the addition graph of order h generated by A and B. The i-th vertex set of the truncated addition graph is

$$V_i = (A + iB) \cap [1, n].$$

Since $1 \in A$ and $0 \in B$, it follows that $V_i \neq \emptyset$ for $i = 0, 1, \ldots, h$. Let

$$\emptyset \neq Z \subseteq V_0 = A \cap [1, n].$$

Then

$$\begin{aligned}Z(m, n) &\leq A(m, n) = A(n) - A(m) \\ &= \gamma n - A(m) \leq \gamma(n - m)\end{aligned}$$

for $m = 0, 1, \ldots, n$, and

$$|\mathrm{im}(Z, V_k)| = (Z + hB)(n)$$

in the graaph G. Let z be the smallest positive element of Z. Then

$$\begin{aligned}(Z + hB)(n) &\geq (\{z\} + hB)(n) \\ &= |\{b' \in hB : z \leq z + b' \leq n\}| \\ &= |\{b' \in hB : 1 \leq 1 + b' \leq n - z + 1\} \\ &= (hB + 1)(n - z + 1) \\ &\geq (n - z + 1)\sigma(hB + 1) \\ &\geq \gamma^{-1} Z(z - 1, n)\sigma(hB + 1) \\ &\geq \gamma^{-1}|Z|\sigma(hB + 1).\end{aligned}$$

Thus,

$$(Z + hB)(n)/|Z| \geq \gamma^{-1}\sigma(hB + 1)$$

for all nonempty subsets Z of V_0, and so

$$D_h(G) \geq \gamma^{-1}\sigma(hB + 1).$$

Plünnecke's inequality implies that for $j = 1, \ldots, h$,

$$\begin{aligned} \gamma^{-1}\sigma(hB + 1) &\leq D_h(G) \leq D_j(G)^{h/j} \\ &\leq ((A + jB)(n)/A(n))^{h/j}, \end{aligned}$$

and so

$$\begin{aligned} (A + jB)(n) &\geq A(n)(\gamma^{-1}\sigma(hB + 1))^{j/h} \\ &= n\gamma(\gamma^{-1}\sigma(hB + 1))^{j/h} \\ &= n\gamma^{1-j/h}\sigma(hB + 1)^{j/h} \\ &\geq n\xi^{1-j/h}\sigma(hB + 1)^{j/h}. \end{aligned}$$

It follows from Lemma 7.5 that

$$\phi(\xi, jB) \geq \xi^{1-j/h}\sigma(hB + 1)^{j/h}.$$

This completes the proof.

Theorem 7.9 (Plünnecke) *Let A be a set of nonnegative integers with $0 < \sigma(A) < 1$, and let B be a basis of order $h \geq 2$. Then*

$$\sigma(A + B) \geq \sigma(A)^{1-1/h}.$$

Proof. Since B is a basis of order $h \geq 2$, $1 \in hB$ and so $0 \in B$. Therefore, $hB + 1 = \mathbf{N}$, and $\sigma(hB + 1) = 1$. Applying the previous Lemma with $j = 1$ gives

$$\phi(\xi, B) \geq \xi^{1-1/h}$$

for all $\xi \in [0, 1]$. Let $\xi = \sigma(A)$. Then

$$\sigma(A + B) \geq \phi(\xi, B) \geq \xi^{1-1/h} = \sigma(A)^{1-1/h}.$$

This completes the proof.

Corollary 7.2 (Erdős) *Let A be a set of nonnegative integers with $0 < \sigma(A) < 1$, and let B be a basis of order $h \geq 2$. Then*

$$\sigma(A + B) \geq \sigma(A) + (1/h)\sigma(A)(1 - \sigma(A)).$$

Remark: Erdős actually obtained a slightly weaker result in which the constant $1/h$ is replaced by $1/2h$.

Proof. It suffices to prove that

$$x + x(1-x)/h \leq x^{1-1/h}$$

for $0 \leq x \leq 1$, or, equivalently, that the function satisfies the inequality

$$f(x) = x^{1/h}(1 + (1-x)/h) \leq 1$$

for $0 \leq x \leq 1$. Since $f(0) = 0$ and $f(1) = 1$, it is enough to show that $f(x)$ is increasing on the unit interval, and this follows immediately from the fact that

$$\begin{aligned} f'(x) &= (1/h)x^{1/h-1}(1 + (1-x)/h) - (1/h)x^{1/h} \\ &= (1/h)x^{1/h-1}(1 + (1-x)/h) - x) \\ &= (1/h)x^{1/h-1}(1 + 1/h)(1-x) \\ &> 0 \end{aligned}$$

for $0 < x < 1$. This completes the proof.

7.8 Notes

Plünnecke's work was published in the monograph [59] and the papers [57, 58, 60]. The proof of Plünnecke's inequality presented in this chapter is from Ruzsa [65], which also contains the applications to sums of finite sets. The proof of Menger's theorem is due to Dirac [19] The *Wirkungsfunktion* was introduced by Stöhr and Wirsing [72], and studied by Plünnecke in great detail in [59]. The proof of Plünnecke's theorem on essential components follows closely the paper of Malouf [44].

7.9 Exercises

1. Draw the addition graphs of level 3 determined by the following sets of integers:

   (a) $A = \{0\}, B = \{0, 1, 2\}$

   (b) $A = \{0\}, B = \{0, 1, 3\}$

   (c) $A = \{0, 1\}, B = \{0, 1, 4\}$

2. Let A and B be finite, nonempty subsets of an abelian group, and let G be the addition graph of level h determined by A and B. Prove that $d^+(v) = 0$ for all $v \in V_h$ and that $d^+(v) = |B|$ for all $v \in V(G) \setminus V_h$.

3. Let G be an addition graph of level h. Prove directly that $D_i \geq 1$ for $i = 1, \ldots, h$.

4. Prove that the inverse of a Plünnecke graph is a Plünnecke graph.

5. Prove that the product of Plünnecke graphs is a Plünnecke graph.

6. Let B be a finite subset of an abelian group, and let $2 \leq j \leq h$. If $|B| = n$ and $|jB| \leq cn$, prove that
$$|hB| \leq c^{h/(j-1)}n.$$

7. Let B be a finite subset of an abelian group, and let $2 \leq j \leq h$. If $|B| = n$ and $|jB| \leq cn^{1+\beta}$, prove that
$$|hB| \leq c^{h/(j-1)}n^{1+\beta h/(j-1)}.$$

8. Let B be a finite subset of an abelian group. Let $1 \leq j \leq h$. Suppose that j divides h. Prove directly (without using Plünnecke's inequality) that $|hB| \leq |jB|^{h/j}$. What inequality can you obtain (without using Plünnecke's inequality) if j does not divide h?

Chapter 8

Freiman's theorem

8.1 Introduction

A simple inverse theorem in additive number theory states that if A is a finite set of integers whose 2-fold sumset is "very small" in the sense that $|2A| \leq 3|A| - 4$, then A is a "large subset of an arithmetic progression. Freiman discovered a deep generalization of this result. His theorem states that if A is a finite set of integers such that $2A$ is small, then A is close to an arithmetic progression in the sense that A is a "dense" subset of a multi-dimensional arithmetic progression.

Freiman's inverse theorem can be stated as follows.

Theorem 8.1 (Freiman) *Let A be a finite set of integers such that $|2A| \leq c|A|$. There exist integers $q_0, q_1, \ldots, q_n, l_1, \ldots, l_n$ such that*

$$A \subseteq Q = \{q_0 + x_1 q_1 + \cdots + x_n q_n : 0 \leq x_i < l_i \text{ for } i = 1, \ldots, n\},$$

where $|Q| \leq c'|A|$, and n and c' depend only on c.

The object of this chapter is to present a beautiful proof due to I. Z. Ruzsa of a generalization of Freiman's theorem.

8.2 Multi-dimensional progressions

Let $q_0, q_1, \ldots, q_n$ be elements of an abelian group G, and let $l_1, \ldots, l_n$ be positive integers. The set

$$\begin{aligned} Q &= Q(q_0, q_1, \ldots, q_n; l_1, \ldots, l_n) \\ &= \{q_0 + x_1 q_1 + \cdots + x_n q_n : 0 \leq x_i < l_i \text{ for } i = 1, \ldots, n\} \end{aligned}$$

is called a *n-dimensional arithmetic progression* in the group G. The *length* of Q is $l_1 \cdots l_n$. Clearly, $|Q| \leq l_1 \cdots l_n$, and Q is called *proper* if $|Q| = l_1 \cdots l_n$.

The representation of a set as a multi-dimensional arithmetic progression is not unique. A set can have more than one such representation, and these representations can have different dimensions and different lengths.

Theorem 8.2 *Let P and P' be multi-dimensional arithmetic progressions of dimensions n and n' and lengths l and l', respectively. Then*

1. *$P + P'$ is a multi-dimensional arithmetic progression of dimension $n + n'$ and length ll'.*
2. *$P - P$ is an arithmetic progression of dimension n and length $|P - P| < 2^n l$.*
3. *Every finite subset F of a group is a subset of an arithmetic progression of dimension $|F|$ and length $2^{|F|}$.*
4. *If P is proper, then $|hP| < h^n |P|$.*

Proof. Let Q be the progression $Q(q_0, q_1, \ldots, q_n; l_1, \ldots, l_n)$, and let Q' be the progression $Q(q'_0, q'_1, \ldots, q'_n; l'_1, \ldots, l'_n)$.

1. This follows from the representation

$$\begin{aligned} Q + Q' &= \{q_0 + q'_0 + x_1 q_1 + \cdots + x_n q_n + x'_1 q'_1 + \cdots + x'_n q'_n : \\ &\quad 0 \le x_i < l_i, 0 \le x'_i < l'_i\}. \end{aligned}$$

2. This follows immediately from the observation that

$$Q - Q = \{x_1 q_1 + \cdots + x_n q_n : -l_i < x_i < l_i \text{ for } i = 1, \ldots, d\}.$$

3. If $F = \{f_1, \ldots, f_s\}$, then

$$F \subseteq Q = \{x_1 f_1 + \cdots + x_s f_s | 0 \le x_i < 2 \text{ for } i = 1, \ldots, s\}.$$

4. Since hQ can be represented in the form

$$hQ = Q(hq_0, q_1, \ldots, q_n; h(l_1 - 1) + 1, \ldots, h(l_n - 1) + 1).$$

it follows that, if Q is proper, then

$$\begin{aligned} |hQ| &\le \prod_{i=1}^{d} (h(l_i - 1) + 1) \\ &< \prod_{i=1}^{d} h l_i \\ &= h^n |Q|. \end{aligned}$$

8.3 Freiman isomorphisms

Freiman introduced the important idea of a "local" isomorphism in additive number theory.

Let G and H be abelian groups, and let $A \subseteq G$ and $B \subseteq H$. Let $h \geq 2$. The map $\phi : A \to B$ is called a *Freiman homomorphism of order* h if

$$\phi(a_1) + \cdots + \phi(a_h) = \phi(a'_1) + \cdots + \phi(a'_h)$$

for all $a_1, \ldots, a_h, a'_1, \ldots, a'_h \in A$ such that

$$a_1 + \cdots + a_h = a'_1 + \cdots + a'_h.$$

In this case, ϕ induces a well-defined map $\phi^{(h)} : hA \to hB$.

For example, let $A = [0, k-1] = \{0, 1, \ldots, k-1\}$, and let $B = \{q_0 + xq_1 : 0 \leq x < k\}$. Then the map $\phi(x) = q_0 + xq_1$ is a Freiman isomorphism of order h for all $h \geq 2$.

Let $A = \{(0,0), (1,0), (0,1)\} \subseteq \mathbf{Z}^2\}$ and let $B = \{0, 1, 3\} \subseteq \mathbf{Z}$. Define $\phi : A \to B$ by $\phi(0,0) = 0$, $\phi(1,0) = 1$, and $\phi(0,1) = 3$. The ϕ is a Freiman isomorphism of order 2, but not of order 3.

If $\phi : A \to B$ is a one-to-one correspondence such that

$$a_1 + \cdots + a_h = a'_1 + \cdots + a'_h$$

if and only if

$$\phi(a_1) + \cdots + \phi(a_h) = \phi(a'_1) + \cdots + \phi(a'_h),$$

then ϕ is a *Freiman isomorphism of order* h, and the induced map $\phi^{(h)} : hA \to hB$ is also a one-to-one correspondence. A Freiman isomorphism of order 2 will be called simply a *Freiman isomorphism.*

Let $\phi : A \to B$ be a Freiman isomorphism of order h. Then ϕ is also a Freiman isomorphism of order h' for every $h' \leq h$. If $A' \subseteq A$ and $B' = \phi(A')$, then the map $\phi : A' \to B'$ is also a Freiman isomorphism of order h.

If $f : G \to H$ is a group homomorphism, then f is a Freiman homorphism of order h for all $h \geq 2$, and $f^{(h)} = f$. If f is a group isomorphism, then f is a Freiman isomorphism of order h for all $h \geq 2$.

If $\phi : G \to H$ is an affine map, that is, a map of the form $\phi(g) = g_0 + f(g)$, where $g_0 \in G$ and $f : G \to H$ is a group homomorphism (resp. isomorphism), then ϕ is a Freiman homomorphism (resp. isomorphism) of order h for all $h \geq 2$, and $\phi^h(g) = hg_0 + f(g)$.

For example, let $q_1 \neq 0$, and let $P = \{q_0 + xq_1 : 0 \leq x < l\}$ be a 1-dimensional arithmetic progression in the group G. Define $\phi : [0, l-1] \to P$ by $\phi(x) = q_0 + xq_1$. Then ϕ is the restriction of an affine map from $\mathbf{Z}$ into G, and is a Freiman isomorphism of order h for all $h \geq 2$.

Let $l_1, \ldots, l_n$ be positive integers, and let $\mathbf{P}$ be the fundamental parallelepiped of the lattice generated by the vectors $l_1\mathbf{e}_1, \ldots, l_n\mathbf{e}_d$. Let $I(\mathbf{P})$ be the integer parallelepiped defined by

$$I(\mathbf{P}) = \mathbf{P} \cap \mathbf{Z}^d = \{(x_1, \ldots, x_n) \in \mathbf{Z}^d : 0 \leq x_1 < l_i \text{ for } i = 1, \ldots, d\}.$$

Then

$$hI(\mathbf{P}) = \{(x_1, \ldots, x_n) \in \mathbf{Z}^d : 0 \leq x_1 < h(l_i - 1) + 1 \text{ for } i = 1, \ldots, d\},$$

and so

$$|hI(\mathbf{P})| \leq \prod_{i=1}^{d} (h(l_i - 1) + 1) < h^n |I(\mathbf{P})|.$$

Let $q_0, q_1, \ldots, q_n \in \mathbf{Z}$, and let P be the n-dimensional arithmetic progression $P = P(q_0, q_1, \ldots, q_n; l_1, \ldots, l_n)$. Define the map $\phi : I(\mathbf{P}) \to P$ by

$$\phi(x_1, \ldots, x_n) = q_0 + x_1 q_1 + \cdots + x_n q_n. \tag{8.1}$$

Since ϕ is the restriction of an affine map from $\mathbf{Z}^d$ into $\mathbf{Z}$, it follows that ϕ is a Freiman homomorphism of order h for all $h \geq 2$.

Theorem 8.3 *Let $h \geq 2$, and let $I(\mathbf{P})$ be an integer parallelepiped of dimension d. Then there exists a d-dimensional arithmetic progression P such that $I(\mathbf{P})$ and P are Freiman isomorphic of order h.*

Proof. We shall construct integers $q_0, q_1, \ldots, q_n$ so that the affine map $\phi : I(\mathbf{P}) \to P$ defined by (8.1) is a Freiman isomorphism of order h. Let q_0 be any integer, and choose positive integers $q_1, \ldots, q_n$ so that

$$\sum_{j=1}^{m-1} h(l_j - 1) q_j < q_m \tag{8.2}$$

for $m = 2, \ldots, d$. Let $\mathbf{x}_i = (x_{i,1}, \ldots, x_{i,d}) \in I(\mathbf{P})$ and $\mathbf{y}_i = (y_{i,1}, \ldots, y_{i,d}) \in I(\mathbf{P})$ for $i = i, \ldots, h$, and suppose that

$$\phi(\mathbf{x}_1) + \cdots + \phi(\mathbf{x}_h) = \phi(\mathbf{y}_1) + \cdots + \phi(\mathbf{y}_h).$$

This implies that

$$hq_0 + \sum_{j=1}^{d} \sum_{i=1}^{h} x_{i,j} q_j = hq_0 + \sum_{j=1}^{d} \sum_{i=1}^{h} y_{i,j} q_j.$$

Let

$$w_j = \sum_{i=1}^{h} x_{i,j} - \sum_{i=1}^{h} y_{i,j}.$$

Then

$$|w_j| \leq h(l_j - 1),$$

and

$$\sum_{j=1}^{n} w_j q_j = 0.$$

Suppose that $w_j \neq 0$ for some j, and let m be the greatest integer for which $w_m \neq 0$. Then

$$-w_m q_m = \sum_{j=1}^{m-1} w_j q_j,$$

and so

$$\begin{aligned} |w_m q_m| &= |\sum_{j=1}^{m-1} w_j q_j| \\ &\leq \sum_{j=1}^{m-1} (h(l_j - 1)q_j \\ &< q_m \\ &\leq |w_m q_m|, \end{aligned}$$

which is impossible. Thus, $w_j = 0$ for all j, which implies that

$$\mathbf{x}_1 + \cdots + \mathbf{x}_h = \mathbf{y}_1 + \cdots + \mathbf{y}_h.$$

Thus, every d-dimensional integer parallelepiped is Freiman isomorphic of order h to a d-dimensional arithmetic progression.

Corollary 8.1 *Let $h \geq 2$, and let A be a finite set of lattice points. Then A is Freiman isomorphic of order h to a set of integers.*

Proof. The set A is a subset of some n-dimensional integer parallelepiped $I(\mathbf{P})$, which is Freiman isomorphic of order h to an n-dimensional arithmetic progression P. Then A is Freiman isomorphic to its image under this mapping.

Corollary 8.2 *Let $h \geq 2$, and let A be a finite subset of a torsion-free abelian group. Then A is isomorphic of order h to a set of integers.*

Proof. Let G be the group generated by A. Since G is finitely generated, it follows that G is isomorphic to $\mathbf{Z}^n$ for some n, and so there is a Freiman isomorphism of order h between A and some finite set of integer lattice points, which, by the previous Corollary, is Freiman isomorphic of order h to a set of integers.

Theorem 8.4 *Let G and H be abelian groups, and let P be a d-dimensional arithmetic progression contained in G. Let $\phi : P \to Q \subseteq H$ be a Freiman homomorphism of order h. Then $\phi(P)$ is a d-dimensional arithmetic progression in H. If ϕ is a Freiman isomorphism, then P is a proper d-dimensional arithmetic progression if and only if $\phi(P)$ is also a proper d-dimensional arithmetic progression.*

Proof. Let $P = \{q_0, q_1, \ldots, q_n; l_1, \ldots, l_n\}$. We define $q'_0, q'_1, \ldots, q'_n \in H$ by

$$\begin{aligned} q'_0 &= \phi(q_0), \\ q'_i &= \phi(q_0 + q_i) - \phi(q_0) \end{aligned}$$

for $i = 1, \ldots, d$. The set $P' = \{q'_0, q'_1, \ldots, q'_n; l_1, \ldots, l_n\}$ is a d-dimensional arithmetic progression in H. We shall show that $P' = \phi(P)$ and that

$$\phi(q_0 + x_1 q_1 + \cdots + x_n q_n) = q'_0 + x_1 q'_1 + \cdots + x_n q'_n$$

for all $q_0 + x_1 q_1 + \cdots + x_n q_n \in P$.

The proof is by induction on $m = \sum_{i=1}^{d} x_i$. It follows from the definition of $q'_0, q'_1, \ldots, q'_s$ that the statement is true for $m = 0$ and $m = 1$. Assume that the result holds for some $m \geq 1$. Let $a = q_0 + x_1 q_1 + \cdots + x_n q_n \in P$ with $\sum_{i=1}^{d} x_i = m + 1$. Choose j such that $x_j \geq 1$, and let $a' = a - q_j$. By the induction hypothesis for m, we have

$$\phi(a') = q'_0 + x_1 q'_1 + \cdots + x_{j-1} q'_{j-1} + (x_j - 1) q'_j + x_{j+1} q'_{j+1} + \cdots + x_n q'_n + x_n q'_n.$$

Since

$$a + q_0 = a' + (q_0 + q_j)$$

and since a Freiman homomorphism of order $h \geq 2$ is also a Freiman homomorphism of order 2, it follows that

$$\phi(a) + \phi(q_0) = \phi(a') + \phi(q_0 + q_j)$$

and so

$$\begin{aligned} \phi(a) &= \phi(a') + \phi(q_0 + q_j) - \phi(q_0) \\ &= \phi(a') + q'_j \\ &= q'_0 + x_1 q'_1 + \cdots + x_n q'_n. \end{aligned}$$

Thus, the statement holds for all $m \geq 0$, and $\phi(P) = P'$.

If ϕ is a Freiman isomorphism of order h, then $|P| = |\phi(P)|$, and so $\phi(P)$ is proper if and only if P is proper. This completes the proof.

Theorem 8.5 *Let $h' = h(k + l)$, where h, k, and l are positive integers. Let G and H be abelian groups, and let $A \subseteq G$ and $B \subseteq H$ be finite sets that are Freiman isomorphic of order h'. Then the difference sets $kA - lA$ and $kB - lB$ are Freiman isomorphic of order h.*

Proof. Let $\phi : A \to B$ be a Freiman isomorphism of order h'. The the maps $\phi^{(k)} : kA \to kB$ and $\phi^{(l)} : lA \to lB$ induced by ϕ are Freiman isomorphisms of orders k and l, respectively. Let $d = u - v \in kA - lA$, where $u \in kA$ and $v \in lA$. Define $\phi : kA - lA \to kB - lB$ by $\phi(d) = \phi(u - v) = \phi^{(k)}(u) - \phi^{(l)}(v)$.

