CHAPTER 1

ELEMENTARY ESTIMATES

In this chapter we introduce the basic notation used throughout the course, and we establish some elementary, but important bounds. Do not be misled by the word 'elementary', as many of our arguments are elementary in the sense that they use only simple notions, but the arguments themselves can be very involved and long. However, in this first chapter we avoid the use of complicated reasoning.

Let Z be a commutative ring, and A and B are subsets of Z. Typically we work with the integers $\mathbb{Z}$, or the residue classes mod m, that is with $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$, though sometimes with other rings and fields like $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{Z}^d$. In many cases we only need the additive group of the ring, or there is no ring at all, only a group written additively (the name 'additive combinatorics' comes from here). The *sumset* of A and B is defined by

$$A + B = \{x \in Z : \, x = a + b \text{ with some } a \in A,\, b \in B\}.$$

Typically we just write $A + B = \{a + b : \, a \in A,\, b \in B\}$ with the understanding that elements are not repeated in $A + B$. For $A + A$ we sometimes write $2A$. Similarly

$$\begin{gathered} kA = A + A + \cdots + A, \\ b + A = \{b\} + A = \{b + a : \, a \in A\}, \text{ a translate of } A, \\ A - B = \{a - b : \, a \in A,\, b \in B\}, \text{ the difference set }, \\ A \cdot B = \{ab : \, a \in A,\, b \in B\}, \text{ the product set }, \\ A/B = \{a/b : \, a \in A,\, b \in B\}, \text{ the quotient set }, \\ k \cdot A = \{k\} \cdot A = \{ka;\, a \in A\}, \text{ a dilate of } A. \end{gathered}$$

Note that with this convention $\mathbb{Z}/m\mathbb{Z}$ was an abuse of notation, from now on we use the short $\mathbb{Z}_m$ to denote the ring of residue classes mod m. Actually we do not plan to work in general groups or rings, always in one of the above list, but the reader is very welcome to double check the arguments if they can indeed be carried out in a more general context.

Theorem 1.1: Let A and B be two finite subsets of the real numbers $\mathbb{R}$. We have

$$|A| + |B| - 1 \le |A + B| \le |A||B|,$$

and equality can hold on both places. Moreover the lower bound is sharp (equality holds) iff (if and only if) A and B are arithmetic progressions of the same difference.

Proof: The upper bound is obvious and reflects the fact that as a worth case all sums in the form $a + b : \, a \in A,\, b \in B$ are different. This is indeed possible, for example, let

$$A = \{0, 1, \ldots, r - 1\}, \quad B = r \cdot \{0, 1, \ldots, s - 1\} = \{0, r, 2r, \ldots, (s - 1)r\},$$

and as is easily seen

$$A + B = \{0, 1, \ldots, rs - 1\}.$$

Let us concentrate on the lower bound. We can list the elements of the sets as

$$A = \{a_1 < a_2 < \cdots < a_r\}, \quad B = \{b_1 < b_2 < \cdots < b_s\}.$$

The next chain of inequalities contains well visibly $r + s - 1$ different elements of $A + B$.

$$a_1 + b_1 < a_1 + b_2 < a_1 + b_3 < \cdots < a_1 + b_s < a_2 + b_s < a_3 + b_s < \cdots < a_r + b_s\,.$$

This proves the lower bound and it is not hard to see that if $A = \{1, \ldots, r\}$ and $B = \{1, \ldots, s\}$ then $A + B = \{2, \ldots, r + s\}$. It is not hard to come up with several other chains of $r + s - 1$ different elements of $A + B$ in increasing order. For example,

$$a_1 + b_1 < a_2 + b_1 < a_2 + b_2 < a_2 + b_3 < \cdots < a_2 + b_s < a_3 + b_s < \cdots < a_r + b_s\,.$$

If $|A + B| = r + s - 1$ then the above two chains should be the same. The first terms of the chains, as well as the last $r - 1$ terms visibly agree, but the rest determines $s - 1$ equations, namely

$$\begin{aligned} a_1 + b_2 &= a_2 + b_1\,, \\ a_1 + b_3 &= a_2 + b_2\,, \\ &\vdots \\ a_1 + b_s &= a_2 + b_{s-1}\,. \end{aligned}$$

Rearranging these equations we arrive at

$$\begin{aligned} b_2 - b_1 &= a_2 - a_1\,, \\ b_3 - b_2 &= a_2 - a_1\,, \\ &\vdots \\ b_s - b_{s-1} &= a_2 - a_1\,. \end{aligned}$$

This means that the set B is indeed an arithmetic progression with difference $a_2 - a_1$. Either we can show a third chain of inequalities to cover the rest of the proof, but it is more elegant to argue that the roles of A and B are symmetric in the statement, we can interchange them to show that the set A is indeed an arithmetic progression with difference $b_2 - b_1$. This concludes the proof.