Thus,

$$\phi(a_1 + \cdots + a_k - a'_1 - \cdots - a'_l) = \phi(a_1) + \cdots + \phi(a_k) - \phi(a'_1) - \cdots \phi(a'_l).$$

We shall show that ϕ is a Freiman isomorphism of order h. Let $\phi^{(h)} : h(kA - lA) \to h(kB - lB)$ be the map induced by ϕ. Let d_i, $d_i' \in kA - lA$ for $i = 1, \ldots, h$. Then $d_i = u_i - v_i$ and $d_i' = u_i' - v_i'$, where $u_i, u_i' \in kA$ and $v_i, v_i' \in lA$ for $i = 1, \ldots, h$, and so

$$\begin{aligned} \phi^{(h)}(d_1 + \cdots + d_h) &= \phi(d_1) + \cdots + \phi(d_h) \\ &= \phi^{(k)}(u_1) + \cdots + \phi^{(k)}(u_h) - \phi^{(l)}(v_1) - \phi^{(l)}(v_h) \end{aligned}$$

and

$$\begin{aligned} \phi^{(h)}(d_1' + \cdots + d_h') &= \phi(d_1') + \cdots + \phi(d_h') \\ &= \phi^{(k)}(u_1') + \cdots + \phi^{(k)}(u_h') - \phi^{(l)}(v_1') - \phi^{(l)}(v_h'). \end{aligned}$$

It follows that

$$\phi^{(h)}(d_1 + \cdots + d_h) = \phi^{(h)}(d_1' + \cdots + d_h')$$

if and only if

$$\begin{aligned} &\phi^{(k)}(u_1) + \cdots + \phi^{(k)}(u_h) + \phi^{(l)}(v_1') + \phi^{(l)}(v_h') \\ &= \phi^{(k)}(u_1') + \cdots + \phi^{(k)}(u_h') + \phi^{(l)}(v_1) + \phi^{(l)}(v_h). \end{aligned}$$

Since ϕ is a Freiman isomorphism of order $h' = h(k + l)$, the last equation holds if and only if

$$u_1 + \cdots + u_h + v_1' + \cdots + v_h' = u_1' + \cdots + u_h' + v_1 + \cdots + v_h,$$

and this is equivalent to the statement that

$$d_1 + \cdots + d_h = d_1' + \cdots + d_h'.$$

This proves that ϕ is a Freiman isomorphism of order h.

8.4 Ruzsa's proof

Theorem 8.6 *Let A be a finite set of integers. Let $h \geq 2$, and let $D = hA - hA$. For every $m > 4h|D|$ there exists a set $A' \subseteq A$ such that $|A'| \geq |A|/h$ and A' is Freiman isomorphic of order h to a set of residues modulo m.*

Proof. Let $m > 4h|D|$, and let p be a prime number such that

$$p > \max\{m, 2h \max_{a \in A} |a|\}.$$

Let $1 \leq q \leq p - 1$. We shall construct a map $\phi : \mathbf{Z} \to \mathbf{Z}/m\mathbf{Z}$, and then prove that there is a "large" subset $A' \subseteq A \subseteq \mathbf{Z}$ such that ϕ restricted to A' is a Freiman isomorphism of order h. The map ϕ will be the composition of four maps:

$$\mathbf{Z} \stackrel{\alpha}{\to} \mathbf{Z}/p\mathbf{Z} \stackrel{\beta_q}{\to} \mathbf{Z}/p\mathbf{Z} \stackrel{\gamma}{\to} \mathbf{Z} \stackrel{\delta}{\to} \mathbf{Z}/m\mathbf{Z},$$

where α, β_q, γ, and δ are defined below.

Let $\alpha : \mathbf{Z} \to \mathbf{Z}/p\mathbf{Z}$ be the natural map that sends a to $a \pmod{p}$. Since α is a group homomorphism, it is also a Freiman homomorphism of order h. Although α is not a group isomorphism, we can show that α restricted to A is a Freiman isomorphism of order h. Let $a_1, \ldots, a_h, a_1^{'}, \ldots, a_h^{'} \in A$, and suppose that

$$\alpha(a_1) + \cdots + \alpha(a_h) = \alpha(a_1^{'}) + \cdots + \alpha(a_h^{'}).$$

This implies that

$$(a_1 + \cdots + a_h) - (a_1^{'} + \cdots + a_h^{'}) \equiv 0 \pmod{p}.$$

Since

$$|(a_1 + \cdots + a_h) - (a_1^{'} + \cdots + a_h^{'})| \leq 2h \max_{a \in A} |a|\} < p,$$

it follows that

$$(a_1 + \cdots + a_h) - (a_1^{'} + \cdots + a_h^{'}) = 0,$$

and so $\alpha : A \to \alpha(A) \subseteq \mathbf{Z}/p\mathbf{Z}$ is a Freiman isomorphism of order h.

For $1 \leq q \leq p-1$, let $\beta_q : \mathbf{Z}/p\mathbf{Z} \to \mathbf{Z}/p\mathbf{Z}$ be the map that sends $a \pmod{p}$ to $aq \pmod{p}$ for every $a \in \mathbf{Z}/p\mathbf{Z}$. The map β_q is a group isomorphism, and therefore a Freiman isomorphism of order h on any subset of $\mathbf{Z}/p\mathbf{Z}$.

Let $\gamma : \mathbf{Z}/p\mathbf{Z} \to \mathbf{Z}$ be the map that sends the congruence class $a \pmod{p}$ to its least nonnegative representative. The image of γ is $[0, p-1]$. The map γ is neither a group homomorphism nor a Freiman homomorphism of order h. We can, however, write $\mathbf{Z}/p\mathbf{Z}$ as the union of h subsets such that γ restricted to each of these subsets is a Freiman isomorphism of order h. For $i = 1, \ldots, h$, let

$$U_i = \gamma^{-1}[(i-1)(p-1)/h, i(p-1)/h] \subset \mathbf{Z}/p\mathbf{Z}.$$

Since

$$[0, p-1] = \cup_{i=1}^{h} [(i-1)(p-1)/h, i(p-1)/h],$$

it follows that

$$\mathbf{Z}/p\mathbf{Z} = \cup_{i=1}^{h} U_i.$$

Suppose that $u_j, u_j^{'} \in U_i$ for $j = 1, \ldots, h$, and

$$u_1 + \cdots + u_h = u_1^{'} + \cdots + u_h^{'}$$

in $\mathbf{Z}/p\mathbf{Z}$. Then

$$\gamma(u_1) + \cdots + \gamma(u_h) \equiv \gamma(u_1^{'}) + \cdots + \gamma(u_h^{'}) \pmod{p}$$

in $\mathbf{Z}$. Since

$$\gamma(u_1) + \cdots + \gamma(u_h) \in [(i-1)(p-1), i(p-1)]$$

and

$$\gamma(u_1^{'}) + \cdots + \gamma(u_h^{'}) \in [(i-1)(p-1), i(p-1)],$$

it follows that

$$\gamma(u_1) + \cdots + \gamma(u_h) = \gamma(u_1^{'}) + \cdots + \gamma(u_h^{'})$$

and so γ is a Freiman homomorphism of order h. Conversely, if $u_1, \ldots, u_h, u_1', \ldots, u_h' \in \mathbf{Z}/p\mathbf{Z}$ and

$$\gamma(u_1) + \cdots + \gamma(u_h) = \gamma(u_1^{'}) + \cdots + \gamma(u_h^{'})$$

in $\mathbf{Z}$, then

$$u_1 + \cdots + u_h = u_1^{'} + \cdots + u_h^{'}$$

in $\mathbf{Z}/p\mathbf{Z}$, and so γ restricted to each set U_i is a Freiman isomorphism of order h.

Let $A_i = A \cap (\beta_q\alpha)^{-1}(U_i)$ for $i = 1, \ldots, h$. Then $A = \cup_{i=1}^{h} A_i$, and so $|A_i| \geq |A|/h$ for some i. Let $A' = A_i$, and define $\theta_q : A \to \mathbf{Z}$ by $\theta_q = \gamma\beta_q\alpha$. Let

$$V_q' = \theta_q(A') = \gamma\beta_q\alpha(A') \subseteq [0, p-1] \subseteq \mathbf{Z}.$$

Then $\theta_q : A' \to V_q'$ is a Freiman isomorphism of order h.

Let $V_q = \theta_q(A) \subseteq [0, p-1]$. Then

$$\theta_q(a) = aq - \left[\frac{aq}{p}\right] p$$

for all $a \in A$. Let $\delta : \mathbf{Z} \to \mathbf{Z}/m\mathbf{Z}$ be the natural map that sends a to $a \pmod m$. We shall prove that δ restricted to V_q is a Freiman isomorphism of order h for some $q \in [1, p-1]$.

Let $q \in [1, p-1]$, and suppose that $\delta : V_q \to \mathbf{Z}/m\mathbf{Z}$ is not a Freiman isomorphism of order h. This means that there exist integers $v_1, \ldots, v_h, v_1', \ldots, v_h' \in V_q$ such that

$$\delta(v_1) + \cdots + \delta(v_h) = \delta(v_1') + \cdots + \delta(v_h'),$$

but

$$v_1 + \cdots + v_h \neq v_1' + \cdots + v_h'.$$

Define

$$v^* = (v_1 + \cdots + v_h) - (v_1' + \cdots + v_h').$$

Then

$$|v^*| \leq h(p-1) \tag{8.3}$$

and

$$v^* \equiv 0 \pmod m, \tag{8.4}$$

but

$$v^* \neq 0. \tag{8.5}$$

Choose $a_i, a_i' \in A$ such that $\theta_q(a_i) = v_i$ and $\theta_q(a_i') = v_i'$ for $i = 1, \ldots, h$. Define

$$a^* = (a_1 + \cdots + a_h) - (a_1' + \cdots + a_h').$$

Then $a^* \in D = hA - hA$. For $i = 1, \ldots, h$,

$$v_i \equiv a_i q \pmod{p}$$

and

$$v_i' \equiv a_i' q \pmod{p},$$

hence

$$v^* \equiv a^* q \pmod{p}.$$

If $a^* \equiv 0 \pmod{p}$, then

$$v^* \equiv 0 \pmod{p}. \tag{8.6}$$

Since p is prime and $0 < m < p$, congruences (8.4) and (8.6) imply that

$$v^* \equiv 0 \pmod{mp}.$$

Since

$$|v^*| \leq h(p-1) < hp < mp,$$

it follows that $v^* = 0$, which contradicts (8.5). Therefore,

$$a^* \not\equiv 0 \pmod{p}.$$

Recall that $\gamma(a^* q \pmod{p})$ is the least nonnegative integer in the congruence class $a^* q \pmod{p}$. By (8.4),

$$v^* = \gamma(a^* q \pmod{p}) + xp \equiv 0 \pmod{m} \tag{8.7}$$

and inequality (8.3) implies that

$$-h \leq x \leq h - 1.$$

We associate to v^* the triple

$$(q, a^*, x),$$

where $1 \leq q \leq p-1$, $a^* \in D$, $a^* \not\equiv 0 \pmod{p}$, and $-h \leq x \leq h-1$. Let us count the number of such triples. Fix an integer $x \in [-h, h-1]$ and an integer $a^* \in D$ such that $a^* \not\equiv 0 \pmod{p}$. There are less than $2h|D|$ such choices. Since $a^* \not\equiv 0 \pmod{p}$, the set of integers of the form $\gamma(a^* q \pmod{p})$ for $q \in [1, p-1]$ is precisely $[1, p-1]$, and so there are at most

$$\left(\frac{p-1}{m}\right) + 1 \leq \frac{2(p-1)}{m}$$

numbers $q \in [1, p-1]$ that satisfy the congruence (8.7). Thus, the number of triples (q, a^*, x) is at most

$$4h|D| \left(\frac{p-1}{m}\right) < p - 1$$

since $m > 4h|D|$. Therefore, at least one integer $q \in [1, p-1]$ occurs in none of the triples, and for this q the map $\delta : \theta_q(A) \to A/m\mathbf{Z}$ is a Freiman isomorphism of order h. Since $\theta_q : A' \to V' \subseteq \mathbf{Z}$ is also a Freiman isomorphism of order h, it follows that there is a Freiman

isomorphism of order h from A' into $\mathbf{Z}/m\mathbf{Z}$. Moreover, $|A'| \geq |A|/h$. This completes the proof.

For $x \in \mathbf{Z}$, let $\|x\|$ denote the distance from x to the nearest integer. If $x, y \in \mathbf{Z}$ and $x \equiv y \pmod{m}$, then $\|x/m\| = \|y/m\|$. It follows that this "distance to the nearest integer" function is well-defined on congruence classes modulo m. If $g \in G = \mathbf{Z}/m\mathbf{Z}$ and x is an integer in the congruence class g, we define $\|g/m\| = \|x/m\|$. For $r_1, \ldots, r_k \in G$, we define the *Bohr neighborhood*

$$B(r_1, \ldots, r_k; \epsilon) = \{g \in G : \|gr_i/m\| \leq \epsilon \text{ for } i = 1, \ldots, k\}.$$

Theorem 8.7 (Bogolyubov) *Let A be a nonempty subset of $\mathbf{Z}/m\mathbf{Z}$, and let $|A| = \lambda m$. For some integer $k < \lambda^{-2}$ there exist congruence classes $r_1, \ldots, r_k \in G$ such that*

$$B(r_1, \ldots, r_k; 1/4) \subseteq 2A - 2A.$$

Proof. Let $G = \mathbf{Z}/m\mathbf{Z}$. For $r \in G$, we consider the additive character χ_r by

$$\chi_r(g) = e^{2\pi i r g/m}.$$

This function is well-defined on congruence classes r and g modulo m, and $\chi_0(g) = 1$ for all $g \in G$. Let

$$S_r(A) = \sum_{a \in A} \chi_r(a).$$

Then

$$S_0(A) = |A|$$

and

$$\sum_{r \in G} |S_r(A)|^2 = \sum_{r \in G} \sum_{a, a' \in A} e^{2\pi i r(a-a')/m} = |G||A| = \lambda m^2.$$

Let $g \in G$. Then

$$\sum_{r \in G} |S_r(A)|^4 \chi_r(g) = \sum_{r \in G} \sum_{a_1, a_2, a_3, a_4 \in A} e^{2\pi i r(g - a_1 - a_2 + a_3 + a_4)/m} \neq 0$$

if and only if $g \in 2A - 2A$. Let

$$R_1 = \{r \in G, r \neq 0 : |S_r| \geq \sqrt{\lambda}\ |A|\}$$

and

$$R_2 = \{r \in G : |S_r| < \sqrt{\lambda}\ |A|\}.$$

Since $S_0 = |A|$, it follows that $0 \notin R_2$, and so $R_2 \neq G$. It follows that

$$\begin{aligned} |\sum_{r \in R_2} |S_r(A)|^4 \chi_r(g)| &= \sum_{r \in R_2} |S_r(A)|^4 \\ &\leq \lambda |A|^2 |\sum_{r \in R_2} |S_r(A)|^2 \\ &< \lambda |A|^2 |\sum_{r \in G} |S_r(A)|^2 \\ &= \lambda |A|^2 |G||A| \\ &= |A|^4. \end{aligned}$$

Let $R_1 = \{r_1, r_2, \ldots, r_k\}$. If $g \in B(r_1, \ldots, r_k; 1/4)$, then $\|r_i g/m\| \leq 1/4$, and so

$$\Re(\chi_{r_i}(g)) = \Re(e^{2\pi i r_i g/m}) = \cos(2\pi r_i g/m) \geq 0.$$

It follows that

$$\begin{aligned}
\Re\left(\sum_{r\in G} |S_r(A)|^4 \chi_r(g)\right) &= |A|^4 + \sum_{r\in R_1} |S_r(A)|^4 \Re(\chi_r(g)) + \\
&\quad + \Re\left(\sum_{r\in R_2} |S_r(A)|^4 \chi_r(g)\right) \\
&\geq |A|^4 + \Re\left(\sum_{r\in R_2} |S_r(A)|^4 \chi_r(g)\right) \\
&\geq |A|^4 - |\sum_{r\in R_2} |S_r(A)|^4 \chi_r(g)| \\
&> 0.
\end{aligned}$$

Therefore,

$$\sum_{r\in G} |S_r(A)|^4 \chi_r(g) \neq 0$$

for all $g \in B(r_1, \ldots, r_k; 1/4)$, and so $B(r_1, \ldots, r_k; 1/4) \subseteq 2A - 2A$.

Finally, we must estimate $k = |R_1|$. Since $|S_r(A)| \geq \sqrt{\lambda}\,|A|$ for all $r \in R_1$, it follows that

$$k\lambda|A|^2 \leq \sum_{r\in R_1} |S_r(A)|^2 < \sum_{r\in G} |S_r(A)|^2 = |G||A| = \lambda^{-1}|A|^2$$

and so $k < \lambda^{-2}$. This completes the proof.

Theorem 8.8 *Let $r_1, \ldots, r_k$ be congruence classes in $\mathbf{Z}/m\mathbf{Z}$ such that $(r_1, \ldots, r_k, m) = 1$. Let*

$$\delta = 1/(4k)^k.$$

There exist congruence classes $q_0, q_1, \ldots, q_k \in \mathbf{Z}/m\mathbf{Z}$ and positive integers $l_1, \ldots, l_k$ such that

$$P(q_0, q_1, \ldots, q_k; l_1, \ldots, l_k) \subseteq B(r_1, \ldots, r_k; 1/4) \subseteq \mathbf{Z}m/\mathbf{Z}$$

and

$$|P| = l_1 \cdots l_k > \delta m.$$

Proof. The proof uses results from the geometry of numbers. Let $\mathbf{u} = (u_1, \ldots, u_k)$ and $\mathbf{v} = (v_1, \ldots, v_k)$ be vectors in the lattice $\mathbf{Z}^k$. We write $\mathbf{u} \equiv \mathbf{v} \pmod{m}$ if $u_i \equiv v_i \pmod{m}$ for $i = 1, \ldots, k$. Let M be the lattice of all vectors $\mathbf{v} = (v_1, \ldots, v_k) \in \mathbf{Z}^k$ such that $v_i \equiv 0 \pmod{m}$ for $i = 1, \ldots, k$. Then $\det(M) = m^k$.

Let $\mathbf{r} = (r_1', \ldots, r_k')$, where r_i' is an integer in the congruence class r_i for $i = 1, \ldots, k$. Then

$$(r_1', \ldots, r_k', m) = 1. \tag{8.8}$$

Let Λ be the set of all vectors $\mathbf{u} \in \mathbf{Z}^k$ such that $\mathbf{u} \equiv t\mathbf{r} \pmod{m}$ for some $t = 0, 1, \ldots, m-1$. Thus,

$$\Lambda = \cup_{t=0}^{m-1}(t\mathbf{r} + M).$$

The set Λ is a lattice, and M is a sublattice of Λ. The condition (8.8) implies that the m vectors $\mathbf{0}, \mathbf{r}, 2\mathbf{r}, \ldots, (m-1)\mathbf{r}$ are pairwise incongruent modulo m, and so the cosets $(t\mathbf{r} + M)$ are pairwise disjoint. It follows that the index of Λ in M is $[\Lambda : M] = m$ and, by Theorem 6.3,

$$\det(\Lambda) = \det(M)/[\Lambda : M] = m^{k-1}.$$

Let $K \subseteq \mathbf{R}^k$ be the cube consisting of all vectors $(x_1, \ldots, x_k)$ such that $|x_i| \leq 1/4$ for $i = 1, \ldots, k$. Then K is a compact, convex set symmetric with respect to the origin, and $\text{vol}(K) = 1/4^k$. Let $\lambda_1, \ldots, \lambda_k$ be the successive minima of K with respect to the lattice Λ, and let $\mathbf{b}_1, \ldots, \mathbf{b}_k$ be the corresponding set of linearly independent vectors in Λ. Let

$$\mathbf{b}_i = (b_{i,1}, \ldots, b_{i,k}) \in \lambda_i K \cap \Lambda$$

for $i = 1, \ldots, k$. It follows from Minkowski's second theorem (Theorem 6.5) that

$$\lambda_1 \cdots \lambda_k \leq 2^k \det(\Lambda)/\ \text{vol}(K) = 4^k m^{k-1}.$$

Since

$$\mathbf{b}_i \in \lambda_i K,$$

it follows that

$$|b_{i,j}| \leq \lambda_i/4$$

for $i, j = 1, \ldots, k$. Since

$$\mathbf{b}_i \in \Lambda,$$

it follows that

$$\mathbf{b}_i \equiv q_i\mathbf{r} \pmod{m}$$

for some integer $q_i \in [0, m-1]$. Then

$$b_{i,j} \equiv q_i r'_j \pmod{m}$$

for $i, j = 1, \ldots, k$. Let

$$l'_i = \left[\frac{m}{k\lambda_i}\right],$$

and let

$$P = \{q_1x_1 + \cdots + q_kx_k : |x_i| \leq l'_i\}.$$

We shall show that $P \subseteq B = B(r_1, \ldots, r_k; 1/4)$. Let $x = q_1x_1 + \cdots + q_kx_k \in P$. Then

$$xr_j = \sum_{i=1}^{k} x_i q_i r_i \equiv \sum_{i=1}^{k} x_i b_{i,j} \pmod{m},$$

and so

$$\begin{aligned}
\|xr_j/m\| &= \|\sum_{i=1}^{k} x_i b_{i,j}/m\| \\
&\leq |\sum_{i=1}^{k} x_i b_{i,j}/m| \\
&\leq \sum_{i=1}^{k} |x_i b_{i,j}/m| \\
&\leq \sum_{i=1}^{k} l_i' |b_{i,j}|/m \\
&\leq \sum_{i=1}^{k} l_i' \lambda_i/4m \\
&\leq \sum_{i=1}^{k} 1/4k \\
&= 1/4.
\end{aligned}$$

This means that $x \in B$, and so $P \subseteq B$

We shall show that the elements of P represent distinct congruence classes modulo m. Suppose that

$$q_1 x_1 + \cdots + q_k x_k \equiv q_1 y_1 + \cdots + q_k \pmod{m},$$

where $0 \leq x_i, y_i \leq l_i'$ for $i = 1, \ldots, k$. Let $z_i = x_i - y_i$. Then $|z_i| \leq 2l_i'$ and

$$\sum_{i=1}^{k} q_i z_i = q_1 z_1 + \cdots + q_k z_k \equiv 0 \pmod{m}.$$

It follows that

$$\sum_{i=1}^{k} q_i r_j z_i = \sum_{i=1}^{k} b_{i,j} z_i \equiv 0 \pmod{m}$$

for $j = 1, \ldots, k$. Since

$$\begin{aligned}
|\sum_{i=1}^{k} b_{i,j} z_i| &\leq \sum_{i=1}^{k} |b_{i,j}||z_i| \\
&\leq \sum_{i=1}^{k} (2l_i')(\lambda_i/4) \\
&\leq (1/2) \sum_{i=1}^{k} m/k \\
&= m/2,
\end{aligned}$$

it follows that $\sum_{i=1}^{k} b_{i,j} z_i = 0$ for $j = 1, \ldots, k$, and so

$$\sum_{i=1}^{k} z_i \mathbf{b}_i = \mathbf{0}.$$

Since the vectors $\mathbf{b_i}$ are linearly independent, we conclude that $z_i = 0$ for all i, and so

$$|P| = (2l_1' + 1) \cdots (2l_k' + 1).$$

Let $l_i = 2l_i' + 1$ and $q_0 = -\sum_{i=1} k l_i' q_i$. Then P is the proper k-dimensional arithmetic progression

$$P = P(q_0, q_1, \ldots, q_k; l_1, \ldots, l_k).$$

Moreover,

$$\begin{aligned} |P| &= l_1 \cdots l_k \\ &> \prod_{i=1}^{k} (l_i' + 1) \\ &> \prod_{i=1}^{k} (m/k\lambda_i) \\ &= (m/k)^k (\prod_{i=1}^{k} \lambda_i)^{-1} \\ &\geq (m/k)^k (4^k m^{k-1})^{-1} \\ &= m/(4k)^k \\ &= \delta m. \end{aligned}$$

This completes the proof.

Theorem 8.9 *Let p be a prime, and let A be a nonempty set of residues modulo p with $|A| = \lambda p$. For some $k \leq \lambda^{-2}$ there exists a proper k-dimensional arithmetic progression P such that*

$$P \subseteq 2A - 2A$$

and

$$|P| > \delta p$$

where

$$\delta = (4k)^{-k} > (\lambda^2/4)^{1/\lambda^2}.$$

Proof. By Theorem 8.7, for some integer $k < \lambda^{-2}$ there exist congruence classes $r_1, \ldots, r_k$ modulo p such that

$$B = B(r_1, \ldots, r_k; 1/4) \subseteq 2A - 2A.$$

Since p is prime and $k \geq 2$, it follows that

$$(r_1, \ldots, r_k, p) = 1.$$

By Theorem 8.8, there exists a proper k-dimensional arithmetic progression P such that $P \subseteq B$ and $|P| > \delta p$, where

$$\delta = (4k)^{-k} > (\lambda^2/4)^{1/\lambda^2}.$$

This completes the proof.

Theorem 8.10 *Let A and B be finite subsets of a torsion-free abelian group, and suppose that $|A| = |B|$ and $|A + B| \leq c|A|$. Then A is a subset of a d-dimensional arithmetic progression of length at most $l|A|$, where d and l depend only on c.*

Proof. Let G be the group generated by A. Since G is a finitely generated torsion-free abelian group, it follows from Corollary 8.2 that there is a Freiman isomorphism of order 32 between A and some set M of integers. Since $32 = 2(8 + 8)$, Theorem 8.5 implies that the difference sets $D_{8,8}(A) = 8A - 8A$ and $D_{8,8}(M) = 8M - 8M$ are Freiman isomorphic of order 2, and so $|D_{8,8}(M)| = |D_{8,8}(A)|$. It follows from Theorem 7.8 that

$$|D_{8,8}(M)| = |D_{8,8}(A)| \leq c^{16}|A| = c^{16}|M|.$$

By Bertrand's postulate, there exists a prime number p such that

$$|M| < 16|M| \leq 16|D_{8,8}(M)| < p < 32|D_{8,8}(M)| \leq 32c^{16}|M|.$$

By Theorem 8.6 with $h = 8$, there exists a set $M' \subseteq M$ such that $|M'| \geq |M|/8$ and M' is Freiman isomorphic of order 8 to a set R of congruence classes modulo p. Define λ by $\lambda p = |R|$. Then

$$\lambda p = |R| = |M'| \geq |M|/8 > p/(8 \cdot 32c^{16}).$$

Then $\lambda > 2^{-8}c^{-16}$, and

$$\lambda^{-2} < 2^{16}c^{32}.$$

By Theorem 8.9, the difference set $2R - 2R$ contains a proper k-dimensional arithmetic progression P' of length $|P'| > \delta p > \delta|M| = \delta|A|$, where

$$k < \lambda^{-2} < 2^{16}c^{32}$$

and

$$\delta = (4k)^{-k} > (2^{18}c^{32})^{-2^{16}c^{32}}$$

Since $8 = 2(2+2)$, it follows from Theorem 8.5 that the difference sets $2R-2R$ and $2M'-2M'$ are Freiman isomorphic of order 2.

The sets M and A are Freiman isomorphic of order 32, hence also of order 8. Let A' be the image of M' under this isomorphism. By Theorem 8.5, the difference sets $2M' - 2M'$ and $2A' - 2A'$ are Freiman isomorphic of order 2, and so $2R - 2R$ and $2A' - 2A'$ are Freiman

isomorphic. Let P be the image of P' under this isomorphism. By Theorem 8.4, the set P is a proper k-dimensional arithmetic progression,

$$P \subseteq 2A' - 2A' \subseteq 2A - 2A \subseteq G,$$

and

$$\delta|A| < |P'| = |P| \leq |2A - 2A| \leq c^4|A|.$$

Let $A^* = \{a_1, \ldots, a_s\}$ be a maximal set of elements in A such that the sets $a_i + P$ are pairwise disjoint. Since

$$a_i + P \subseteq A^* + P \subseteq A + P \subseteq 3A - 2A$$

for $i = 1, \ldots, s$, and

$$s|P| = \sum_{i=1}^{s} |(a_i + P)| = |A^* + P| \leq |3A - 2A| \leq c^5|A|,$$

it follows that

$$s \leq c^5|A|/|P| < c^5|A|/\delta|A| = c^5(4k)^k.$$

The set A^* is a subset of the s-dimensional arithmetic progression

$$P^* = \{x_1a_1 + \cdots + x_sa_s : 0 \leq x_i < 2 \text{ for } i = 1, \ldots, s\}$$

of length 2^s.

For every $a \in A$ there is a a_i such that

$$(a + P) \cap (a_i + A) \neq \emptyset,$$

and so there exist integers $p, p' \in P$ such that $a + p = a_i + p'$. Then

$$a = a_i + p' - p \in A^* + P - P \subseteq P^* + P - P.$$

Since $P - P$ is a k-dimensional arithmetic progression of length

$$|P - P| < 2^k|P| \leq 2^kc^4|A|,$$

it follows that A is a subset of an arithmetic progression of dimension $d = k + s$ and length less than $2^{s+k}c^4|A|$, where k and s depend only on c. This completes the proof.

Freiman's theorem 8.1 is the special case of the preceding result, when A is a finite set of integers and $B = A$.

8.5 Notes

Freiman's theorem first appeared in a short paper [28] in 1964 and in a 1966 monograph [29]. Freiman published a revised proof in the proceedings of the New York Number Theory Seminar [30]. Ruzsa's proof appears in [66] and [67]. Bogolyubov's theorem [4] is part of his proof that every almost periodic function can be uniformly approximated by trigonometric polynomials. Expositions of this work appear in Corduneanu [12], Jessen [34], and Maak [43].

8.6 Exercises

1. Let $\phi : A \to B$ be a *Freiman isomorphism of order* h, and let $A' \subseteq A$ and $B' = \phi(A')$. Prove that $\phi : A' \to B'$ is also a Freiman isomorphism of order h.

2. Let $h' \leq h$ and let $\phi : A \to B$ be a Freiman isomorphism of order h. Prove that ϕ is also a Freiman isomorphism of order h'.

3. (Freiman [29]) Let $A = \{0, 1, h+1\}$ and $B = \{0, 1, h+2\}$. Prove that A and B are Freiman isomorphic of order h but not of order $h+1$.

4. Let $P = [0, 35]$. Represent P as a d-dimensional arithmetic progression for $d = 1, 2, 3, 4, 6, 9$.

5. Let P and P' be multi-dimensional arithmetic progresions of dimensions d and d', respectively. Show that $P - P'$ is an arithmetic progression of dimension $d + d'$.

6. Let Ω be some condition, such that, if the finite subset A of an abelian group satisfies Ω, then $A \subseteq Q$, where Q is an n-dimensional arithmetic progression such that $|Q| \leq c'|A|$, where n and c' depend only on Ω. Prove that $|2A| \leq 2^n c'|A|$.

7. Let $|A|$ be a finite subset A of an abelian group such that

$$c_1|A| < |2A| < c_2|A| \log |A|.$$

   Show that A is not a large subset of a multi-dimensional arithmetic progression.

8. Let $\phi : \mathbf{Z}/m\mathbf{Z} \to \mathbf{Z}$ be the map that sends each congruence class to its least nonnegative residue. Show that ϕ is not a Freiman homomorphism of order 2.

9. Let A be a finite set of lattice points in the plane whose elements do not all lie on a line. Prove there does not exist a set B of integers such that A and B are Freiman isomorphic of order h for all $h \geq 2$.

10. Let A be a set of integers. Let $h \geq 2$, and let $N = |hA - hA|$. Prove that there exists a set $A'' \subseteq A$ such that $|A''| \geq n/h^2$ and A'' is Freiman isomorphic of order h to a subset of $[1, 2N]$. (Hint: Use Theorem 8.6.)

Chapter 9

Large subsets of $A \times B$

9.1 Introduction

Let A and B be nonempty subsets of a torsion-free abelian group, and let $|A| = |B| = k$. For $W \subseteq A \times B$, let

$$S(W) = \{a + b : (a, b) \in W\}.$$

Freiman's theorem states that if $|S(A \times B)| = |A + B| \leq ck$, then A is contained an n-dimensional arithmetic progression Q such that $|Q| \leq c'|A|$, where n and c' are parameters that depend only on c.

Balog and Szemerédi[1] proved a theorem that implies a Freiman-type result for large subsets of $A \times B$: Let $c_1 > 0$ and $c_2 > 0$. If $W \subseteq A \times B$, $|W| \geq c_1 k^2$, and $|S(W)| \leq c_2 k$, then there exists a set $A' \subseteq A$ such that $|A'| \geq c_1' k$ and $|2A'| \leq c_2' k$, where c_1' and c_2' are positive constants that depend only on c_1 and c_2. It follows that

$$|2A'| \leq c_2'|k| = (c_2'/c_1')c_1'|k| \leq c|A'|,$$

where $c = c_2'/c_1'$. Applying Freiman's theorem, we conclude that A' is a "large" subset of a multi-dimensional arithmetic progression. In this chapter we prove the theorem of Balog and Szemerédi. The proof uses an important result in graph theory, also discovered by Szemerédi, known as the Regularity Lemma.

9.2 The Regularity Lemma

Let $G = (V, E)$ be a *graph*, or, more precisely, an undirected graph with loops but with no multiple edges. Then V is a finite set of vertices, and E is the set of edges, where each edge is a set $\{v, v'\}$ of not necessarily distinct elements of V. The vertices v, v' are called the *endpoints* of the edge $e = \{v, v'\}$, and an edge e is *adjacent* to a vertex v if $v \in e$. The *degree* of a vertex v is the number of edges adjacent to v. There is at most one edge between any two vertices of the graph. In this section we shall prove an important result due to Szemerédi about the existence of a remarkable class of partitions of the vertex set of a graph.

Let A and B be subsets of V. We denote by $e(A,B)$ the number of edges with one endpoint in A and the other endpoint in B. If A and B are disjoint nonempty sets, we define the *density of edges* between A and B by

$$d(A,B) = e(A,B)/|A||B|,$$

where $|A|$ denotes the cardinality of the set A. Since $0 \leq e(A,B) \leq |A||B|$, it follows that $0 \leq d(A,B) \leq 1$.

We require the following lemmas.

Lemma 9.1 *Let $G = (V,E)$ be a graph, and let A and B be disjoint nonempty subsets of V. Let $A' \subseteq A$ and $B' \subseteq B$ satisfy*

$$\begin{aligned} |A'| &\geq (1-\delta)|A| \\ |B'| &\geq (1-\delta)|B|, \end{aligned}$$

where $0 < \delta < 1$. Then

$$\begin{aligned} |d(A,B) - d(A',B')| &\leq 2\delta/(1-\delta)^2 \\ |d(A,B)^2 - d(A',B')^2| &\leq 4\delta/(1-\delta)^2. \end{aligned}$$

In particular, if $\delta \leq 1/2$, then

$$|d(A,B) - d(A',B')| \leq 8\delta$$

and

$$|d(A,B)^2 - d(A',B')^2| \leq 16\delta.$$

Proof. Let $A'' = A \setminus A'$ and $B'' = B \setminus B'$. Then $|A''| = |A| - |A'| \leq \delta|A|$ and $|B''| \leq \delta|B|$, and

$$\begin{aligned} e(A,B) &= e(A',B') + e(A',B'') + e(A'',B') + e(A'',B'') \\ &= e(A',B') + e(A',B'') + e(A'',B) \\ &\leq e(A',B') + e(A,B'') + e(A'',B) \\ &\leq e(A',B') + |A||B''| + |A''||B| \\ &\leq e(A',B') + 2\delta|A||B|. \end{aligned}$$

It follows that

$$\begin{aligned} d(A,B) &= e(A,B)/|A||B| \\ &\leq e(A',B')/|A||B| + 2\delta \\ &\leq e(A',B')/|A'||B'| + 2\delta \\ &= d(A',B') + 2\delta \\ &\leq d(A',B') + 2\delta/(1-\delta)^2 \end{aligned}$$

and so

$$d(A,B) - d(A',B') \leq 2\delta/(1-\delta)^2.$$

Similarly,

$$\begin{aligned} d(A',B') &= e(A',B')/|A'||B'| \\ &\leq e(A,B)/|A'||B'| \\ &\leq e(A,B)/\left((1-\delta)^2|A||B|\right) \\ &= d(A,B)/(1-\delta)^2 \end{aligned}$$

and so

$$\begin{aligned} d(A',B') - d(A,B) &\leq d(A,B)\left(1/(1-\delta)^2 - 1\right) \\ &\leq \left(1/(1-\delta)^2 - 1\right) \\ &= (2\delta - \delta^2)/(1-\delta)^2 \\ &\leq 2\delta/(1-\delta)^2. \end{aligned}$$

Therefore,

$$|d(A,B) - d(A',B')| \leq 2\delta/(1-\delta)^2$$

and

$$\begin{aligned} |d(A,B)^2 - d(A',B')^2| &= |d(A,B) + d(A',B')||d(A,B) - d(A',B')| \\ &\leq 2|d(A,B) - d(A',B')| \\ &\leq 4\delta/(1-\delta)^2. \end{aligned}$$

If $0 < \delta \leq 1/2$, then $1/(1-\delta)^2 \leq 4$. This completes the proof.

Lemma 9.2 (Schwartz's inequality) *Let $x_1, \ldots, x_n$ be real numbers. Then*

$$\sum_{i=1}^{n} x_i^2 \geq (1/n)\left(\sum_{i=1}^{n} x_i\right)^2. \tag{9.1}$$

For $m = 1, \ldots, n-1$, let

$$\Delta = (1/m)\left(\sum_{i=1}^{m} x_i\right) - (1/n)\left(\sum_{i=1}^{n} x_i\right).$$

Then

$$\sum_{i=1}^{n} x_i^2 \geq (1/n)\left(\sum_{i=1}^{n} x_i\right)^2 + mn\Delta^2/(n-m). \tag{9.2}$$

Proof. Let $S_1(n) = \sum_{i=1}^n x_i$ and $S_2(n) = \sum_{i=1}^n x_i^2$. Since

$$\begin{aligned} 0 &\leq \sum_{i=1}^n (x_i - S_1(n)/n)^2 \\ &= \sum_{i=1}^n (x_i^2 - 2x_i S_1(n)/n + S_1(n)^2/n^2) \\ &= S_2(n) - 2S_1(n)^2/n + nS_1^2/n^2 \\ &= S_2(n) - S_1(n)^2/n, \end{aligned}$$

it follows that

$$S_2(n) \geq S_1(n)^2/n,$$

which is the first inequality. This implies that

$$\begin{aligned} S_2(n) - S_2(m) &= \sum_{i=m+1}^n x_i^2 \\ &\geq (1/(n-m)) \left(\sum_{i=m+1}^n x_i \right)^2 \\ &= (S_1(n) - S_1(m))^2/(n-m), \end{aligned}$$

and so

$$\begin{aligned} S_2(n) &= S_2(m) + (S_2(n) - S_2(m)) \\ &\geq S_1(m)^2/m + (S_1(n) - S_1(m))^2/(n-m) \\ &= S_1(n)^2/(n-m) - 2S_1(n)S_1(m)/(n-m) \\ &\quad + nS_1(m)^2/m(n-m) \\ &= S_1(n)^2/n + (m/n(n-m))S_1(n)^2 - (2/(n-m))S_1(n)S_1(m) \\ &\quad + (n/m(n-m))S_1(m)^2 \\ &= (1/n)S_1(n)^2 + (mn/(n-m))(S_1(n)/n - S_1(m)/m)^2 \\ &= S_1(n)^2/n + mn\Delta^2/(n-m). \end{aligned}$$

This proves the second inequality.

Lemma 9.3 *Let $G = (V, E)$ be a graph, and let*

$$A = \cup_{i=1}^q A_i \subseteq V$$

and

$$B = \cup_{j=1}^r B_j \subseteq V,$$

where

$$|A_i| = a \geq 1$$

for $i = 1, \ldots, q$,

$$|B_j| = b \geq 1$$

for $j = 1, \ldots, r$, *and*

$$A_i \cap B_j = \emptyset$$

for all i *and* j. *Then*

$$(1/qr) \sum_{i=1}^{q} \sum_{j=1}^{r} d(A_i, B_j)^2 \geq d(A, B)^2. \tag{9.3}$$

Let $0 < \theta < 1$, *and let* q' *and* r' *be integers such that* $0 < \theta q \leq q' < q$ *and* $0 < \theta r \leq r' < r$. *Let* $A' = \cup_{i=1}^{q'} A_i$ *and* $B' = \cup_{j=1}^{r'} B_j$. *Then*

$$(1/qr) \sum_{i=1}^{q} \sum_{j=1}^{r} d(A_i, B_j)^2 \geq d(A, B)^2 + \theta^2 (d(A, B) - d(A', B'))^2. \tag{9.4}$$

Proof. Since $|A| = qa$ and $|B| = rb$, it follows that

$$\begin{aligned}
(1/qr) \sum_{i=1}^{q} \sum_{j=1}^{r} d(A_i, B_j) &= (1/qr) \sum_{i=1}^{q} \sum_{j=1}^{r} e(A_i, B_j)/|A_i||B_j| \\
&= 1/(qarb) \left(\sum_{i=1}^{q} \sum_{j=1}^{r} e(A_i, B_j) \right) \\
&= e(A, B)/|A||B| \\
&= d(A, B).
\end{aligned}$$

Similarly,

$$(1/q'r') \sum_{i=1}^{q'} \sum_{j=1}^{r'} d(A_i, B_j) = d(A', B').$$

Then

$$\begin{aligned}
\Delta &= (1/qr) \sum_{i=1}^{q} \sum_{j=1}^{r} d(A_i, B_j) - (1/q'r') \sum_{i=1}^{q'} \sum_{j=1}^{r'} d(A_i, B_j) \\
&= d(A, B) - d(A', B').
\end{aligned}$$

It follows from (9.1) of Lemma 9.2 that

$$\begin{aligned}
(1/qr) \sum_{i=1}^{q} \sum_{j=1}^{r} d(A_i, B_j)^2 &\geq (1/qr)^2 \left(\sum_{i=1}^{q} \sum_{j=1}^{r} d(A_i, B_j) \right)^2 \\
&= d(A, B)^2.
\end{aligned}$$

This proves the first inequality.

The conditions $q' \geq \theta q$ and $r' \geq \theta r$ imply that

$$q'r'/(qr - q'r') \geq qr\theta^2/(qr - q'r') \geq \theta^2.$$

To obtain the second inequality, we apply (9.2) of Lemma 9.2 with $n = qr$ and $m = q'r'$. This gives

$$\begin{aligned}(1/qr)\sum_{i=1}^{q}\sum_{j=1}^{r} d(A_i, B_j)^2 &\geq (1/qr)^2 \left(\sum_{i=1}^{q}\sum_{j=1}^{r} d(A_i, B_j)\right)^2 \\ &\quad + q'r'\Delta^2/(qr - q'r') \\ &\geq d(A, B)^2 + \theta^2\Delta^2 \\ &= d(A, B)^2 + \theta^2(d(A, B) - d(A', B'))^2.\end{aligned}$$

This completes the proof.

Let $G = (V, E)$ be a graph, and let $\mathcal{P}$ be a partition of the vertex set V into $m + 1$ sets $C_0, C_1, \ldots, C_m$. The partition $\mathcal{P}$ will be called *equitable* if $|C_s| = |C_t|$ for $1 \leq s < t \leq m$. The set C_0 is called the *exceptional set* of the partition. The *partition density* of the equitable partition $\mathcal{P}$ is defined by

$$d(\mathcal{P}) = m^{-2} \sum_{1 \leq s < t \leq m} d(C_s, C_t)^2.$$

Since there are $m(m-1)/2$ summands and each summand satisfies $0 \leq d(C_s, C_t) \leq 1$, it follows that

$$0 \leq d(\mathcal{P}) < 1/2.$$

Let A and B be disjoint nonempty subsets of the vertex set V, and let $\varepsilon > 0$. The pair (A, B) will be called *ε-regular* if the conditions

$$X \subseteq A, |X| \geq \varepsilon|A|$$

and

$$Y \subseteq B, |Y| \geq \varepsilon|B|$$

imply that

$$|d(A, B) - d(X, Y)| < \varepsilon.$$

An equitable partition of V into $m + 1$ pairwise disjoint sets $C_0, C_1, \ldots, C_m$ will be called *ε-regular* if

$$|C_0| \leq \varepsilon|V|,$$

and if the pair (C_s, C_t) is ε-regular for all but at most εm^2 pairs (C_s, C_t) with $1 \leq s < t \leq m$.

The following result is the heart of the proof of the Regularity Lemma.