□

Theorem 1.2: Let $A_1, \ldots, A_k$, $k \geq 2$ be finite subsets of the real numbers $\mathbb{R}$. We have

$$|A_1| + \cdots + |A_k| - k + 1 \leq |A_1 + \cdots + A_k| \leq |A_1| \ldots |A_k|,$$

and equality can hold on both places. Moreover the lower bound is sharp iff $A_1, \dots, A_k$ are arithmetic progressions of the same difference.

Proof: This is a typical proof by induction. Note that the case $k = 2$ is precisely Theorem 1.1, and therefore this case is proved. We can suppose that $k > 2$, and that the statement is true for $k - 1$ and for 2. Let $A = A_1 + \cdots + A_{k-1}$ and $B = A_k$. First observe that $A + B = A_1 + \cdots + A_k$. Theorem 1.1 and the induction hypothesis say that

$$|A_1| + \cdots + |A_{k-1}| - k + 2 + |A_k| - 1 \leq |A| + |B| - 1 \leq |A + B| \leq |A||B| \leq |A_1| \dots |A_k|.$$

The two inequalities are therefore proved. Next observe that the only way to have equality in the lower bound is the following.

$$|A_1| + \cdots + |A_{k-1}| - k + 2 = |A_1 + \cdots + A_{k-1}| = |A|\,,$$
$$|A| + |B| - 1 = |A + B|\,.$$

By the induction hypothesis the first relation implies that $A_1, \dots, A_{k-1}$ are arithmetic progressions of the same difference, and therefore $A_1 + \cdots + A_{k-1} = A$ is also an arithmetic progression of the same difference, while the second relation implies that $B = A_k$ is also an arithmetic progression of the same difference.

□

Corollary 1.3: Let A be a finite subset of the real numbers $\mathbb{R}$, and $k \geq 2$ be an integer. We have

$$k|A| - k + 1 \leq |kA| \leq |A|^k,$$

moreover the lower bound is sharp iff A is an arithmetic progression.

□

The upper bound here is obviously not sharp unless $|A| = 1$, as $a_1 + a_2 = a_2 + a_1$, for example, and so not all sums can be different. We leave as an exercise to show that the sharp upper bound is $\binom{|A|+k-1}{k}$. It is a very interesting question to construct the largest set of integers $A \subset \{1, \dots, N\}$ such that $|A + A| = \binom{|A|+1}{2}$. Such an extremal set is called a 'Sidon set' and we plan to return to their study in Chapter 2.

Realize the main pattern of Corollary 1.3. It says if $A + A$ is smallest possible then A has a well defined structure, namely A should be an arithmetic progression. It was Gregori Freiman in the 1950's who recognized if $A + A$ is small, say $\leq C|A|$, then A still has a special structure, though not that strong as above. One of the main purpose of this course is to understand and to prove Freiman's structure theorem, but we do it only in Chapter 10. Right now we try to guess the result.

Our strategy is simply investigating examples. If A is a set of r members of an arithmetic progression of length $2r$, say, then $A + A$ is part of an arithmetic progression of length $4r - 1$, and thus $|A + A| \leq 4r - 1$. However, hoping that $|A + A| \leq C|A|$ implies that A is part of an arithmetic progression of length $C|A|/2$ is too optimistic.

Consider the following example. $A = \{0, 1, \dots, r - 2, N\}$ with $N > r - 2$ so $|A| = r$. One can easily see that $A + A = \{0, 1, \dots, 2r - 4\} \cup \{N, N + 1, \dots, N + r - 2\} \cup \{2N\}$. If $r - 2 < N \leq 2r - 4$ then $|A + A| = N + r \leq 3r - 4$, while if $N > 2r - 4$ then

$|A+A| = 3r-3$. In other words, we have found an example of a set A, which is r members of an $N+1\,(\leq 2r-3)$ elements arithmetic progression such that $|A+A| = N+r \leq 3r-4$, however we have also found an example of a set A such that $|A+A| = 3r-3$ but A can not be covered with any arithmetic progression of reasonable size. One can prove that the first part of this discovery is actually the extremal situation whenever $|A+A| \leq 3|A|-4$. For larger sumset, however, the arithmetic progression is no longer the underlining structure.

Next consider

$$A = \{a_0 + a_1 n_1 + a_2 n_2 + \cdots + a_d n_d : 0 \leq n_j < N_j \text{ for } j = 1, \ldots, d\},$$

where a_0, a_1, ... ,a_d are fixed real numbers (or integers, or ...) and N_1, N_2, ... ,N_d are integers ≥ 2. This is a *generalized arithmetic progression* of *dimension* d and *volume* $\mathrm{vol}(A) = N_1 \cdots N_d$. Obviously $|A| \leq \mathrm{vol}(A)$. A is *proper* if all the above elements are different, in which case $|A| = \mathrm{vol}(A) = N_1 \cdots N_d$. Realize that $A+A$ is another generalized arithmetic progression, namely

$$A + A = \{2a_0 + a_1 n_1 + a_2 n_2 + \cdots + a_d n_d : 0 \leq n_j < 2N_j - 1 \text{ for } j = 1, \ldots, d\}$$

(this is not necessary proper, even if A was). Nevertheless $|A+A| \leq (2N_1-1)\cdots(2N_d-1)$ and if A is proper then $|A+A| < 2^d|A|$, so a proper generalized arithmetic progression is another example of small doubling (that is $A+A$ small). This is the ultimate example, as we will see in Chapter 10 that *if $A+A$ is 'small' then A is a 'large' subset of a generalized arithmetic progression of 'small' dimension and volume.*

In the rest of this chapter we study some consequences of small doubling. Some of them would follow from the structure theorem described above, but here we prove them without referring to it.