Lemma 9.4 *Let $0 < \varepsilon < 1$, and let the integer m satisfy*

$$4^m \geq 2^{10}\varepsilon^{-5}.$$

Let $G = (V, E)$ be a graph with k vertices, and let $\mathcal{P}$ be an equitable partition of V into $m+1$ classes $C_0, C_1, \ldots, C_m$ such that

$$|C_s| \geq 4^{2m}$$

for $s = 1, \ldots, m$. If the number of ε-irregular pairs (C_s, C_t) with $1 \leq s < t \leq m$ is greater than εm^2, then there is an equitable partition $\mathcal{P}'$ of V into $m4^m + 1$ classes such that the exceptional class has cardinality less than

$$|C_0| + k/4^m,$$

and the partition density of $\mathcal{P}$ satisfies

$$d(\mathcal{P}') \geq d(\mathcal{P}) + \varepsilon^5/32.$$

Proof. Let $q = 4^m$, and let

$$e = [|C_s|/q]$$

for $s = 1, \ldots, m$. Then

$$q^2 = 4^{2m} \leq |C_s| = eq + r < (e+1)q$$

where

$$0 \leq r < q \leq e.$$

It follows that

$$q = [|C_s|/e].$$

Let $1 \leq s < t \leq m$. If the pair (C_s, C_t) is ε-irregular, then we choose sets $X_s(t) \subseteq C_s$ and $X_t(s) \subseteq C_t$ such that

$$|X_s(t)| \geq \varepsilon |C_s| \tag{9.5}$$

$$|X_t(s)| \geq \varepsilon |C_t| \tag{9.6}$$

and

$$|d(C_s, C_t) - d(X_s(t), X_t(s))| \geq \varepsilon. \tag{9.7}$$

If the pair (C_s, C_t) is ε-regular, let $X_s(t) = X_t(s) = \emptyset$. Thus, for each $s = 1, \ldots, m$, we have constructed $m - 1$ sets

$$X_s(1), \ldots, X_s(s-1), X_s(s+1), \ldots, X_s(m)$$

contained in C_s. These sets determine a partition of C_s into 2^{m-1} pairwise disjoint sets (some of which may be empty) in the following way: Let $\Lambda = \{0,1\}^{m-1}$ be the set of all $(m-1)$-tuples of 0's and 1's. Then $|\Lambda| = 2^{m-1}$. For $\lambda = (\lambda_1, \ldots, \lambda_{s-1}, \lambda_{s+1}, \ldots, \lambda_m) \in \Lambda$, let $Y_s(\lambda)$ consist of all $v \in C_s$ such that $v \in X_s(j)$ if $\lambda_j = 1$ and $v \notin X_s(j)$ if $\lambda_j = 0$.

We shall use the 2^{m-1} sets $Y_s(\lambda)$ to construct an equitable partition $\mathcal{P}'$ that will consist of $m4^m$ sets of cardinality e, and an exceptional set of cardinality less than $|C_0| + k/4^m$.

In each set $Y_s(\lambda)$, we choose

$$q_\lambda = [|Y_s(\lambda)|/e]$$

pairwise disjoint sets, each of cardinality exactly e. Let $Y'_s(\lambda)$ be the union of these q_λ sets. Then

$$|Y_s(\lambda) \setminus Y'_s(\lambda)| < e.$$

Since the set

$$C_s \setminus \cup_{\lambda \in \Lambda} Y'_s(\lambda)$$

has cardinality exactly

$$|C_s| - \sum_{\lambda \in \Lambda} |Y'_s(\lambda)| = (eq + r) - e \sum_{\lambda \in \Lambda} q_\lambda = e(q - \sum_{\lambda \in \Lambda} q_\lambda) + r,$$

we can choose an additional

$$q - \sum_{\lambda \in \Lambda} q_\lambda$$

pairwise disjoint sets of cardinality e in C_s. This construction produces $q = 4^m$ pairwise disjoint sets, which we shall denote by $C'_s(i)$ for $i = 1, \ldots, q$. Let

$$C'_s = \cup_{i=1}^{q} C'_s(i) \subseteq C_s.$$

Then

$$C'_0 = V \setminus \cup_{s=1}^{m} C'_s \supseteq V \setminus \cup_{s=1}^{m} C_s = C_0$$

and

$$|C_s \setminus C'_s| = r < e \leq |C_s|/q,$$

hence

$$\begin{aligned} |C'_0| &= |C_0| + \sum_{s=1}^{m} |C_s \setminus C'_s| \\ &< |C_0| + \sum_{s=1}^{m} |C_s|/q \\ &\leq |C_0| + |V \setminus C_0|/q \\ &\leq |C_0| + k/4^m. \end{aligned}$$

Let $\mathcal{P}'$ be the partition of V that consists of the $mq+1 = m4^m+1$ sets $C'_s(i)$ for $1 \leq s \leq m$ and $1 \leq i \leq q$, together with the exceptional set C'_0. It remains to prove the partition density inequality

$$d(\mathcal{P}') \geq d(\mathcal{P}) + \varepsilon^5/32.$$

Since

$$|C_s \setminus C'_s| = r < e \leq |C_s|/q$$

for $s = 1, \ldots, m$, it follows that

$$|C'_s| > (1 - 1/q)|C_s|,$$

where $0 < 1/q < 1/2$. By Lemma 9.1,

$$|d(C_s, C_t) - d(C'_s, C'_t)| \leq 8/q \leq \varepsilon^5/2^7 < \varepsilon/4$$

and

$$|d(C_s, C_t)^2 - d(C'_s, C'_t)^2| \leq 16/q \leq \varepsilon^5/64$$

for $1 \leq s < t \leq m$. Applying inequality (9.3) of Lemma 9.3 with $q = r$, $A_i = C'_s(i)$, and $B_j = C'_t(j)$, we obtain

$$\begin{aligned} q^{-2} \sum_{i=1}^{q} \sum_{j=1}^{q} d(C'_s(i), C'_t(j))^2 &\geq d(C'_s, C'_t)^2 \\ &> d(C_s, C_t)^2 - \varepsilon^5/64 \end{aligned}$$

for all pairs (C_s, C_t) with $1 \leq s < t \leq m$.

Now let (C_s, C_t) be an ε-irregular pair, and let $X_s = X_s(t) \subseteq C_s$ and $X_t = X_t(s) \subseteq C_t$ be the sets that were chosen to satisfy conditions (9.5)–(9.7). In particular, $|X_s| \geq \varepsilon |C_s|$. Recall that the set C_s was partitioned into 2^{m-1} sets $Y_s(\lambda)$, where $\lambda = (\lambda_1, \ldots, \lambda_{s-1}, \lambda_{s+1}, \ldots, \lambda_m) \in \Lambda$. The set $X_s = X_s(t)$ is the union of the 2^{m-2} sets $Y_s(\lambda)$ with $\lambda_t = 1$. Each of these sets $Y_s(\lambda)$ contains $[|Y_s(\lambda)|/e]$ pairwise disjoint sets belonging to the partition $\mathcal{P}'$. Let $C'_s(1), \ldots, C'_s(q')$ denote the sets in the partition $\mathcal{P}'$ that are contained in $Y_s(\lambda)$ for some $Y_s(\lambda) \subseteq X_s$, and let

$$X'_s = \cup_{i=1}^{q'} C'_s(i).$$

Let $\sum$ denote the sum over all $\lambda \in \Lambda$ with $\lambda_t = 1$. Then

$$|X'_s| = q'e = \sum [|Y_s(\lambda)|/e]\, e.$$

Since

$$\begin{aligned} |X_s \setminus X'_s| &= \sum |Y_s(\lambda)| - |X'_s| \\ &= \sum |Y_s(\lambda)| - \sum [|Y_s(\lambda)|/e]\, e \\ &= \sum (|Y_s(\lambda)| - [|Y_s(\lambda)|/e]\, e) \\ &< \sum e \\ &= 2^{m-2} e \\ &\leq 2^{m-2} |C_s|/q \\ &= 2^{-m-2} |C_s| \\ &\leq 2^{-m-2} \varepsilon^{-1} |X_s| \\ &= q^{-1/2} 2^{-2} \varepsilon^{-1} |X_s| \\ &\leq (\varepsilon^5 2^{-10})^{1/2} 2^{-2} \varepsilon^{-1} |X_s| \\ &< 2^{-7} \varepsilon |X_s|, \end{aligned}$$

it follows that

$$|X'_s| > (1 - 2^{-7}\varepsilon)|X_s|,$$

and

$$\begin{aligned} q' &= |X'_s|/e \\ &> (1 - 2^{-7}\varepsilon)|X_s|/e \\ &\geq (1 - 2^{-7}\varepsilon)\varepsilon|C_s|/e \\ &> (1 - 2^{-7})\varepsilon q \\ &> (\varepsilon/2)q. \end{aligned}$$

Similarly, let X'_t be the union of the r' sets $C'_t(1), \ldots, C'_t(r')$ of the partition $\mathcal{P}'$ that are contained in X_t. Then

$$|X'_t| > (1 - 2^{-7}\varepsilon)|X_t|,$$

and

$$r' \geq (\varepsilon/2)q.$$

Therefore,

$$|d(X_s, X_t) - d(X'_s, X'_t)| < 8\varepsilon 2^{-7} = \varepsilon/16 < \varepsilon/4.$$

Since the pair (C_s, C_t) is not ε-regular, we obtain

$$\begin{aligned} \varepsilon &\leq |d(C_s, C_t) - d(X_s, X_t)| \\ &\leq |d(C_s, C_t) - d(C'_s, C'_t)| + |d(C'_s, C'_t) - d(X'_s, X'_t)| \\ &\quad + |d(X'_s, X'_t) - d(X_s, X_t)| \\ &< \varepsilon/4 + |d(C'_s, C'_t) - d(X'_s, X'_t)| + \varepsilon/4 \end{aligned}$$

and so

$$|d(C'_s, C'_t) - d(X'_s, X'_t)| > \varepsilon/2.$$

Applying inequality (9.4) of Lemma 9.3 with $r = q$ and $\theta = \varepsilon/2$, we see that if (C_s, C_t) is an ε-irregular pair, then

$$\begin{aligned} q^{-2} \sum_{i,j=1}^{q} d(C'_s(i), C'_t(j))^2 &\geq d(C'_s, C'_t)^2 + (\varepsilon/2)^2 (d(C'_s, C'_t) - d(X'_s, X'_t))^2 \\ &> d(C'_s, C'_t)^2 + \varepsilon^4/16 \\ &> d(C_s, C_t)^2 - \varepsilon^5/64 + \varepsilon^4/16. \end{aligned}$$

For $1 \leq s < t \leq m$, let $\chi(s,t) = 1$ if the pair (C_s, C_t) is ε-irregular, and let $\chi(s,t) = 0$ if the pair (C_s, C_t) is ε-regular. Then

$$\sum_{1 \leq s < t \leq m} \chi(s,t) \geq \varepsilon m^2,$$

and

$$q^{-2}\sum_{i,j=1}^{q} d(C'_s(i), d(C'_t(j))^2 > d(C_s, C_t)^2 - \varepsilon^5/64 + \chi(s,t)\varepsilon^4/16.$$

We can now estimate the partition density $d(\mathcal{P}')$:

$$\begin{aligned}
d(\mathcal{P}') &= (mq)^{-2}\left(\sum_{1\le s<t\le m}\sum_{i,j=1}^{q} d(C'_s(i), C'_t(j))^2 + \sum_{s=1}^{m}\sum_{1\le i<j\le q} d(C'_s(i), C_s(j))^2\right)\\
&\ge m^{-2}\sum_{1\le s<t\le m} q^{-2}\sum_{i,j=1}^{q} d(C'_s(i), C'_t(j))^2\\
&\ge m^{-2}\sum_{1\le s<t\le m}\Big(d(C_s, C_t)^2 - \varepsilon^5/64 + \chi(s,t)\varepsilon^4/16\Big)\\
&\ge m^{-2}\left(\sum_{1\le s<t\le m} d(C_s, C_t)^2 - \binom{m}{2}\varepsilon^5/64 + (\varepsilon m^2)(\varepsilon^4/16)\right)\\
&\ge m^{-2}\sum_{1\le s<t\le m}(d(C_s, C_t)^2 - m^{-2}\binom{m}{2}\varepsilon^5/64) + m^{-2}m^2\varepsilon^5/16\\
&> d(\mathcal{P}) + \varepsilon^5/32.
\end{aligned}$$

This completes the proof of the Lemma.

Theorem 9.1 (The Regularity Lemma) *Let $0 < \varepsilon < 1$ and let $m' \ge 1$. There exist numbers $K = K(\varepsilon, m')$ and $M = M(\varepsilon, m')$ such that, if $G = (V, E)$ is a graph with $|V| \ge K$ vertices, then there exists an ε-regular partition of V into $m+1$ sets, where $m' \le m \le M$.*

Proof. Let $t' = [16\varepsilon^{-5}]$. We construct a sequence $m_0, m_1, m_2, \ldots$ of integers as follows: Let m_0 be an integer such that

$$m_0 \ge m'$$

and

$$4^{m_0} \ge \max\{2^{10}\varepsilon^{-5}, 2^{t'+2}\varepsilon^{-1}\},$$

and let

$$m_{t+1} = m_t 4^{m_t}$$

for $t = 0, 1, 2, 3, \ldots$. We define numbers M and K by

$$M = M(\varepsilon, m') = m_{t'}$$

and

$$K = K(\varepsilon, m') = \max\{2m_0/\varepsilon, m_{t'}16^{m_{t'}}/(1-\varepsilon)\}.$$

Let $G = (V, E)$ be a graph with $|V| = k \ge K$. Let T be the set of nonnegative integers t with the property that there exists an equitable partition $\mathcal{P}$ of V into $m_t + 1$ sets such that

$$d(\mathcal{P}) \ge t\varepsilon^5/32 \tag{9.8}$$

and the exceptional set C_0 of the partition has cardinality

$$|C_0| < \varepsilon k(1 - 2^{-t-1}). \tag{9.9}$$

Consider any partition $\mathcal{P}_0$ of V that consists of m_0 pairwise disjoint sets of size $[k/m_0]$, together with an exceptional set C_0 of cardinality less than m_0. Then

$$|C_0| < m_0 = \varepsilon K/2 \leq \varepsilon k/2 = \varepsilon k(1 - 1/2).$$

Since $d(\mathcal{P}_0) \geq 0$, it follows that $t = 0$ satisfies conditions (9.8) and (9.9). Thus, $0 \in T$.

Since $d(\mathcal{P}) < 1/2$, if t satisfies condition (9.8), then

$$t \leq t' = [16\varepsilon^{-5}].$$

It follows that condition (9.8) is satisfied for only finitely many positive integers. Thus, the set T is finite and there exists a greatest integer $t \leq t'$ such that conditions (9.8) and (9.9) are satisfied for some partition $\mathcal{P}$ of V into $m_t + 1$ sets. Let $\mathcal{P} = \{C_0, C_1, \ldots, C_{m_t}\}$. Then

$$\begin{aligned} |C_s| &= (|V| - |C_0|)/m_t \\ &> k(1 - \varepsilon)/m_t \\ &\geq K(1 - \varepsilon)/m_{t'} \\ &\geq 16^{m_{t'}} \\ &\geq 16^{m_t} \end{aligned}$$

for $s = 1, \ldots, m_t$, and

$$4^{m_t} \geq 4^{m_0} > 2^{10}\varepsilon^{-5}.$$

Since the exceptional set of the partition $\mathcal{P}$ satisfies $|C_0| < \varepsilon k$, it follows that, if the partition $\mathcal{P}$ is not ε-regular, then the number of ε-irregular pairs (C_s, C_t) is greater that εm^2. Lemma 9.4 implies that there exists an equitable partition $\mathcal{P}'$ of V into $m_t 4^{m_t} + 1 = m_{t+1} + 1$ sets such that

$$d(\mathcal{P}') \geq d(\mathcal{P}) + \varepsilon^5/32 \geq (t + 1)\varepsilon^5/32$$

and the exceptional set C_0' of $\mathcal{P}'$ satisfies

$$\begin{aligned} |C_0'| &< |C_0| + k/4^{m_t} \\ &< \varepsilon k(1 - 2^{-t-1}) + k/4^{m_t} \\ &\leq \varepsilon k(1 - 2^{-t-1} + \varepsilon^{-1}4^{-m_t}) \\ &\leq \varepsilon k(1 - 2^{-t-1} + \varepsilon^{-1}4^{-m_0}) \\ &\leq \varepsilon k(1 - 2^{-t-1} + 2^{-t'-2}) \\ &\leq \varepsilon k(1 - 2^{-t-1} + 2^{-t-2}) \\ &= \varepsilon k(1 - 2^{-t-2}). \end{aligned}$$

This implies that $t + 1$ satisfies conditions (9.8) and (9.9), which contradicts the maximality of t. Therefore, $\mathcal{P}$ is ε-regular. This completes the proof.

9.3 The Balog-Szemerédi theorem

Theorem 9.2 *Let c_1, c_2, λ, and μ be positive real numbers. There exist positive numbers c_1', c_2' and K that depend only on c_1, c_2, λ, and μ with the following property: Let $k \geq K$, and let A and B be finite subsets of an abelian group such that*

$$\lambda k \leq |A| \leq \mu k$$

and

$$\lambda k \leq |B| \leq \mu k,$$

let W be a subset of $A \times B$ such that

$$|W| \geq c_1 k^2,$$

and let

$$S = S(W) = \{a + b : (a, b) \in W\}$$

satisfy

$$|S| \leq c_2 k. \tag{9.10}$$

Then there exists a set $A' \subseteq A$ such that

$$|A'| \geq c_1' k$$

and

$$|2A'| \leq c_2' k.$$

In particular,

$$|2A'| \leq c|A'|,$$

where $c = c_2'/c_1'$.

Proof. Let

$$0 < \varepsilon < \min\{1, \frac{c_1^2}{16\mu c_2(12\mu^2 + 2\mu c_2)}, \frac{c_1^2}{64\mu^3 c_2}\}$$

and

$$m' = [1/\varepsilon] + 1 > 1/\varepsilon.$$

Let $M(\varepsilon, m')$ and $K(\varepsilon, m')$ be the numbers constructed in the Regularity Lemma. Let $M = M(\varepsilon, m')$ and $K = K(\varepsilon, m')/\lambda$. Then M and K depend only on c_1, c_2, λ and μ.

For $s \in S$, let $r(s)$ denote the number of representations of s in the form $a + b$, where $(a, b) \in W$, that is,

$$r(s) = |\{(a, b) \in W : a + b = s\}|.$$

For each $a \in A$ there is at most one $b \in B$ such that $a + b = s$, and so

$$1 \leq r(s) \leq |A| \leq \mu k.$$

Let

$$c_3 = c_1/2c_2 > 0,$$

and let

$$S' = \{s \in S : r(s) > c_3 k\}.$$

Then

$$\begin{aligned} c_1 k^2 &\leq |W| = \sum_{s \in S} r(s) \\ &= \sum_{s \in S \setminus S'} r(s) + \sum_{s \in S'} r(s) \\ &\leq c_3 k |S \setminus S'| + \mu k |S'| \\ &\leq c_3 k |S| + \mu k |S'| \\ &\leq c_2 c_3 k^2 + \mu k |S'|, \end{aligned}$$

and so

$$|S'| \geq ((c_1 - c_2 c_3)/\mu)k = (c_1/2\mu)k.$$

Let

$$W' = \{(a,b) \in W : a + b \in S'\}.$$

Then

$$|W'| = \sum_{s \in S'} r(s) > c_3 k |S'| \geq (c_1 c_3/2\mu)k^2 = c_4 k^2,$$

where

$$c_4 = c_1 c_3/2\mu = c_1^2/4\mu c_2.$$

Let $G = (V, E)$ be the graph with vertex set

$$V = A \cup B$$

and edge set

$$E = \{\{a,b\} : (a,b) \in W'\}.$$

Since A and B are not necessarily disjoint, it follows that

$$\lambda k \leq |A| \leq |V| \leq |A| + |B| \leq 2\mu k.$$

If $e = \{a,b\} \in E$, then the set W' contains at least one and possibly both of the ordered pairs (a,b) and (b,a), and so

$$(c_4/2)k^2 \leq |W'|/2 \leq |E| \leq |W'|.$$

Since $|V| \geq \lambda k \geq \lambda K = K(\varepsilon, m')$, the Regularity Lemma implies that there exists a partition of V into $m+1$ pairwise disjoint sets $C_0, C_1, \ldots, C_m$, where

$$m' \leq m \leq M,$$

$$|C_0| \leq \varepsilon |V| \leq \varepsilon 2\mu k,$$

and

$$(1-\varepsilon)\lambda k/m \leq (1-\varepsilon)|V|/m \leq |C_i| = (|V| - |C_0|)/m \leq 2\mu k/m$$

for $i = 1, \ldots, m$. Moreover, for $1 \leq i < j \leq m$, there are at most εm^2 pairs (C_i, C_j) that fail to be ε-regular.

We shall construct a subset E' of the edge set E by deleting the following four classes of edges from E:

1. Delete all edges with at least one endpoint in C_0. Since the degree of a vertex is at most $|V|$, it follows that number of edges removed from E is at most
$$|C_0||V| \leq \varepsilon |V|^2.$$

2. For $i = 1, \ldots, m$, delete all edges with both endpoints in the same set C_i. The number of edges removed from E is at most
$$\sum_{i=1}^{m} |C_i|^2 \leq m(|V|/m)^2 = |V|^2/m \leq |V|^2/m' < \varepsilon |V|^2.$$

3. Let $1 \leq i < j \leq m$. If (C_i, C_j) is not an ε-regular pair, delete all edges between C_i and C_j. Since $e(C_1, C_j) \leq |C_i||C_j| \leq |V|^2/m^2$, and since there are at most εm^2 irregular pairs, the number of edges removed from E in this way does not exceed
$$\varepsilon m^2 |V|^2/m^2 = \varepsilon |V|^2.$$

4. We use the set S' to "color" the edges of E by assigning to the edge $\{a, b\} \in E$ the color $s = a + b \in S'$. For $i = 1, \ldots, m$ and for a fixed color $s \in S'$, we consider the number of edges of color s with at least one endpoint in the set C_i. If this number is less than $\varepsilon |C_i|$, then we delete all these edges. The total number of edges of all colors deleted from the sets $C_1, \ldots, C_m$ is at most
$$|S'| \sum_{i=1}^{m} \varepsilon |C_i| \leq m|S|\varepsilon |V|/m = \varepsilon |S||V|.$$

Let E' be the set of edges remaining after the four deletions just described. Then

$$|E \setminus E'| < 3\varepsilon |V|^2 + \varepsilon |S||V| \leq (12\mu^2 + 2\mu c_2)\varepsilon k^2$$

and so

$$\begin{aligned} |E'| &> |E| - (3\varepsilon |V|^2 + \varepsilon |S||V|) \\ &\geq (c_4/2)k^2 - (12\mu^2 + 2\mu c_2)\varepsilon k^2 \\ &= ((c_4/2) - (12\mu^2 + 2\mu c_2)\varepsilon)k^2 \\ &= ((c_1^2/8\mu c_2) - (12\mu^2 + 2\mu c_2)\varepsilon)k^2 \\ &> (c_1^2/16\mu c_2)k^2 \\ &= c_5 k^2, \end{aligned}$$

where

$$c_5 = c_1^2/16\mu c_2 > 0.$$

All of the edges in E' are edges between ε-regular pairs (C_s, C_t), where $1 \leq s < t \leq m$. Let $e'(C_s, C_t)$ (resp. $e(C_s, C_t)$) denote the number of edges in E' (resp. in E) with one endpoint in C_s and the other endpoint in C_t. Since there are at most $\binom{m}{2}$ ε-regular pairs, there must exist some ε-regular pair, say, (C_1, C_2), such that the number of edges in E' between C_1 and C_2 is

$$e'(C_1, C_2) \geq |E'|/\binom{m}{2} \geq c_5 k^2/\binom{m}{2} > c_5 k^2/m^2.$$

Therefore,

$$\begin{aligned} d(C_1, C_2) &= e(C_1, C_2)/|C_1||C_2| \\ &\geq e'(C_1, C_2)/|C_1||C_2| \\ &\geq (c_5 k^2/m^2)/(2\mu k/m)^2 \\ &= c_5/4\mu^2 \\ &= c_1^2/64\mu^3 c_2 \\ &> \varepsilon. \end{aligned}$$

Let S'' be the set of colors of edges in E' between C_1 and C_2:

$$S'' = \{a + b : \{a, b\} \in e'(C_1, C_2)\}.$$

Since $r(s) \leq \mu k$ for every color $s \in S$, it follows that

$$|S''| \geq e'(C_1, C_2)/\mu k \geq c_5 k/\mu m^2 \geq (c_5/\mu M^2)k = c_6 k.$$

We shall prove that the sumset $2S'' = S'' + S''$ is small. Let $s^* \in S''$. We fix some representation $s^* = s_1 + s_2 \in 2S''$, where $s_1 \in S''$ and $s_2 \in S''$. Since s_1 is the color of an edge that has "survived" step 4 of the deletion process described above, there are at least $\varepsilon|C_1|$ edges in E' of color s_1 with exactly one endpoint in the set C_1. Let X_1 be the set of these endpoints. Then

$$X_1 \subseteq C_1$$

and

$$|X_1| \geq \varepsilon|C_1|.$$

Similarly, there are at least $\varepsilon|C_2|$ edges in E' of color s_2 with exactly one endpoint in the set C_2. Let X_2 be the set of these endpoints. Then

$$X_2 \subseteq C_2$$

and

$$|X_2| \geq \varepsilon |C_2|.$$

Since the pair (C_1, C_2) is ε-regular, it follows that

$$d(X_1, X_2) \geq d(C_1, C_2) - \varepsilon \geq (c_1^2/64\mu^3 c_2) - \varepsilon = c_7 > 0.$$

Therefore,

$$\begin{aligned} e(X_1, X_2) &= d(X_1, X_2)|X_1||X_2| \\ &\geq c_7\varepsilon^2|C_1||C_2| \\ &\geq c_7\varepsilon^2(1-\varepsilon)^2\lambda^2 k^2 m^{-2} \\ &\geq (c_7\varepsilon^2(1-\varepsilon)^2\lambda^2 M^{-2})k^2 \\ &= c_8 k^2. \end{aligned}$$

Let Ω denote the set of all ordered triples of the form

$$(s, v_1, v_2),$$

where $s \in S$ and $v_1, v_2 \in V$. Then

$$|\Omega| = |S||V|^2 \leq c_2 4\mu^2 k^3.$$

Let $\{v_1, v_2\} \in E$ be an edge between the sets X_1 and X_2, where $v_1 \in X_1$ and $v_2 \in X_2$. Then there exist vertices $v_1' \in V$ and $v_2' \in V$ such that $\{v_1, v_1'\} \in E'$ with $v_1 + v_1' = s_1 \in S'$ and $\{v_2, v_2'\} \in E'$ with $v_2 + v_2' = s_2 \in S'$. Let $s = v_1 + v_2 \in S$. To the edge $\{v_1, v_2\}$ we associate the ordered triple

$$(s, v_1', v_2') \in \Omega.$$

Observe that

$$s + v_1' + v_2' = (v_1 + v_1') + (v_2 + v_2') = s_1 + s_2 = s^* \in 2S''.$$

Conversely, if $(s, v_1', v_2') \in \Omega$ is a triple constructed in this way from an edge $\{v_1, v_2\}$ between X_1 and X_2, then $v_1 = s_1 - v_1'$ and $v_2 = s_2 - v_2'$. Therefore, the number of distinct triples $(s, v_1', v_2') \in \Omega$ such that $s + v_1' + v_2' = s_1 + s_2 = s^*$ is at least $e(X_1, X_2) \geq c_8 k^2$. Applying this construction to each element $s^* \in 2S''$, we obtain $c_8 k^2 |2S''|$ distinct triples in Ω. Since

$$c_8 k^2 |2S''| \leq |\Omega| \leq 4c_2\mu^2 k^3,$$

it follows that

$$|2S''| \leq (4c_2\mu^2/c_8)k = c_2' k.$$

For $b \in B$, let $R(b)$ denote the number of $s \in S''$ such that $s = a + b$ for some $a \in A$. Choose $b' \in B$ such that $R(b') = \max\{R(b) : b \in B\}$. If $s \in S'' \subseteq S'$, then $r(s) \geq c_3 k$, and

so

$$\begin{aligned} c_3c_6k^2 &\leq c_3k|S''| \\ &\leq \sum_{s\in S''} r(s) \\ &= |\{(a,b)\in W : a+b\in S''\}| \\ &\leq |\{(a,b)\in A\times B : a+b\in S''\}| \\ &\leq \sum_{b\in B} R(b) \\ &\leq |B|R(b') \\ &\leq \mu k R(b'). \end{aligned}$$

Therefore, $R(b') \geq (c_3c_6/\mu)k = c_1'k$. Let

$$A' = \{a\in A : a+b'\in S''\}.$$

Then

$$|A'| = R(b') \geq c_1'k$$

and

$$|2A'| = |2(A'+\{b\})| \leq |2S''| \leq c_2'k.$$

This completes the proof of the Theorem.

Corollary 9.1 *Let c_1, c_2, λ, and μ be positive real numbers. There exist positive numbers c_1', c', n, and K that depend only on c_1, c_2, λ and μ such that the following holds: Let A and B be finite subsets of a torsion-free abelian group, let $W \subseteq A\times B$, and let $S(W) = \{a+b : (a,b)\in W\}$. For $k\geq K$, if*

$$\lambda k \leq |A| \leq \mu k,$$
$$\lambda k \leq |B| \leq \mu k,$$
$$|W| \geq c_1k^2,$$

and

$$|S(W)| \leq c_2k,$$

then there exists an n-dimensional arithmetic progression Q such that

$$|A\cap Q| \geq c_1'|A|$$

and

$$|A\cap Q| \geq c'|Q|.$$

Proof. This follows immediately from Freiman's Theorem.

Open problems. It is not known if Theorem 9.2 remains true if inequality (9.10) is replaced with the weaker condition

$$|S| \leq c_2k^{2-\delta}$$

for some $\delta > 0$. Nor is it known if the Theorem can be generalized to the sum of $h\geq 3$ finite subsets of the integers or of $h\geq 3$ finite subsets of an arbitrary abelian group.

9.4 Notes

The Regularity Lemma for ε-regular partitions of the vertices of a graph is due to Szemerédi[74]. He used a variant of this result in his proof [73] that every set of integers of positive upper density contains arbitrarily long finite arithmetic progressions. Chung[11] and Frankl and Rödl[25] have generalized the Regularity Lemma to hypergraphs. The version of Freiman's theorem for large subsets of $A \times B$ was proved by Balog and Szemerédi[1] in the case that $A = B$ is a finite set of integers. The somewhat more general result given in this chapter requires only minor modifications of their proof.

9.5 Exercises

1. Let A and B be finite subsets of an abelian group such that $|A| = |B| = k$. Let $S = A + B$. Let $r(s)$ denote the number of ordered pairs $(a, b) \in A \times B$ such that $a + b = s$. Suppose that
$$\sum_{s \in S} r(s)^2 \geq c_1 k^3.$$
   (a) Prove that there exist constants $d_1 > 0$ and $d_2 > 0$ such that, if $S' = \{s \in S : r(s) \geq d_1 k^{1/2}\}$, then $|S'| \geq d_2 k$.
   (b) Suppose that $|S| \leq c_2 k$. Prove that there exist constants $d_3 > 0$ and $d_4 > 0$ such that, if $S' = \{s \in S : r(s) \geq d_3 k\}$, then $|S'| \geq d_4 k$.

2. Let $h \geq 2$, and let $A_1, \ldots, A_h$ be finite subsets of an abelian group such that $|A_i| = k$ for $i = 1, 2, \ldots, h$. Let $S = A_1 + \cdots + A_h$. Let $r(s)$ denote the number of ordered pairs $(a_1, \ldots, a_h) \in A_1 \times \cdots \times A_h$ such that $a_1 + \cdots + a_h = s$. Prove the following:
   (a) $k \leq |S| \leq k^h$.
   (b) $r(s) \leq k^{h-1}$ for all $s \in S$.
   (c) $\sum_{s \in S} r(s) = k^h$.

3. Let $h \geq 2$, and let $A_1, \ldots, A_h$ be finite subsets of an abelian group such that $|A_i| = k$ for $i = 1, 2, \ldots, h$. Let $S = A_1 + \cdots + A_h$. Let $r(s)$ denote the number of ordered pairs $(a_1, \ldots, a_h) \in A_1 \times \cdots \times A_h$ such that $a_1 + \cdots + a_h = s$. Suppose that
$$\sum_{s \in S} r(s)^m \geq c_1 k^{h+\alpha}$$
   and
$$|S| \leq c_2 k^{h-\delta}.$$
   Let $S' = \{s \in S : r(s) \geq (c_1/2c_2) k^{(\alpha+\delta)/m}\}$. Prove that $|S'| \geq (c_1/2)k^{\beta}$, where $\beta = h + \alpha - (h-1)m$.

4. Let $x_1, \ldots, x_n$ be real numbers, and let $p_1, \ldots, p_n$ be positive real numbers such that $p_1 + \cdots + p_n = 1$. Prove that

$$\sum_{i=1}^{n} p_i x_i^2 \geq (\sum_{i=1}^{n} p_i x_i)^2.$$

5. Let $x_1, \ldots, x_n$ be real numbers, and let $p_1, \ldots, p_n$ be positive real numbers such that $p_1 + \cdots + p_n = 1$. For $m = 1, \ldots, n-1$, let $P(m) = \sum_{i=1}^{m} p_i$ and

$$\Delta = (\sum_{i=1}^{m} p_i x_i)/P(m) - \sum_{i=1}^{n} p_i x_i.$$

Prove that

$$\sum_{i=1}^{n} p_i x_i^2 \geq (\sum_{i=1}^{n} p_i x_i)^2 + P(m)\Delta^2/(1 - P(m)).$$

Chapter 10

The geometry of sumsets

10.1 Introduction

Freiman's theorem states that if a finite set A has a small sumset $2A$, then A must have an arithmetical structure in the sense that A is a large subset of a multi-dimensional arithmetic progression. In this chapter we shall prove that if A is a finite set of vectors in Euclidean space, and if the cardinality of the sumset $2A$ is very small, then A will have a geometrical structure. More precisely, if $|2A| < c|A|$, where $1 < c < 2^n$, then a positive proportion of the elements of A must lie on a hyperplane, or, equivalently, A is a subset of a bounded number of parallel hyperplanes. This result is independent of Freiman's theorem, and, indeed, played an essential role in the original proof of that theorem.

Let $n \geq 2$, and let V be an n-dimensional Euclidean space with inner product $(\ ,\ )$. Let h be a nonzero vector in V, and let $\gamma \in \mathbf{R}$. The hyperplane H defined by h and γ is the set

$$H = \{v \in V : (h, v) = \gamma\}.$$

Let A be a finite subset of V. Denote the cardinality of A by $|A|$. The *h-fold sumset* hA is the set

$$hA = \{a_1 + a_2 + \ldots + a_h : a_i \in A \text{ for } i = 1, 2, \ldots, h\}.$$

Theorem 10.1 *Let $n \geq 2$, and let*

$$1 < c < 2^n.$$

There exist constants $k_0^* = k_0^*(n, c)$ and $\varepsilon_0 = \varepsilon_0(n, c) > 0$ such that, if A is a finite subset of an n-dimensional Euclidean space V and if A satisfies

$$|A| \geq k_0^*$$

and

$$|2A| \leq c|A|,$$

then there exists a hyperplane H in V such that

$$|A \cap H| > \varepsilon_0 |A|.$$

The following example, in the case $h = 2$, shows that the upper bound on c in Theorem 10.1 is best possible.

Theorem 10.2 *Let $n \geq 2$, and let V be an n-dimensional Euclidean space. Let $h \geq 2$. For any numbers k_0 and $\varepsilon_0 > 0$ there exists a finite subset A of V such that*

$$|A| \geq k_0$$

and

$$|hA| < h^n |A|,$$

but

$$|A \cap H| \leq \varepsilon_0 |A|$$

for every hyperplane H in V.

Proof. Choose an orthonormal basis $\{e_1, \ldots, e_n\}$ for V, and let $u, v \in V$, where $u = \sum_{i=1}^n u_i e_i$ and $v = \sum_{i=1}^n v_i e_i$. Then $(u, v) = \sum_{i=1}^n u_i v_i$ is the inner product on V. Choose k_0 and $\varepsilon_0 > 0$, and let $t \in \mathbf{Z}$ satisfy

$$t \geq \max\left\{k_0^{1/n}, \varepsilon_0^{-1}\right\}.$$

Let

$$A = \left\{\sum_{i=1}^n v_i e_i \in V : v_i \in \{0, 1, \ldots, t-1\} \text{ for } i = 1, 2, \ldots, n\right\}.$$

Then

$$hA = \left\{\sum_{i=1}^n v_i e_i \in V : v_i \in \{0, 1, \ldots, ht-h\} \text{ for } i = 1, 2, \ldots, n\right\}.$$

It follows that

$$|A| = t^n \geq k_0$$

and

$$|hA| = (ht - h + 1)^n < h^n t^n = h^n |A|.$$

Let H be the hyperplane defined by $h = \sum_{i=1}^n h_i e_i$ and $\gamma \in \mathbf{R}$. Then $h_j \neq 0$ for some j. If $v = \sum_{i=1}^n v_i e_i \in A \cap H$, then $\sum_{i=1}^n h_i v_i = \gamma$, and the integer v_j is uniquely determined by the $n-1$ integers $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$. Since $v_i \in \{0, 1, \ldots, t-1\}$ for $i = 1, \ldots, n$, it follows that

$$|A \cap H| \leq t^{n-1} \leq \varepsilon_0 t^n = \varepsilon_0 |A|.$$

This completes the proof.

It is not known if there exists an $\delta > 0$ such that Theorem 10.1 remains true if the condition $2|A| \leq c|A|$ is replaced by the condition $2|A| \leq c|A|^{1+\delta}$. Nor is it known how to generalize Theorem 10.1 h-fold sumsets. For example, let $n \geq 2$, $h \geq 3$, and let $1 < c < h^n$. Do there exist constants $k_1^* = k_1^*(n, h, c)$ and $\varepsilon_1 = \varepsilon_1(n, h, c) > 0$ such that, if A is a finite subset of an n-dimensional Euclidean space V that satisfies $|A| \geq k_1^*$ and $|hA| \leq c|A|^{h-1}$, then there exists a hyperplane H in V such that $|A \cap H| > \varepsilon_1 |A|$?

10.2 Linearly independent hyperplanes

Let $n \geq 2$, and let V be an n-dimensional Euclidean space with inner product $(\ ,\ )$. Let h be a nonzero vector in V, and let $\gamma \in \mathbf{R}$. Define H, $H^{(+1)}$, and $H^{(-1)}$ as follows:

$$\begin{aligned} H &= \{v \in V : (h, v) = \gamma\} \\ H^{(+1)} &= \{v \in V : (h, v) > \gamma\} \\ H^{(-1)} &= \{v \in V : (h, v) < \gamma\}. \end{aligned}$$

The sets $H^{(+1)}$ and $H^{(-1)}$ are, respectively, the upper and lower open half spaces determined by H. The sets H, $H^{(+1)}$, and $H^{(-1)}$ are convex. The vector h is called a *normal vector* to the hyperplane H. If $0 \in H$, then H is an $(n-1)$-dimensional subspace of V. Note that $0 \in H$ if and only if $\gamma = 0$.

If h_1 and h_2 are both normal vectors to a hyperplane H, then $h_1 = \theta h_2$ for some $\theta \neq 0$. Let $E \subseteq V$. Then E is *parallel* to H if $E + v \subseteq H$ for some vector $v \in V$. Let H_1 and H_2 be hyperplanes with normal vectors h_1 and h_2, respectively. Then H_1 is parallel to H_2 if and only if $h_1 = \theta h_2$ for some $\theta \in \mathbf{R}$, $\theta \neq 0$.

Let $H_1, \ldots, H_m$ be hyperplanes. Let $\{\pm 1\}^m$ denote the set of all m-tuples $(\mu_1, \ldots, \mu_m)$ such that $\mu_i \in \{\pm 1\}$ for $i = 1, \ldots, m$. For $(\mu_1, \ldots \mu_m) \in \{\pm 1\}^m$, let

$$H(\mu_1, \ldots, \mu_m) = \bigcap_{i=1}^{m} H_i^{(\mu_i)}$$

Let $H^* = \bigcup_{i=1}^m H_i$. Then the 2^m sets $H(\mu_1, \ldots, \mu_m)$ are pairwise disjoint, and

$$V \setminus H^* = \bigcup_{(\mu_1, \ldots, \mu_m) \in \{\pm 1\}^m} H(\mu_1, \ldots, \mu_m).$$

Definition. Let V be an n-dimensional Euclidean space, and let $H_1, \ldots, H_m$ be hyperplanes with normal vectors $h_1, \ldots, h_m$, respectively. The hyperplanes $H_1, \ldots, H_m$ are *linearly independent* if the vectors $h_1, \ldots, h_m$ are linearly independent. The hyperplanes $H_1, \ldots, H_m$ are *linearly dependent* if the vectors $h_1, \ldots, h_m$ are linearly dependent.

Lemma 10.1 *Let $H_1, \ldots, H_m$ be hyperplanes in an n-dimensional Euclidean space V. Suppose that $0 \in H_i$ for $i = 1, \ldots, m$. The hyperplanes $H_1, \ldots, H_m$ are linearly independent if and only if*

$$H(\mu_1, \ldots, \mu_m) \neq \emptyset$$

for all $(\mu_1, \ldots, \mu_m) \in \{\pm 1\}^m$.

Proof. There exist nonzero vectors $h_1, \ldots, h_m$ in V such that

$$H_i = \{v \in V : (h_i, v) = 0\}$$

for $i = 1, \ldots, m$. Suppose that $H_1, \ldots, H_m$ are independent hyperplanes. Then the vectors $h_1, \ldots, h_m$ are independent, and there exists a dual set of vectors $h_1^*, \ldots, h_m^*$ such that

$$\left(h_i, h_j^*\right) = \delta_{i,j} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

for $i, j = 1, \ldots, m$. Let $(\mu_1, \ldots, \mu_m) \in \{\pm 1\}^m$. Let

$$v = \sum_{j=1}^{m} \mu_j h_j^* \in V.$$

Then

$$(h_i, v) = \sum_{j=1}^{m} \mu_j (h_i, h_j^*) = \mu_i$$

and so

$$v \in H_i^{(\mu_i)}$$

for $i = 1, \ldots, m$. It follows that $H(\mu_1, \ldots \mu_m) \neq \emptyset$.

Suppose that $H_1, \ldots H_m$ are dependent hyperplanes. Then the vectors $h_1, \ldots, h_m$ are dependent, and there exists scalars $\alpha_1, \ldots, \alpha_m$ not all zero such that $\sum_{i=1}^m \alpha_i h_i = 0$. Define $(\mu_1, \ldots, \mu_m) \in \{\pm 1\}^m$ by

$$\mu_i = \begin{cases} +1 & \text{if } \alpha_i > 0 \\ -1 & \text{if } \alpha_i \leq 0. \end{cases}$$

Then $\mu_i \alpha_i \geq 0$ for all $i = 1, \ldots, m$, and $\mu_j \alpha_j > 0$ for some j.

I claim that $H(\mu_1, \ldots, \mu_m) = \emptyset$. If not, choose $v \in H(\mu_1, \ldots, \mu_m)$. Then $v \in H_i^{(\mu_i)}$ implies that

$$\alpha_i (h_i, v) \geq 0$$

for $i = 1, \ldots, m$, and

$$\alpha_j (h_j, v) > 0.$$

It follows that

$$0 = (0, v) = \left(\sum_{i=1}^{m} \alpha_i h_i, v \right) = \sum_{i=1}^{m} \alpha_i (h_i, v) > 0,$$

which is impossible. Thus, $H(\mu_1, \ldots, \mu_m) = \emptyset$. This completes the proof.

Notation. Let S be a subset of a Euclidean space V. The *convex hull* of S, denoted $\text{conv}(S)$, is the smallest convex subset of V that contains S.

Lemma 10.2 *Let V be an n-dimensional Euclidean space, and let $H_1, \ldots, H_m$ be hyperplanes in V such that $0 \in H_i$ for $i = 1, \ldots, m$. Let S be a subset of V such that*

$$S \bigcap \left(\bigcap_{i=1}^{m} H_i^{(\mu_i)} \right) \neq \emptyset \tag{10.1}$$

for all $(\mu_1, \ldots, \mu_m) \in \{\pm 1\}^m$. *Then*

$$conv(S) \bigcap \left(\bigcap_{i=1}^{m} H_i \right) \neq \emptyset.$$

Proof. By induction on m. Let $m = 1$, and let h_1 be a normal vector for the hyperplane H_1. There exist vectors $s_1 \in S \cap H_1^{(+1)}$ and $s_2 \in S \cap H_1^{(-1)}$. Then

$$(h_1, s_1) = \alpha_1 > 0$$

and

$$(h_1, s_2) = -\alpha_2 < 0.$$

Let

$$s = \left(\frac{\alpha_2}{\alpha_1 + \alpha_2} \right) s_1 + \left(\frac{\alpha_1}{\alpha_1 + \alpha_2} \right) s_2.$$

Then $s \in \text{conv}(S)$, and $s \in H_1$, since

$$(h_1, s) = \frac{\alpha_2 (h_1, s_1)}{\alpha_1 + \alpha_2} + \frac{\alpha_1 (h_1, s_2)}{\alpha_1 + \alpha_2} = 0.$$

Therefore,

$$\text{conv}(S) \cap H_1 \neq \emptyset.$$

Let $m \geq 2$, and suppose that the Lemma holds for $m - 1$. Define $S^{(+1)}$ and $S^{(-1)}$ by

$$S^{(+1)} = S \cap H_m^{(+1)}$$

and

$$S^{(-1)} = S \cap H_m^{(-1)}.$$

Then (1) implies that

$$S^{(+1)} \cap \left(\bigcap_{i=1}^{m-1} H^{(\mu_i)} \right) \neq \emptyset$$

and

$$S^{(-1)} \cap \left(\bigcap_{i=1}^{m-1} H^{(\mu_i)} \right) \neq \emptyset$$

for all $(\mu_1, \ldots, \mu_{m-1}) \in \{\pm 1\}^{m-1}$. Then

$$\text{conv} \left(S^{(+1)} \right) \cap \left(\bigcap_{i=1}^{m-1} H_i \right) \neq \emptyset,$$

and so there exist vectors $s_1, \ldots, s_k \in S^{(+1)}$ and scalars $\alpha_1, \ldots \alpha_k \in \mathbf{R}$ with $\alpha_i \geq 0$ for $i = 1, \ldots, k$ and $\alpha_1 + \ldots + \alpha_k = 1$, such that

$$s^{(+1)} = \sum_{i=1}^{k} \alpha_i s_i \in \text{conv} \left(S^{(+1)} \right) \cap \left(\bigcap_{i=1}^{m-1} H_i \right).$$

Moreover, $s^{(+1)} \in H_m^{(+1)}$ since $s_i \in H_m^{(+1)}$ for $i = 1, \ldots, k$. Similarly, there exists

$$s^{(-1)} \in \operatorname{conv}\left(S^{(-1)}\right) \cap \left(\bigcap_{i=1}^{m-1} H_i\right)$$

such that $s^{(-1)} \in H_m^{(-1)}$.

Let $T = \{s^{(+1)}, s^{(-1)}\}$. Then $T \cap H_m^{(\mu_m)} \neq \emptyset$ for $\mu_m \in \{\pm 1\}$, and so there exists $s \in \operatorname{conv}(T) \cap H_m$. Since

$$\operatorname{conv}(T) \subseteq \operatorname{conv}(S) \cap \left(\bigcap_{i=1}^{m-1} H_i\right),$$

it follows that

$$s \in \operatorname{conv}(S) \cap \left(\bigcap_{i=1}^{m} H_i\right).$$

This completes the proof.

Lemma 10.3 *Let V be an n-dimensional Euclidean space, and let $H_1, \ldots, H_m$ be linearly independent hyperplanes in V such that $0 \in H_i$ for $i = 1, \ldots, m$. Let $1 \leq r \leq m$, and*

$$1 \leq i_1 < i_2 < \ldots < i_r \leq m.$$

Then

$$\dim\left(\bigcap_{j=1}^{r} H_{i_j}\right) = n - r.$$

In particular, if $n = m$, then

$$\bigcap_{i=1}^{n} H_i = \{0\}.$$

Proof. Let $h_1, \ldots, h_m$ be normal vectors for $H_1, \ldots H_m$, respectively. Let $W = \bigcap_{j=1}^{r} H_{i_j}$, and let

$$W^{\perp} = \{v \in V \mid (v, w) = 0 \text{ for all } w \in W\}.$$

Since $h_{i_j} \in W^{\perp}$ for $j = 1, \ldots, r$, it follows that $\dim(W^{\perp}) \geq r$, and so

$$\dim(W) = n - \dim(W^{\perp}) \leq n - r.$$

We shall prove that $\dim(W) \geq n - r$ by induction on r. Let $r = 1$. Then $W = H_{i_1}$ and $\dim(W) = \dim(H_{i_1}) = n - 1$. Let $2 \leq r \leq n$, and assume that the assertion is true for $r - 1$. Let $W' = \bigcap_{j=1}^{r-1} H_{i_j}$. Then $\dim(W') \geq n - r + 1$. Since $W = W' \cap H_{i_r}$, it follows that

$$\begin{aligned}
\dim(W) &= \dim(W' \cap H_{i_r}) \\
&= \dim(W') + \dim(H_{i_r}) - \dim(W' + H_{i_r}) \\
&\geq \dim(W') + \dim(H_{i_r}) - \dim V \\
&\geq (n - r + 1) + (n - 1) - n \\
&= n - r.
\end{aligned}$$

If $m = n$, then $\dim\left(\bigcap_{i=1}^{n} H_i\right) = 0$ and so $\bigcap_{i=1}^{n} H_i = \{0\}$. This completes the proof.

Lemma 10.4 *Let V be an n-dimensional Euclidean space, and let $H_1, \ldots, H_n$ be hyperplanes in V such that $0 \in H_i$ for $i = 1, \ldots, n$. Let S be a subset of V such that*

$$S \cap \left(\bigcap_{i=1}^{n} H_i^{(\mu_i)} \right) \neq \emptyset$$

for all $(\mu_1, \ldots, \mu_n) \in \{\pm 1\}^n$. Then

$$0 \in conv(S).$$

Proof. It follows from Lemma 1 that the hyperplanes $H_1, \ldots, H_n$ are linearly independent, and so, by Lemmas 2 and 3,

$$\text{conv}(S) \cap \{0\} = \text{conv}(S) \cap \left(\bigcap_{i=1}^{n} H_i \right) \neq \emptyset$$

This completes the proof.

Lemma 10.5 *Let V be an n-dimensional vector space, and let $H_1, \ldots, H_n$ be linearly independent hyperplanes in V with normal vectors $h_1, \ldots, h_n$, respectively. Let $0 \in H_i$ for $i = 1, \ldots, n$. Then there exists a dual basis $h_1^*, \ldots, h_n^*$ for V such that*

$$\left(h_i, h_j^*\right) = \delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

and

$$L_j = \bigcap_{i=1 i \neq j}^{n} H_i$$

is the one-dimensional subspace spanned by h_j^*.

Proof. It follows from Lemma10.3 that $\dim(L_j) = 1$. Let f_j^* be a basis vector for L_j. Then $f_j^* \in L_j$ implies that $f_j^* \in H_i$ for all $i \neq j$, and so $\left(h_i, f_j^*\right) = 0$. If $\left(h_j, f_j^*\right) = 0$, then $f_j^* \in H_j$ and so

$$f_j^* \in L_j \cap H_j = \bigcap_{i=1}^{n} H_i = \{0\},$$

which is impossible since $f_j^* \neq 0$. Therefore, $\left(h_j, f_j^*\right) \neq 0$. Let

$$h_j^* = \frac{f_j^*}{\left(h_j, f_j^*\right)}.$$

Then $(h_i, h_j^*) = \delta_{i,j}$ for $i, j = 1, \ldots, n$, and L_j is spanned by h_j^*. This completes the proof.

Lemma 10.6 *Let V be an n-dimensional vector space, and let $H_1, \ldots, H_n$ be linearly independent hyperplanes in V with $0 \in H_i$ for $i = 1, \ldots, n$. Let $Q_i = H_i \cap H_n$ for $i = 1, \ldots, n-1$. Then $\dim Q_i = n - 2$, and $Q_1, \ldots, Q_{n-1}$ are linearly independent hyperplanes in H_n. Let*

$$L_n = \bigcap_{i=1}^{n-1} H_i,$$

and let $\pi : V \to H_n$ be the projection corresponding to the direct sum decomposition

$$V = H_n \oplus L_n.$$

Let $(\mu_1, \ldots, \mu_{n-1}) \in \{\pm 1\}^{n-1}$, and let $v \in V$. If

$$v \in \bigcap_{i=1}^{n-1} H_i^{(\mu_i)},$$

then

$$\pi(v) \in \bigcap_{i=1}^{n-1} Q_i^{(\mu_i)}.$$

Let S be a subset of V, and let $\pi(S) = \{\pi(s) \mid s \in S\}$. If

$$S \cap \left(\bigcap_{i=1}^{n-1} H_i^{(\mu_i)} \right) \neq \emptyset$$

for all $(\mu_1, \ldots, \mu_{n-1}) \in \{\pm 1\}^{n-1}$, then

$$0 \in conv(\pi(S)).$$

If $S \subseteq H_n^{(+1)}$, then

$$\alpha h_n^* \in conv(S)$$

for some $\alpha > 0$.

Proof. It follows from Lemma 3 that $\dim(Q_i) = n - 2$, and so Q_i is a hyperplane in H_n for $i = 1, \ldots, n-1$.

Let $h_1, \ldots, h_n$ be normal vectors for $H_1, \ldots, H_n$, respectively, and let

$$q_i = h_i - \frac{(h_i, h_n)}{(h_n, h_n)} h_n.$$

Since $h_1, \ldots, h_n$ are linearly independent vectors, it follows that the vectors $q_1, \ldots, q_{n-1}$ are linearly independent. Moreover, $(q_i, h_n) = 0$, and so $q_i \in H_n$ for $i = 1, \ldots, n-1$.

Let $w \in H_n$. Then $(q_i, w) = (h_i, w)$, and so $w \in Q_i$ if and only if $(q_i, w) = 0$. Thus, q_i is a normal vector for Q_i, and the hyperplanes $Q_1, \ldots, Q_{n-1}$ are independent in H_n.

By Lemma 5, there is a basis vector h_n^* for L_n such that $(h_i, h_n^*) = \delta_{i,n}$. If $v \in V$, then

$$v = \pi(v) + \varphi(v)h_n^*,$$

where $\pi(v) \in H_n$ and $\varphi(v) \in \mathbf{R}$. Moreover,

$$\begin{aligned}(q_i, \pi(v)) &= (h_i, \pi(v)) - \frac{(h_i, h_n)}{(h_n, h_n)}(h_n, \pi(v)) \\ &= (h_i, \pi(v)) \\ &= (h_i, v) - \varphi(v)(h_i, h_n^*) \\ &= (h_i, v)\end{aligned}$$

for $i = 1, \ldots, n-1$. It follows that if

$$v \in \bigcap_{i=1}^{n-1} H_i^{(\mu_i)},$$

then

$$\pi(v) \in \bigcap_{i=1}^{n-1} Q_i^{(\mu_i)}.$$

If

$$S \cap \left(\bigcap_{i=1}^{n-1} H_i^{(\mu_i)} \right) \neq \emptyset$$

for all $(\mu_1, \ldots, \mu_{n-1}) \in \{\pm 1\}^{n-1}$, then

$$\pi(S) \cap \left(\bigcap_{i=1}^{n-1} Q_i^{(\mu_i)} \right) \neq \emptyset$$

for all $(\mu_1, \ldots, \mu_{n-1}) \in \{\pm 1\}^{n-1}$. It now follows from Lemma 4 that

$$0 \in \mathrm{conv}(\pi(S)).$$

Therefore, there exist vectors $s_1, \ldots, s_k \in S$ such that $s_i = \pi(s_i) + \varphi(s_i)h_n^*$, and

$$\sum_{i=1}^{k} \alpha_i \pi(s_i) = 0$$

for some nonnegative scalars $\alpha_1, \ldots, \alpha_k$ such that $\alpha_1 + \ldots + \alpha_k = 1$. Then

$$\sum_{i=1}^{k} \alpha_i s_i = \sum_{i=1}^{k} \alpha_i \varphi(s_i) h_n^* \in \mathrm{conv}(S).$$

Let $\alpha = \sum_{i=1}^{k} \alpha_i \varphi(s_i)$. If $S \subseteq H_n^{(+1)}$, then $\varphi(s_i) > 0$ for all $i = 1, \ldots, k$, and so $\alpha > 0$ and $\alpha h_n^* \in \mathrm{conv}(S)$. This completes the proof.

10.3 Parallelepipeds

Let V be an n-dimensional Euclidean space. Let $e_0 \in V$, and let $\{e_1, \ldots, e_n\}$ be a basis for V. The *standard parallelepiped* with center e_0 and basis $\{e_1, \ldots, e_n\}$ is the set

$$P(e_1, \ldots, e_n; e_0) = \left\{ e_0 + \sum_{i=1}^{n} x_i e_i \mid -1 \leq x_i \leq 1 \text{ for } i = 1, \ldots, n \right\}.$$

A subset P of V is a *parallelepiped* if $P = P(e_1, \ldots, e_n; e_0)$ for some vector $e_0 \in V$ and some basis $\{e_1, \ldots e_n\}$ of V.

Let $P = P(e_1, \ldots, e_n; e_0)$. The *vertices* of P are the 2^n vectors in the set

$$\text{vert}(P) = \left\{ e_0 + \sum_{i=1}^{n} \mu_i e_i \;\middle|\; (\mu_1, \ldots, \mu_n) \in \{\pm 1\}^n \right\}.$$

The parallelepiped P is the convex hull of its vertices. The *interior* of P is the set

$$\text{int}(P) = \left\{ e_0 + \sum_{i=1}^{n} x_i e_i \;\middle|\; -1 < x_i < 1 \text{ for } i = 1, \ldots, n \right\}.$$

Lemma 10.7 *Let V be an n-dimensional Euclidean space. Let P be a parallelepiped in V, let $B \subseteq int(P)$, and let*

$$M(B, vert(P)) = \left\{ \frac{b+s}{2} : b \in B, s \in vert(P) \right\}.$$

Then

$$M(B, vert(P)) \subseteq int(P)$$

and

$$|M(B, vert(P))| = 2^n |B|.$$

Proof. Let $P = P(e_1, \ldots, e_n; e_0)$. For $j = 1, 2$, let

$$b_j = e_0 + \sum_{i=1}^{n} x_{ij} e_i \in B \subseteq \text{int}(P)$$

and

$$s_j = e_0 + \sum_{i=1}^{n} \mu_{ij} e_i \in \text{vert}(P).$$

Then $-1 < x_{ij} < 1$ and $\mu_{ij} \in \{\pm 1\}$ for $i = 1, \ldots, n$ and $j = 1, 2$. If

$$\frac{b_1 + s_1}{2} = \frac{b_2 + s_2}{2}$$

then

$$\sum_{i=1}^{n}(x_{i1}+\mu_{i1})e_i = \sum_{i=1}^{n}(x_{i2}+\mu_{i2})e_i$$

and so

$$x_{i1}-x_{i2} = \mu_{i2}-\mu_{i1}$$

for $i = 1, \ldots, n$. Since

$$-2 < x_{i1}-x_{i2} < 2$$

and

$$\mu_{i2}-\mu_{i1} \in \{0, \pm 2\},$$

it follows that

$$x_{i1}-x_{i2} = \mu_{i2}-\mu_{i1} = 0$$

for $i = 1, \ldots, n$, and so $b_1 = b_2$ and $s_1 = s_2$. Therefore,

$$|M(B, \mathrm{vert}(P))| = |B|\,|\mathrm{vert}(P)| = 2^n|B|.$$

Also, $(x_{ij}+\mu_{ij})/2 \in (-1,1)$ for all i and j implies that $M(B, \mathrm{vert}(P)) \subseteq \mathrm{int}(P)$. This completes the proof.

Corresponding to each parallelepiped $P = P(e_1, \ldots, e_n; e_0)$ are the $2n$ *facial hyperplanes*

$$F_{j,\mu_j} = \left\{ e_0 + \mu_j e_j + \sum_{i=1 i\neq j}^{n} x_i e_i \;\middle|\; x_i \in \mathbf{R} \text{ for } i \neq j \right\}$$

where $j = 1, \ldots, n$ and $\mu_j \in \{\pm 1\}$.

Let $\{0, \pm 1\}^n$ denote the set of n-tuples $(\lambda_1, \ldots, \lambda_n)$, where $\lambda_i \in \{0, \pm 1\}$ for $i = 1, \ldots, n$. To each $(\lambda_1, \ldots, \lambda_n) \in \{0, \pm 1\}^n$ there is associated the set $D(\lambda_1, \ldots, \lambda_n)$ consisting of all vectors $e_0 + \sum_{i=1}^{n} x_i e_i \in V$ such that

$$\begin{aligned} x_i > 1 \quad & \text{if } \lambda_i = +1 \\ -1 < x_i < 1 \quad & \text{if } \lambda_i = 0 \\ x_i < -1 \quad & \text{if } \lambda_i = -1. \end{aligned}$$

Let $F^* = \bigcup_{j=1}^{n} \bigcup_{\mu_j = \pm 1} F_{j,\mu_j}$. Then $V \backslash F^*$ is the disjoint union of the 3^n open convex sets $D(\lambda_1, \ldots, \lambda_n)$. In particular,

$$D(0, 0, \ldots, 0) = \mathrm{int}(P).$$

Notation. Let $S \subseteq V$ and $T \subseteq V$. Then

$$T - S = \{t - s \mid s \in S, t \in T\}$$

and

$$T + S = \{t + s \mid s \in S, t \in T\}.$$

Lemma 10.8 *Let V be an n-dimensional Euclidean space, and let $\{f_0, f_1, \ldots, f_n\} \subseteq V$, where $f_0 = 0$. Let*

$$S_n \subseteq S_{n-1} \subseteq \ldots \subseteq S_0$$

be a sequence of sets constructed inductively by setting $S_n = \{f_n\}$ and

$$S_{k-1} = S_k \cup (\{2f_{k-1}\} - S_k)$$

for $k = 1, 2, \ldots, n$. Let $e_i = f_i - f_{i-1}$ for $i = 1, \ldots, n$. Then

$$S_0 = \left\{ \sum_{i=1}^{n} \mu_i e_i \;\middle|\; (\mu_1, \ldots, \mu_n) \in \{\pm 1\}^n \right\}.$$

If the vectors $f_1, \ldots, f_n$ are linearly independent, then

1. (i) *the vectors $e_1, \ldots e_n$ are linearly independent,*
2. (ii) *the convex hull of S_0 is the parallelepiped $P = P(e_1, \ldots, e_n; 0)$,*
3. (iii) *vert$(P) = S_0$, and*
4. (iv) *$|S_k| = 2^{n-k}$ for $k = 0, 1, \ldots, n$.*

Proof. It will be shown inductively that

$$S_k = \{f_k\} + \left\{ \sum_{i=k+1}^{n} \mu_i e_i \;\middle|\; \mu_i \in \{\pm 1\} \text{ for } i = k+1, k+2, \ldots, n \right\}$$

for $k = 0, 1, \ldots, n$. This clearly holds for $k = n$, since $S_n = \{f_n\}$. assume that the relationship holds for k. Then

$$\begin{aligned} \{2f_{k-1}\} - S_k &= \{2f_{k-1} - f_k\} - \left\{ \sum_{i=k+1}^{n} \mu_i e_i \;\middle|\; \mu_i \in \{\pm 1\} \text{ for } i = k+1, \ldots, n \right\} \\ &= \{f_{k-1}\} - \{e_k\} + \left\{ \sum_{i=k+1}^{n} \mu_i e_i \;\middle|\; \mu_i \in \{\pm 1\} \text{ for } i = k+1, \ldots, n \right\}. \end{aligned}$$

Since $\{f_k\} = \{f_{k-1}\} + \{e_k\}$, it follows that

$$\begin{aligned} S_{k-1} &= S_k \cup (\{2f_{k-1}\} - S_k) \\ &= \{f_{k-1}\} + \left\{ \sum_{i=k}^{n} \mu_i e_i \;\middle|\; \mu_i \in \{\pm 1\}^n \text{ for } i = k, \ldots, n \right\}. \end{aligned}$$

This completes the induction. For $k = 0$, we obtain

$$S_0 = \left\{ \sum_{i=1}^{n} \mu_i e_i \;\middle|\; (\mu_1, \ldots, \mu_n) \in \{\pm 1\}^n \right\}.$$

If $\{f_1, \ldots, f_n\}$ is a basis for V, then $\{e_1, \ldots, e_n\}$ is also a basis for V, and

$$|S_k| = 2^{n-k}$$

for $k = 0, 1, \ldots, n$. In particular, $|S_0| = 2^n$. Clearly,

$$\mathrm{conv}(S_0) = P(e_1, \ldots, e_n; 0) = P$$

and

$$\mathrm{vert}(P) = S_0.$$

This completes the proof.

Lemma 10.9 *Let V be an n-dimensional Euclidean space, and let $H_1, \ldots, H_n$ be hyperplanes in V. Let $\{f_0, f_1, \ldots, f_n\} \subseteq V$ satisfy $f_0 = 0$ and*

$$\begin{aligned} f_i &\in H_j \qquad \text{for } i < j \\ f_i &\in H_j^{(+1)} \qquad \text{for } i \geq j. \end{aligned}$$

Then the vectors $f_1, \ldots f_n$ are linearly independent, and the hyperplanes $H_1, \ldots, H_n$ are linearly independent.

Let $S_n \subseteq S_{n-1} \subseteq \ldots \subseteq S_0$ be the sequence of finite subsets of V constructed from $\{f_0, f_1, \ldots, f_n\}$ by the inductive procedure given in the previous Lemma. Suppose that

$$S_k \subseteq \bigcap_{j=1}^{k} H_j^{(+1)}$$

for $k = 1, \ldots, n$. Then

$$S_0 \cap H(\mu_1, \ldots, \mu_n) \neq \emptyset$$

for all $(\mu_1, \ldots, \mu_n) \in \{\pm 1\}^n$.

Let $s \in S_0 \cap H(\mu_1, \ldots, \mu_n)$. Define $S_{n+1} = \emptyset$. Then $s \in S_k \backslash S_{k+1}$ if and only if $\mu_i = 1$ for $i = 1, \ldots, k$ and $\mu_{k+1} = -1$.

Proof. If the vectors $f_1, \ldots, f_n$ are dependent, then

$$f_k = \sum_{i=1}^{k-1} x_i f_i$$

for some $k \geq 1$ and $x_1, \ldots, x_{k-1} \in \mathbf{R}$. Since $\{f_1, \ldots, f_{k-1}\} \subseteq H_k$ and the hyperplane H_k is a subspace (since $0 = f_0 \in H_k$), it follows that $f_k \in H_k$, which is impossible since also $f_k \in H_k^{(+1)}$, and $H_k \cap H_k^{(+1)} = \emptyset$. Therefore, the vectors $f_1, \ldots, f_n$ are independent. It follows from Lemma 8 that

$$|S_k| = 2^{n-k}$$

for $k = 0, 1, \ldots, n$.

We shall show by induction on k that

$$S_k \cap H(1, \ldots, 1, \mu_{k+1}, \ldots, \mu_n) \neq \emptyset \tag{10.2}$$

for all $(\mu_{k+1}, \ldots, \mu_n) \in \{\pm 1\}^{n-k}$. Let $k = n$. Since $f_n \in H_j^{(+1)}$ for all $j = 1, \ldots, n$, and $S_n = \{f_n\}$, it follows that

$$S_n \cap H(1, \ldots, 1) \neq \emptyset.$$

Suppose that (2) holds for some $k \in \{1, \ldots, n\}$. Let $(\mu_{k+1}, \ldots, \mu_n) \in \{\pm 1\}^{n-k}$. There exists

$$s \in S_k \cap H_k^{(+1)} \cap \left(\bigcap_{j=k+1}^{n} H_j^{(\mu_j)} \right).$$

Then $2f_{k-1} - s \in S_{k-1}$. Since $f_{k-1} \in H_j$ for $j = k, k+1, \ldots, n$, it follows that $2f_{k-1} - s \in H_k^{(-1)}$ and $2f_{k-1} - s \in H_j^{(-\mu_j)}$ for $j = k+1, \ldots, n$. Thus,

$$\Big(\{2f_{k-1}\} - S_k \Big) \cap H_k^{(-1)} \cap \left(\bigcap_{j=k+1}^{n} H_j^{(\mu_j)} \right) \neq \emptyset$$

for all $(\mu_{k+1}, \ldots, \mu_n) \in \{\pm 1\}^{n-k}$. Since $S_{k-1} = S_k \cup (\{2f_{k-1}\} - S_k)$ and $S_{k-1} \subseteq \bigcap_{j=1}^{k-1} H_j^{(+1)}$, it follows that

$$S_{k-1} \cap \left(\bigcap_{j=1}^{k-1} H_j^{(+1)} \right) \cap \left(\bigcap_{j=k}^{n} H_j^{(\mu_j)} \right) \neq \emptyset$$

for all $(\mu_k, \ldots, \mu_n) \in \{\pm 1\}^{n-k+1}$. This completes the induction.

In particular, it follows that

$$S_0 \cap H(\mu_1, \ldots, \mu_n) = S_0 \cap \left(\bigcap_{j=1}^{n} H_j^{(\mu_j)} \right) \neq \emptyset$$

for all $(\mu_1, \ldots, \mu_n) \in \{\pm 1\}^n$, and so, by Lemma 1, the hyperplanes $H_1, \ldots, H_n$ are linearly independent in V.

Since $|S_k| = s^{n-k}$ and $S_k \cap H(1, \ldots, 1, \mu_{k+1}, \ldots, \mu_n) \neq \emptyset$ for every $k = 0, 1, \ldots, n$ and every $(\mu_{k+1}, \ldots, \mu_n) \in \{\pm 1\}^n$, it follows that

$$|S_k \cap H(1, \ldots, 1, \mu_{k+1}, \ldots, \mu_n)| = 1$$

for every $k = 0, 1, \ldots, n$ and $(\mu_{k+1}, \ldots, \mu_n) \in \{\pm 1\}^{n-k}$. Since each of the 2^n sets $H(\mu_1, \ldots, \mu_n)$ contains exactly one element of S_0, it follows that if

$$s \in S_0 \cap H(\mu_1, \ldots, \mu_n)$$

and $\mu_i = 1$ for $i = 1, \ldots, k$, then $s \in S_k$. If $\mu_{k+1} = -1$, then $s \notin S_{k+1}$. Thus

$$S_0 \cap H(1, \ldots, 1, -1, \mu_{k+2}, \ldots, \mu_n) \subseteq S_k \backslash S_{k+1}.$$

Since

$$\left|S_0 \cap H(1,\ldots,1,-1,\mu_{k+2},\ldots,\mu_n)\right| = \left|S_k \backslash S_{k+1}\right| = 2^{n-k-1},$$

it follows that

$$S_0 \cap H(1,\ldots,1,-1,\mu_{k+2},\ldots,\mu_n) = S_k \backslash S_{k+1}.$$

This completes the proof.

Lemma 10.10 *Let the vector space V, the hyperplanes $H_1,\ldots,H_n$ and the set $F = \{f_0, f_1, \ldots, f_{n-1}\}$ satisfy the conditions of Lemma 10.9. Let $\{a_1, a_2\} \subseteq H(1,\ldots,1)$. For $i = 1, 2$, let $\{S_k(a_i)\}_{k=0}^n$ be the sequence of subsets of V constructed from the set $F \cup \{a_i\}$ by the method of Lemma 8. If $a_1 \neq a_2$, then*

$$S_0(a_1) \cap S_0(a_2) = \emptyset.$$

Proof. Suppose that $S_0(a_1) \cap S_0(a_2) \neq \emptyset$. Let k be the greatest integer such that $S_k(a_1) \cap S_k(a_2) \neq \emptyset$. If $k = n$, then $S_n(a_i) = \{a_i\}$ implies that $a_1 = a_2$.

Suppose that $k < n$. Choose

$$s \in S_k(a_1) \cap S_k(a_2).$$

Then

$$s \notin S_{k+1}(a_1) \cap S_{k+1}(a_2),$$

and so $s \notin S_{k+1}(a_i)$ for $i = 1$ or $i = 2$, say, $i = 1$. Then

$$s \in S_k(a_1) \backslash S_{k+1}(a_1) = \{2f_k\} - S_{k+1}(a_1).$$

It follows from Lemma 9 that

$$s \in H(1,\ldots,1,-1,\mu_{k+2},\ldots,\mu_n)$$

and so, again by Lemma 9,

$$s \in S_k(a_2) \backslash S_{k+1}(a_2) = \{2f_k\} - S_{k+1}(a_2).$$

Therefore, there exist vectors $s_i' \in S_{k+1}(a_i)$ such that

$$s = 2f_k - s_1' = 2f_k - s_2',$$

and so

$$s_1' = s_2' \in S_{k+1}(a_1) \cap S_{k+1}(a_2),$$

which contradicts the maximality of k. It follows that $S_0(a_1) \cap S_0(a_2) = \emptyset$ if $a_1 \neq a_2$. This completes the proof.

Lemma 10.11 *Let $\{e_1, \ldots e_n\}$ be a basis for the n-dimensional vector space V, and let $P = P(e_1, \ldots, e_n; 0)$ be the standard parallelepiped in V with center at 0. Let $b, c \in V$. If*

$$0 \in P(e_1, \ldots, e_n; b) = P + \{b\},$$

then

$$0 \in P + \{tb\}$$

for all $t \in [0, 1]$. If

$$\Big(P + \{b\}\Big) \cap \Big(P + \{c\}\Big) \neq \emptyset,$$

then

$$vert\Big(P + \{b\}\Big) \cap \Big(P + \{c\}\Big) \neq \emptyset.$$

Proof. If $0 \in P + \{b\}$, then $0 = u + b$ for some $u \in P$. Let $t \in [0, 1]$. Then $tu \in P$ since P is convex, and so

$$0 = u + b = tu + tb \in P + \{tb\}.$$

Let $b = \sum_{i=1}^n b_i e_i$ and $c = \sum_{i=1}^n c_i e_i$. If $\Big(P + \{b\}\Big) \cap \Big(P + \{c\}\Big) \neq \emptyset$, then there exist scalars $x_i, y_i \in [-1, 1]$ for $i = 1, \ldots, n$ such that

$$b + \sum_{i=1}^n x_i e_i = c + \sum_{i=1}^n y_i e_i,$$

and so

$$b_i + x_i = c_i + y_i$$

for $i = 1, \ldots, n$. Then

$$c_i - 1 \leq c_i + y_i = b_i + x_i \leq b_i + 1.$$

If $b_i + 1 \leq c_i + 1$, let $\mu_i = 1$. If $b_i + 1 > c_i + 1$, then

$$c_i - 1 < b_i - 1 \leq b_i + x_i = c_i + y_i \leq c_i + 1$$

and we let $\mu_i = -1$. In both cases,

$$c_i - 1 \leq b_i + \mu_i \leq c_i + 1$$

for $i = 1, \ldots, n$, and so

$$\sum_{i=1}^n (b_i + \mu_i) e_i = b + \sum_{i=1}^n \mu_i e_i \in \text{vert}\Big(P + \{b\}\Big) \cap \Big(P + \{c\}\Big) \neq \emptyset.$$

This completes the proof.

Theorem 10.3 *Let V be an n-dimensional vector space, and let $H_1, \ldots, H_n$ be hyperplanes in V. Let the vectors $f_0, f_1, \ldots, f_{n-1} \in V$ satisfy $f_0 = 0$ and*

$$\begin{aligned} f_i &\in H_j \qquad \text{for } i < j \\ f_i &\in H_j^{(+1)} \qquad \text{for } i \geq j. \end{aligned}$$

Let A_n be a finite subset of $H(1, \ldots, 1)$. To each $a \in A_n$, let $S_0(a)$ be the set constructed from $\{f_0, f_1, \ldots, f_{n-1}, a\}$ by letting $S_n(a) = \{a\}$ and

$$S_k(a) = S_{k+1}(a) \cup \Big(\{2f_k\} - S_{k+1}(a) \Big)$$

for $k = 0, 1, \ldots, n-1$. Suppose that

$$S_k(a) \subseteq \bigcap_{j=1}^{k} H_j^{(+1)}$$

for $k = 1, \ldots, n$ and all $a \in A_n$. Then $\mathrm{conv}(S_0(a)) = P(a)$ is a parallelepiped, and there exists a vector $a^* \in A_n$ such that

$$S_0(a) \cap P(a^*) \neq \emptyset$$

for all $a \in A_n$, and

$$\left| \left(\bigcup_{a \in A_n} S_0(a) \right) \cap P(a^*) \right| \geq |A_n|.$$

Proof. Let $e_i = f_i - f_{i-1}$ for $i = 1, \ldots, n-1$, and let $e_n(a) = a - f_{n-1}$. The set $\{e_1, \ldots, e_{n-1}\}$ is a basis for H_n, and the set $\{e_1, \ldots, e_{n-1}, e_n(a)\}$ is a basis for V. Let

$$P = P(e_1, \ldots, e_{n-1}; 0)$$

and

$$P(a) = P(e_1, \ldots, e_{n-1}, e_n(a); 0)$$

be the standard parallelepipeds in H_n and V, respectively. Then

$$\mathrm{vert}(P(a)) = \left\{ \sum_{i=1}^{n-1} \mu_i e_i + \mu_n e_n(a) \;\Big|\; (\mu_1, \ldots, \mu_n) \in \{\pm 1\}^n \right\} = S_0(a)$$

and

$$P(a) = \mathrm{conv}\Big(\mathrm{vert}(P(a)) \Big) = \mathrm{conv}(S_0(a)).$$

For $\mu_n \in \{\pm 1\}$, let

$$S_0^{(\mu_n)}(a) = \left\{ \sum_{i=1}^{n-1} \mu_i e_i + \mu_n e_n(a) \;\Big|\; (\mu_1, \ldots, \mu_{n-1}) \in \{\pm 1\}^{n-1} \right\}.$$

Then

$$S_0^{(+1)}(a) = \mathrm{vert}(P) + \{e_n(a)\},$$

and so

$$\mathrm{conv}\left(S_0^{(+1)}(a)\right) = P + \{e_n(a)\} \subseteq P(a).$$

Similarly,

$$\mathrm{conv}\left(S_0^{(-1)}(a)\right) = P - \{e_n(a)\} \subseteq P(a).$$

Since $P(a)$ is convex, it follows that

$$P + \{te_n(a)\} \subseteq P(a)$$

for all $t \in [-1, 1]$.

Let $h_1, \ldots, h_n$ be normal vectors to the hyperplanes $H_1, \ldots, H_n$, respectively. Let $L_n = \bigcap_{i=1}^{n-1} H_i$, and let h_n^* be the basis vector for L_n such that $(h_i, h_n^*) = 0$ if $i \neq n$ and $(h_n, h_n^*) = 1$. Let $Pi : V \to H_n$ be the projection corresponding to the direct sum decomposition

$$V = H_n \oplus L_n.$$

If $a \in A_n \subseteq H_n^{+1}$, then $e_n(a) = a - f_{n-1} \in H_n^{+1}$ since $f_{n-1} \in H_n$, and $e_n(a)$ can be written uniquely in the form

$$e_n(a) = g_n(a) + \varphi(a)h_n^*,$$

where $Pi(e_n(a)) = g_n(a) \in H_n$ and $\varphi(a) \in \mathbf{R}$, $\varphi(a) > 0$.

Let $a \in A_n$. By Lemma 9,

$$S_0^{(+1)}(a) \cap \left(\bigcap_{j=1}^{n-1} H_j^{(\mu_i)}\right) \neq \emptyset$$

for all $(\mu_1, \ldots, \mu_{n-1}) \in \{\pm 1\}^{n-1}$. Let $Q_i = H_i \cap H_n$ for $i = 1, \ldots, n-1$. By Lemma 6,

$$\begin{aligned} 0 \in \mathrm{conv}\Big(Pi(S_0^{(+1)}(a))\Big) &= Pi\Big(\mathrm{conv}(S_0^{(+1)}(a))\Big) \\ &= Pi(P + e_n(a)) \\ &= P + Pi(e_n(a)) \\ &= P + g_n(a). \end{aligned}$$

Choose $a^* \in A_n$ such that

$$\varphi(a^*) = \max\Big\{\varphi(a) \mid a \in A_n\Big\}.$$

Let $a \in A_n$, and let $t = \varphi(a)/\varphi(a^*)$. Then $0 < t \leq 1$.

Since $0 \in P + Pi(e_n(a^*))$, it follows from Lemma 11 that

$$\begin{aligned} 0 \in P + tPi(e_n(a^*)) &= P + Pi(te_n(a^*)), \\ &= P + tg_n(a^*) \end{aligned}$$

where

$$\begin{aligned} te_n(a^*) &= t(g_n(a^*) + \varphi(a^*)h_n^*) \\ &= tg_n(a^*) + \varphi(a)h_n^*, \end{aligned}$$

and $tg_n(a^*) \in H_n$. Then

$$0 \in (P + g_n(a)) \cap (P + tg_n(a^*)) \neq \emptyset$$

and so, by Lemma 11,

$$\mathrm{vert}(P + g_n(a)) \cap (P + tg_n(a^*)) \neq \emptyset.$$

It follows that there exists $(\mu_1, \ldots, \mu_{n-1}) \in \{\pm 1\}^{n-1}$ such that

$$\sum_{i=1}^{n-1} \mu_i e_i + g_n(a) \in P + tg_n(a^*).$$

Then

$$\sum_{i=1}^{n-1} \mu_i e_i + g_n(a) + \varphi(a)h_n^* = \sum_{i=1}^{n-1} \mu_i e_i + e_n(a) \in S_0^+(a)$$

and

$$\sum_{i=1}^{n-1} \mu_i e_i + e_n(a) \in P + tg_n(a^*) + \varphi(a)h_n^* = P + te_n(a^*) \subseteq P(a^*).$$

This proves that for all $a \in A_n$

$$S_0^{(+)}(a) \cap P(a^*) \neq \emptyset$$

and so

$$S_0(a) \cap P(a^*) \neq \emptyset.$$

By Lemma 10, if $a_1, a_2 \in A_n$ and $a_1 \neq a_2$, then $S_0(a_1) \cap S_0(a_2) = \emptyset$. This implies that

$$\left| \left(\bigcup_{a \in A_n} S_0(a) \right) \cap P(a^*) \right| \geq |A_n|.$$

This completes the proof of Theorem 3.

10.4 Proof of the theorem

The proof of Theorem 10.1 depends on the geometrical results contained in Theorem 10.3, and on the following Lemma, a proof of which is sketched in Freiman [29].

Lemma 10.12 *Let $n \geq 2$ and let $1 < c < 2^n$. Define $\varepsilon_0 = \varepsilon_0(n, c) > 0$ and $k^* = k^*(n, c)$ by*

$$\varepsilon_0 = \frac{2^n - c}{4n(3^n c + 2nc + 1)(4c)^{2^n - 1}}$$

and

$$k^* = (4c)^{2^n - 1}.$$

If V is an n-dimensional vector space, and if $A \subseteq V$ satisfies

$$|A| = k_0 \geq k^*,$$

$$|2A| \leq c|A|,$$

and

$$|A \cap H| \leq \varepsilon_0 |A|$$

for every hyperplane H in V, then there exists $B \subseteq A$ such that

$$|B| \geq \varepsilon_0 |A|$$

and

$$|2B| \leq (c - \varepsilon_0)|B|.$$

Proof. Let W be any finite, nonempty subset of V. Let

$$M(W) = \left\{ \frac{w_1 + w_2}{2} : w_1, w_2 \in W \right\} = \{v/2 : v \in 2W\}.$$

Clearly, $|M(W)| = |2W|$. If $\mathbf{C}$ is a convex subset of V and $W \subseteq \mathbf{C}$, then $M(W) \subseteq \mathbf{C}$. Let $r_W(v)$ denote the number of sets $\{w_1, w_2\} \subseteq W$ such that $(w_1 + w_2)/2 = v$. Then

$$\begin{aligned} \frac{|W|^2}{2} &< \frac{|W|(|W|+1)}{2} = \sum_{v \in M(W)} r_W(v) \\ &\leq |M(W)| \max\{r_W(v : v \in M(W)\}. \end{aligned}$$

Let $A \subseteq V$ satisfy the three hypotheses of the Lemma. We shall prove by induction that there exist hyperplanes $H_1, \ldots, H_n$ in V, vectors $f_0, f_1, \ldots, f_{n-1} \in V$ with $f_0 = 0$, and sets $A_1, \ldots, A_n$ such that

1. $f_i \in H_j$ for $0 \leq i < j \leq n$,

2. $f_i \in H_j^{(+1)}$ for $1 \leq j \leq i \leq n-1$,

3. $A_n \subseteq A_{n-1} \subseteq \ldots \subseteq A_1 \subseteq A_0 = A$,

4. $k_j = |A_j| > \frac{k_0}{(4c)^{2^j - 1}}$ for $j = 1, \ldots, n$,

5. $A_j \subseteq \bigcap_{i=1}^{j} H_i^{(+1)}$ for $j = 1, \ldots, n$, and

6. $A_{j+1} \cup (\{2f_j\} - A_{j+1}) \subseteq A_j$ for $j = 0, 1, \ldots, n-1$.

Let $A_0 = A$. Choose $f_0 \in A_0$ such that

$$r_{A_0}(f_0) = \max\{r_{A_0}(v) \mid v \in M(A_0)\}.$$

Then

$$\frac{k_0^2}{2} < \sum_{v \in M(A_0)} r_{A_0}(v) \leq r_{A_0}(f_0)|M(A_0)| \leq r_{A_0}(f_0)ck_0$$

and so

$$r_{A_0}(f_0) > \frac{k_0}{2c}.$$

Replacing A by $A - \{f_0\}$, we can assume without loss of generality that $f_0 = 0$.

Let H_1 be any hyperplane in V such that $f_0 \in H_1$. If $(w_1 + w_2)/2 = f_0$, then either $\{w_1, w_2\} \subseteq H_1$ or

$$|\{w_1, w_2\} \cap H_1^{(+1)}| = |\{w_1, w_2\} \cap H_1^{(-1)}| = 1.$$

Let

$$A_1 = \{w \in A_0 \mid w \in H_1^{(+1)} \text{ and } 2f_0 - w \in A_0\}.$$

Then

$$A_1 \subseteq A_0 \cap H_1^{(+1)} \subseteq H_1^{(+1)}$$

and

$$A_1 \cup (\{2f_0\} - A_1) \subseteq A_0.$$

Since $|A_0 \cap H_1| \leq \varepsilon_0 k_0$ and $\varepsilon_0 < (4c)^{-1}$, it follows that

$$\begin{aligned} k_1 = |A_1| &\geq r_{A_0}(f_0) - |A_0 \cap H_1| \\ &> \frac{k_0}{2c} - \varepsilon_0 k_0 \\ &> \frac{k_0}{4c}. \end{aligned}$$

Thus, H_1, f_0, and A_1, satisfy conditions (1) – (6).

Suppose that $1 \leq m \leq n-1$, and that we have constructed hyperplanes $H_1, \ldots, H_m$, vectors $f_0, f_1, \ldots, f_{m-1}$, and sets $A_1, \ldots, A_m$ that satisfy conditions (i) – (vi). Since $A_m \subseteq \bigcap_{i=1}^m H_i^{(+1)}$ and $\bigcap_{i=1}^m H_i^{(+1)}$ is convex, it follows that

$$M(A_m) \subseteq \bigcap_{i=1}^m H_i^{(+1)}.$$

Choose $f_m \in M(A_m)$ such that

$$r_{A_m}(f_m) = \max\{r_{A_m}(v) \mid v \in M(A_m)\}.$$

Then

$$\begin{aligned}\frac{1}{2}\left(\frac{k_0}{(4c)^{2^m-1}}\right)^2 &< \frac{k_m^2}{2} < r_{A_m}(f_m)|M(A_m)| \\ &\leq r_{A_m}(f_m)|M(A_0)| \\ &\leq r_{A_m}(f_m)ck_0\end{aligned}$$

and so

$$r_{A_m}(f_m) > \frac{2k_0}{(4c)^{2^{m+1}-1}}.$$

Let H_{m+1} be any hyperplane in V such that

$$\{f_0, f_1, \ldots, f_m\} \subseteq H_{m+1}.$$

Let

$$A_{m+1} = \{w \in A_m \mid w \in H_{m+1}^{(+1)} \text{ and } 2f_m - w \in A_m\}.$$

Then

$$A_{m+1} \subseteq A_m \cap H_{m+1}^{(+1)} \subseteq \bigcap_{i=1}^{m+1} H_i^{(+1)}$$

and

$$A_{m+1} \cup (\{2f_m\} - A_{m+1}) \subseteq A_m.$$

Since

$$|A_m \cap H_{m+1}| \leq |A_0 \cap H_{m+1}| \leq \varepsilon_0 k_0$$

and

$$\varepsilon_0 < \frac{1}{(4c)^{2^{m+1}-1}},$$

it follows that

$$\begin{aligned}k_{m+1} &= |A_{m+1}| \geq r_{A_m}(f_m) - |A_m \cap H_{m+1}| \\ &> \frac{2k_0}{(4c)^{2^{m+1}-1}} - \varepsilon_0 k_0 \\ &> \frac{k_0}{(4c)^{2^{m+1}-1}}.\end{aligned}$$

Thus, the hyperplanes $H_1, \ldots, H_{m+1}$, the vectors $f_0, f_1, \ldots, f_m$, and the sets $A_1, \ldots, A_{m+1}$ satisfy conditions (i) – (vi) . This completes the induction.

Let $a \in A_n$. Construct sets

$$S_n(a) \subseteq S_{n-1}(a) \subseteq \ldots \subseteq S_0(a) \subseteq A$$

by setting $S_n(a) = \{a\}$ and

$$S_j(a) = S_{j+1}(a) \cup (\{2f_j\} - S_{j+1}(a))$$

for $j = 0, 1, \ldots, n-1$. We shall prove by induction that

$$S_m(a) \subseteq A_m \subseteq \bigcap_{i=1}^{m} H_i^{(+1)}$$

for $m = 0, 1, \ldots, n$. Clearly,

$$S_n(a) = \{a\} \subseteq A_n \subseteq \bigcap_{i=1}^{n} H_i^{(+1)}.$$

Suppose that

$$S_{m+1}(a) \subseteq A_{m+1} \subseteq \bigcap_{i=1}^{m+1} H_i^{(+1)}$$

for some $m < n$. Then

$$\{2f_m\} - S_{m+1}(a) \subseteq \{2f_m\} - A_{m+1} \subseteq A_m \subseteq \bigcap_{i=1}^{m} H_i^{(+1)}$$

and so

$$S_m(a) = S_{m+1}(a) \cup (\{2f_m\} - S_{m+1}(a)) \subseteq A_m \subseteq \bigcap_{i=1}^{m} H_i^{(+1)}.$$

This completes the induction.

We have shown that the hyperplanes $H_1, \ldots, H_n$, the vectors $f_0, f_1, \ldots, f_{n-1}$, the set $A_n \subset H(1, \ldots, 1)$, and the sets

$$\{S_k(a) \mid a \in A_n \text{ and } k = 0, 1, \ldots, n\}$$

satisfy the hypotheses of Theorem 3. Therefore, there exists a vector $a^* \in A_n$ such that the parallelepiped

$$P(a^*) = \mathrm{conv}(S_0(a^*))$$

has the property that

$$\begin{aligned} |A \cap P(a^*)| &\geq \left| \left(\bigcup_{a \in A_n} S(a) \right) \cap P(a^*) \right| \\ &\geq |A_n| = k_n \\ &> \frac{k_0}{(4c)^{2^n - 1}}. \end{aligned}$$

The parallelepiped $P(a^*)$ determines the $2n$ facial hyperplanes F_{j,μ_j}, where $j = 1, \ldots, n$ and $\mu_j \in \{\pm 1\}$. Let

$$F^* = \bigcup_{j=1 \mu_j = \pm 1}^{n} F_{j,\mu_j}.$$

Let $B_0 = A \cap \mathrm{int}(P(a^*))$. Since $|A \cap H| \leq \varepsilon_0 k_0$ for every hyperplane H, and

$$\varepsilon_0 < \frac{1}{4n(4c)^{2^n-1}},$$

it follows that $|A \cap F^*| \leq 2n\varepsilon_0 k_0$,and so

$$\begin{aligned} |B_0| &= |A \cap \mathrm{int}(P(a^*))| \\ &\geq |A \cap P(a^*)| - |A \cap F^*| \\ &> \frac{k_0}{(4c)^{2^n-1}} - 2n\varepsilon_0 k_0 \\ &> \frac{k_0}{2(4c)^{2^n-1}}. \end{aligned}$$

Since

$$\mathrm{vert}(P(a^*)) = S_0(a^*) \subseteq A,$$

it follows from Lemma 7 that

$$M(B_0, S_0(a^*)) = \{\frac{b+s}{2} : b \in B_0, s \in S_0(a^*)\} \subseteq \mathrm{int}(P(a^*)) \cap M(A)$$

and

$$|M(B_0, S_0(a^*))| = 2^n |B_0|.$$

The $2n$ facial hyperplanes F_{j,μ_j} partition $V \setminus F^*$ into 3^n pairwise disjoint open convex sets $D(\lambda_1, \ldots, \lambda_n)$, where $(\lambda_1, \ldots, \lambda_n) \in \{0, \pm 1\}^n$ and

$$D(0, \ldots, 0) = \mathrm{int}(P(a^*)).$$

Then

$$B_0 = A \cap D(0, \ldots, 0).$$

Let $B_1, \ldots, B_{3^n-1}$ be the pairwise disjoint sets

$$A \cap D(\lambda_1, \ldots, \lambda_n),$$

where $(\lambda_1, \ldots, \lambda_n) \in \{0, \pm 1\}^n$ and $(\lambda_1, \ldots, \lambda_n) \neq (0, \ldots, 0)$. Since the sets $D(\lambda_1, \ldots, \lambda_n)$ are convex, it follows that

$$M(B_i) \cap M(B_j) = \emptyset$$

for $1 \leq i < j \leq 3^n - 1$, and

$$M(B_0, S_0(a^*)) \cap M(B_i) = \emptyset$$

for $i = 1, \ldots, 3^n - 1$.

We shall prove that, for some $i \in \{1, 2, \ldots, 3^n - 1\}$, the set B_i satisfies the conditions

$$|B_i| \geq \varepsilon_0 |A|$$

and

$$|2B_i| = |M(B_i)| \leq (c - \varepsilon_0)|B_i|.$$

Suppose not. Then $|B_i| \geq \varepsilon_0 k_0$ implies that $|M(B_i)| > (c - \varepsilon_0)|B_i|$. Let $\sum'$ denote the sum over all $i \in [1, 3^n - 1]$ such that $|B_i| \geq \varepsilon_0 k_0$. It follows that

$$\begin{aligned}
ck_0 &\geq |2A| = |M(A)| \\
&\geq |M(B_0, S(a^*))| + \sum_{i=1}^{3^n-1} |M(B_i)| \\
&\geq 2^n|B_0| + \sum{}' |M(B_i)| \\
&\geq 2^n|B_0| + (c - \varepsilon_0) \sum{}' |B_i| \\
&> 2^n|B_0| + c \sum{}' |B_i| - \varepsilon_0 k_0 \\
&> 2^n|B_0| + c \sum_{i=1}^{3^n-1} |B_i| - 3^n c\varepsilon_0 k_0 - \varepsilon_0 k_0 \\
&= (2^n - c)|B_0| + c \sum_{i=0}^{3^n-1} |B_i| - 3^n c\varepsilon_0 k_0 - \varepsilon_0 k_0 \\
&= (2^n - c)|B_0| + c(k_0 - |F^* \cap A|) - 3^n c\varepsilon_0 k_0 - \varepsilon_0 k_0 \\
&\geq (2^n - c)|B_0| + ck_0 - (3^n c + 2nc + 1)\varepsilon_0 k_0.
\end{aligned}$$

This implies that

$$(3^n c + 2nc + 1)\varepsilon_0 k_0 > (2^n - c)|B_0| > \frac{(2^n - c)k_0}{2(4c)^{2^n-1}}$$

and so

$$\varepsilon_0 > \frac{2^n - c}{2(3^n c + 2nc + 1)(4c)^{2^n-1}}.$$

This contradicts the choice of

$$\varepsilon_0 = \frac{2^n - c}{4n(3^n c + 2nc + 1)(4c)^{2^n-1}}.$$

Therefore, there exists a set $B \subseteq A$ such that $|B| \geq \varepsilon_0|A|$ and $|2B| \leq (c - \varepsilon_o)|B|$. This completes the proof.

Proof of Theorem 1. Let $\varepsilon_0 = \varepsilon_0(n, c) > 0$ be the constant defined in Lemma 12. Let $t = t(n, c)$ be the unique positive integer such that

$$t - 1 < \frac{c - 1}{\varepsilon_0} \leq t.$$

Let $k^* = (4c)^{2^n-1}$, and let

$$k_0^* = k_0^*(n, c) = \varepsilon_0^{-t} k^*.$$

Let A be a subset of V such that

$$|A| \geq k_0^*$$

and

$$|2A| \leq c|A|.$$

Suppose that

$$|A \cap H| \leq \varepsilon_0 |A|$$

for every hyperplane H in V. By Lemma 12, there exists a set $B \subseteq A$ such that

$$|B| \geq \varepsilon_0 |A| \geq \varepsilon_0^{-(t-1)} k^*$$

and

$$|2B| \leq (c - \varepsilon_0)|B|.$$

Let $B = W_1$ and $A = W_0$. Let $2 \leq j \leq t$. Suppose that we have constructed sets

$$W_{j-1} \subseteq W_{j-2} \subseteq \ldots \subseteq W_1 \subseteq W_0 = A$$

such that

$$|W_i| \geq \varepsilon_0^{-(t-i)} k^*$$

and

$$|2W_i| \leq (c - i\varepsilon_0)|W_i|$$

for $i = 1, \ldots, j-1$. Since $|W_{j-1}| \geq k^*$ and $|2W_{j-1}| \leq (c - (j-1)\varepsilon_0)|W_{j-1}|$, where

$$1 < c - (j-1)\varepsilon_0 < c,$$

it follows from Lemma 12 that there exists a set $W_j \subseteq W_{j-1}$ such that

$$|W_j| \geq \varepsilon_0 |W_{j-1}| \geq \varepsilon_0^{-(t-j)} k^*$$

and

$$|2W_j| \leq (c - j\varepsilon_0)|W_j|.$$

In particular, when $j = t$, we obtain a set W_t such that

$$|W_t| \geq k^* \geq 2$$

and

$$|2W_t| \leq (c - t\varepsilon_0)|W_t| \leq |W_t|.$$

This is impossible, since

$$|2W_t| > |W_t|$$

if W_t is any finite set of vectors satisfying $|W_t| \geq 2$. Therefore, there exists a hyperplane H in V such that $|A \cap H| > \varepsilon_0 |A|$. This completes the proof of Theorem 10.1.

Theorem 10.4 *Let $n \geq 2$, and let*

$$1 < c < 2^n.$$

There exist constants $k_0^* = k_0^*(n,c)$ and $\ell = \ell(n,c)$ such that, if A is a finite subset of an n-dimensional Euclidean space V and if A satisfies

$$|A| \geq k_0^*$$

and

$$|2A| \leq c|A|,$$

then there exist ℓ parallel hyperplanes $H_1, \ldots, H_\ell$ in V such that

$$A \subseteq \bigcup_{i=1}^{\ell} H_i.$$

Proof. By Theorem 10.1, there exists a hyperplane H in V and a number $\varepsilon_0 = \varepsilon_0(n,c) > 0$ such that $|A \cap H| > \varepsilon_0|A|$. Then $H = \{v \in V : (h,v) = \gamma\}$, where h is a nonzero vector in V and $\gamma \in \mathbf{R}$. If $a \in A$ and $(h,a) = \gamma'$, then a lies in the hyperplane $H' = \{v \in V : (h,v) = \gamma'\}$, and this hyperplane is parallel to H. Let $H = H_1, H_2, \ldots, H_\ell$ be a set of pairwise disjoint hyperplanes, each parallel to H, such that

$$A \subseteq \bigcup_{i=1}^{\ell} H_i.$$

Choose $a_i \in A \cap H_i$ for $i = 1, \ldots, \ell$. Then

$$\bigcup_{i=1}^{\ell} (a_i + (A \cap H)) \subseteq 2A$$

and

$$|a_i + (A \cap H)| = |A \cap H| > \varepsilon_0 k.$$

If $b \in a_i + (A \cap H)$, then $(h,b) = \gamma_i + \gamma$, and so

$$(a_i + (A \cap H)) \cap (a_j + (A \cap H)) = \emptyset$$

for $i \neq j$. It follows that

$$\ell\varepsilon_0|A| < \sum_{i=1}^{\ell} |a_i + (A \cap H)| \leq |2A| \leq c|A|,$$

and so

$$\ell < c/\varepsilon_0.$$

This completes the proof.

10.5 Notes

Theorem 10.1, in the case where A is a finite set of integer lattice points, appears as Lemma 2.12 in Freiman's monograph [29], but the proof in the book is difficult to follow. Fishburn [24] has published an independent proof for the special case $n = 2$, but this proof does not generalize to higher dimensions. The proof given in this chapter for all $n \geq 2$ and for arbitrary finite subsets of V comes from Nathanson [54], whose approach follows closely that of Freiman, and includes complete proofs of various geometrical results on which Freiman's argument depends.

10.6 Exercises

1. Let V be an n-dimensional Euclidean space with inner product $(\ ,\ )$, and let $e_1, e_2, \ldots, e_n$ be a basis for V. Construct a basis $e_1^*, e_2^*, \ldots, e_n^*$ for V such that $(e_i, e_j^*) = \delta_{i,j}$ for $1 \leq i, j \leq n$.

2. Let H be a hyperplane, and let $v \in H$. Prove that $H - v$ is a subspace.

3. Let H be a hyperplane with $0 \in H$. Prove that H is a subspace of dimension $n - 1$.

4. Prove that a standard parallelepiped is the convex hull of its vertices.

5. Prove that Theorem 10.4 implies Theorem 10.1.

Bibliography

[1] A. Balog and E. Szemerédi, preprint.

[2] R. P. Bambah, A. Woods, and H. Zassenhaus, Three proofs of Minkowski's second inequality in the geometry of numbers, *J. Austral. Math. Soc.* **5** (1965), 453–462.

[3] H. F. Blichfeldt, A new principle in the geometry of numbers, with some application, *Trans. Amer. Math. Soc.* **15** (1914), 227–235.

[4] N. N. Bogolyubov, Sur quelques propriétés arithmétiques des presque-périodes, *Ann. Chaire Math. Phys. Kiev* **4** (1939), 185–194.

[5] L. V. Brailovsky and G. A. Freiman, On a product of finite subsets in a torsion-free group, *J. Algebra* **130** (1990), 462–476.

[6] J. W. S. Cassels, *An Introduction to the Geometry of Numbers*, Springer-Verlag, Berlin, 1959.

[7] J. W. S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge Univ. Press, Cambridge, 1965.

[8] A. L. Cauchy, Recherches sur les nombres, *J. École polytech.* **9** (1813), 99–116.

[9] I. Chowla, A theorem on the addition of residue classes: Application to the number $\Gamma(k)$ in Waring's problem, *Proc. Indian Acad. Sci.* **2** (1935), 242-243.

[10] I. Chowla, *Proc. Nat. Acad. Sci. India A* **13** (1943), 195–220.

[11] F. R. K. Chung, Regularity lemmas for hypergraphs and quasi-randomness, *Random Structures and Algorithms* **2** (1991), 241–252.

[12] C. Corduneanu, *Almost Periodic Functions*, Interscience, New York, 1968.

[13] J. G. van der Corput, On sets of integers, I, II, III, *Proc. Konig. Neder. Akad. Wetensch.* **50** (1947), 252–261, 340–350, 429–435.

[14] I. Danicic, An elementary proof of Minkowski's second inequality, *J. Austral. Math. Soc.* **10** (1969), 177–181.

[15] H. Davenport, On the addition of residue classes, *J. London Math. Soc.* **10** (1935), 30–32.

[16] H. Davenport, Minkowski's inequality for the minima associated with a convex body, *Quart. J. Math.* **10** (1939), 119–121.

[17] H. Davenport, A historical note, *J. London Math. Soc.* **22** (1947), 100–101.

[18] P. H. Diananda, Critical subsets of finite abelian groups, *J. London Math. Soc.* **43** (1968), 479–81.

[19] G. A. Dirac, Short proof of Menger's graph theorem, *Mathematika* **13** (1966), 42–44.

[20] F. J. Dyson, A theorem on the densities of sets of integers, *J. London Math. Soc.* **20** (1945), 8–14.

[21] H. G. Eggleston, *Convexity*, Cambridge University Press, Cambridge, 1966.

[22] P. Erdős, On the arithmetical density of the sum of two sequences, one of which forms a basis for the integers, *Acta Arith.* **1** (1936), 197–200.

[23] P. Erdős and G. Freiman, On two additive problems, *J. Number Theory* **34** (1990), 1–12.

[24] P. C. Fishburn, On a contribution of Freiman to additive number theory, *J. Number Theory* **35** (1990), 325–334.

[25] P. Frankl and V. Rödl, The uniformity lemmas for hypergraphs, *Graphs and Combinatorics* **8** (1992), 309–312.

[26] G. A. Freiman, On the addition of finite sets. I, *Izv. Vysh. Zaved. Matematika* no. 6 (13) (1959), 202–213.

[27] G. A. Freiman, Inverse problems of additive number theory. VI. On the addition of finite sets. III, *Izv. Vysh. Ucheb. Zaved. Matematika* no. 3 (28) (1962), 151–157.

[28] G. A. Freiman, Addition of finite sets, *Doklady Akad. Nauk SSSR* **158** (1964), 1038–1041 = *Soviet Math.–Doklady* **5** (1964), 1366–1370.

[29] G. A. Freiman, *Foundations of a Structural Theory of Set Addition* (in Russian), Kazan Gos. Ped. Inst., Kazan, 1966 = Translations of Mathematical Monographs, Vol. 37, American Mathematical Society, Providence, 1973.

[30] G. A. Freiman, What is the structure of K if $K + K$ is small? in: D. V. Chudnovsky, G. V. Chudnovsky, H. Cohn, and M. B. Nathanson, eds., *Number Theory, New York 1984–85*, Lecture Notes in Mathematics, Vol. 1240, Springer-Verlag, Heidelberg, 1987, pp. 109–134.

[31] G. A. Freiman, Sumsets and powers of 2, preprint.

[32] G. A. Freiman, H. Halberstam, and I. Z. Ruzsa, Integer sum sets containing long arithmetic progressions, *J. London Math. Soc.* **46** (1992), 193–201.

[33] H. Halberstam and K. F. Roth, *Sequences*, Vol. 1, Clarendon, Oxford, 1966; 2nd ed., Springer–Verlag, Heidelberg, 1983.

[34] B. Jessen, Review of paper by N. N. Bogolyubov, *Math. Reviews* **8** (1947), p. 512.

[35] J. H. B. Kemperman, On complexes in a semigroup, *Indag. Math.* **18** (1956), 247–254.

[36] J. H. B. Kemperman, On small subsets of an abelian group, *Acta Math.* **103** (1960), 63–88.

[37] A. Khinchin, Über ein metrisches Problem der additiven Zahlentheorie, *Mat. Sbornik N.S.* **10** (1933) 180–189.

[38] M. Kneser, Abschätzungen der asymptotischen Dichte von Summenmengen, *Math. Zeit.* **58** (1953), 459–484.

[39] M. Kneser, Anwendung eines Satzes von Mann aauf die Geometrie der Zahlen, *Proc. Int. Cong. Math. Amsterdam* **2** (1954), 32.

[40] M. Kneser, Ein Satz über abelsche Gruppen mit Anwendungen auf die Geometrie der Zahlen, *Math. Zeit.* **61** (1955), 429–434.

[41] M. Kneser, Summenmengen in lokalkompakten abelsche Gruppen, *Math. Zeit.* **66** (1956), 88–110.

[42] S. Lang, *Algebra*, Addison-Wesley

[43] W. Maak, *Fastperiodische Funktionen*, 2nd ed., Springer-Verlag, Heidelberg, 1967.

[44] J. L. Malouf, The density of the sum of a basis and a set of positive density; a bound of Helmut Plünnecke, preprint.

[45] H. B. Mann, On products of sets of group elements, *Canadian J. Math.***4** (1952), 64–66.

[46] H. B. Mann, *Addition Theorems*, Interscience Publishers, John Wiley, New York, 1965.

[47] H. Minkowski, *Geometrie der Zahlen*, Teubner, Leipzig, 1896, 1910; reprinted by Chelsea, New York, 1953.

[48] L. J. Mordell, On some arithmetical results in the geometry of numbers, *Compositio Math.* **1** (1935248–253.),

[49] J. C. M. Nash, Freiman's theorem answers a question of Erdős, *J. Number Theory* **27** (1987), 7–8.

[50] M. B. Nathanson, Sums of finite sets of integers, *Amer. Math. Monthly* **79** (1972), 1010–1012.

[51] M. B. Nathanson, Best possible results on the density of sumsets, in: B. Berndt, H. Diamond, H. Halberstam, and A. Hildebrand, eds., *Analytic Number Theory*, Birkhäuser, Boston, 1990, pp. 395–403.

[52] M. B. Nathanson, On a problem of Rohrbach for finite groups, *J. Number Theory* (1992), to appear.

[53] M. B. Nathanson, The simplest inverse problems in additive number theory, in: A. Pollington and W. Moran, eds., *Number Theory with an Emphasis on the Markoff Spectrum*, Marcel Dekker, New York, 1993, pp. 191–206.

[54] M. B. Nathanson, An inverse theorem for sums of sets of lattice points, *J. Number Theory* **43** (1993), to appear.

[55] M. B. Nathanson, On a conjecture of Ruzsa and Turjányi, preprint.

[56] M. B. Nathanson and A. Sárkőzy, Sumsets containing long arithmetic progressions and powers of 2, *Acta Arith.* **54** (1989), 147–154.

[57] H. Plünnecke, Über ein metrisches Problem der additiven Zahlentheorie, *J. reine angew. Math.* **197** (1957). 97–103.

[58] H. Plünnecke, Über die Dichte der Summe zweier Mengen, deren eine die Dichte null hat, *J. reine angew. Math.* **205** (1960), 1–20.

[59] H. Plünnecke, *Eigenschaften und Abschätzungen von Wirkungsfunktionen*, Berichte der Gesellschaft für Mathematik und Datenverarbeitung, Nr. 22, Bonn, 1969.

[60] H. Plünnecke, Eine zahlentheoretische Anwendung der Graphtheorie, *J. reine angew. Math.* **243** (1970), 171–183.

[61] L.P. Postnikova, Fluctuations in the distribution of fractional parts, *Doklady Akad. Nauk SSSR* **161** (1965), 1282–1284 = *Soviet Math.–Doklady* **6** (1965), 597–600.

[62] I. Z. Ruzsa, On the cardinality of $A + A$ and $A - A$, in: *Combinatorics (Keszthely, 1976)*, Coll. Math. Soc. J. Bolyai **18** North Holland Publishing Company, Amsterdam, 1979, pp. 933-938.

[63] I. Z. Ruzsa, Sets of sums and differences, in: *Séminaire de Théorie des Nombres de Paris 1982–1983*, Birkhäuser, Boston, 1984, pp. 267-273.

[64] I. Z. Ruzsa, On the number of sums and differences, *Acta Math. Sci. Hungar.*

[65] I. Z. Ruzsa, An application of graph theory to additive number theory, *Scientia, Ser. A* **3** (1989), 97–109.

[66] I. Z. Ruzsa, Arithmetic progressions and the number of sums, *Periodica Math. Hungar.*, to appear.

[67] I. Z. Ruzsa, Generalized arithmetic progressions and sumsets, to appear.

[68] I. Z. Ruzsa, Sums of finite sets, to appear.

[69] J. C. Shepherdson, On the addition of elements of a sequence, *J. London Math. Soc.* **22** (1947), 85–88.

[70] L. G. Shnirel'man, Über additive Eigenschaften von Zahlen, *Math. Ann.* **107** (1933), 649-690.

[71] C. L. Siegel, *Lectures on the geometry of numbers*, Springer-Verlag, Berlin, 1989.

[72] A. Stöhr and E. Wirsing, Beispiele von wesentlichen Komponenten, die keine Basen sind, *J. reine angew. Math.* **196** (1956), 96-98.

[73] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.

[74] E. Szemerédi, Regular partitions of graphs, in: J.-C. Bermond, et al. eds., *Problèmes Combinatoires et Théorie des Graphes*, Colloques internationaux C.N.R.S., No. 260, Paris, 1978, pp. 399-401.

[75] A. G. Vosper, The critical pairs of subsets of a group of prime order, *J. London Math. Soc.* **31** (1956), 200–205; Addendum 280–282.

[76] H. Wegmann, Ein Dichtebegriff für Mengen dyadisch rationaler Zahlen, *Math. Zeit.* **92** (1966), 51-64.