Theorem 1.4: Let A, B and C be finite subsets of the real numbers. We have

$$|A-C||B| \leq |A-B||B-C|.$$

Proof: For all $d \in A-C$ we fix a representation $d = a_d - c_d$ where $a_d \in A$ and $c_d \in C$. We define a map between $(A-C)\times B$ and $(A-B)\times(B-C)$, namely a pair (d,b), where $d \in A-C$ and $b \in B$ is mapped to the pair $(a_d - b, b - c_d)$, where well visibly $a_d - b \in A-B$ and $b - c_d \in B-C$. We show that this map is injective, that is an image determines its unique inverse. Indeed, if (u,v) is an image under this map then $u+v$ should be d, and than a_d or c_d determine b by $b = a_d - u$ or $b = c_d + v$. Injectivity implies that the size of the image space should be bigger or equal to the size of the domain, which is precisely the stated inequality.

□

We want to stress (though this type of remarks are usually left to the reader) that Theorem 1.4 is valid in any group, we not even used commutativity in the proof. Another remark, one can recognize a kind of triangle inequality in this statement. Indeed, if $\Delta(A,B)$ is defined as $\log \frac{|A-B|}{\sqrt{|A||B|}}$ then the present theorem states $\Delta(A,C) \leq \Delta(A,B) + \Delta(B,C)$.

Putting $C = A$ and than $B = -A$ into Theorem 1.4 we get

Corollary 1.5:

$$|A-A| \le \frac{|A-B|^2}{|B|}, \text{ for all } B, \tag{1.1}$$

$$\frac{|A-A|}{|A|} \le \frac{|A+A|^2}{|A|^2}. \tag{1.2}$$

□

This last inequality (1.2) reads as if A has small doubling then $A-A$ is small. This is just the first of a host of similar statements.

Corollary 1.6:

$$\frac{|mA-nA|}{|A|} \le \left(\frac{|3A|}{|A|}\right)^{m+n-2} \left(\frac{|2A|}{|A|}\right)^{m+n-4}, \text{ for all } m,\, n \ge 2, \tag{1.3}$$

$$\frac{|nA|}{|A|} \le \left(\frac{|3A|}{|A|}\right)^{n-2} \left(\frac{|2A|}{|A|}\right)^{n-3}, \text{ for all } n \ge 3. \tag{1.4}$$

Proof: Put rA to A, $-A$ to B, and sA or $-sA$ to C in Theorem 1.4, where r, $s \ge 1$ are integers. We get

$$|(r+s)A||A| \le |(r+1)A||sA-A|, \tag{1.5}$$

$$|rA-sA||A| \le |(r+1)A||(s+1)A|. \tag{1.6}$$

Note that (1.5) is trivially true for $r=0$ as well, because $|sA| \le |sA-A|$. Next put $r=n-2$, $s=2$ to (1.5), where $n \ge 2$, and $r=2$, $s=1$ to (1.6), we get

$$|nA||A| \le |(n-1)A||2A-A|, \quad |2A-A||A| \le |3A||2A|,$$

and the two inequalities together give

$$\frac{|nA|}{|A|} \le \frac{|(n-1)A|}{|A|}\frac{|3A|}{|A|}\frac{|2A|}{|A|}. \tag{1.7}$$

(1.4) follows from (1.7) by induction, and (1.3) follows by putting (1.4) into (1.6).

□

If A is a generalized arithmetic progression then any combination $mA-nA$ is small. Corollary 1.6 reads as if $2A$ and $3A$ are small then any combination $mA-nA$ is small, however, we need an extra twist to estimate $|3A|$ by $|2A|$. At least we see that if $2A$ and $3A$ are both small then A shares some property with generalized arithmetic progressions. Next we show another similar statement, mainly to illustrate another argument.

Theorem 1.7: For any finite subsets A and B of the real numbers there is a set $X \subset B$ such that $|X| \le |A+B|/|A|$ and $B \subset A-A+X$.

Proof: We can name X directly. Let $X \subset B$ be the largest subset such that the sets $A + x : x \in X$ are mutually disjoint. As $A + x \subset A + B$ for all $x \in B$ and $|A + x| = |A|$ we have $|X||A| \leq |A + B|$ as is stated. Now for any $b \in B$ we have $(A + b) \cap (A + x) \neq \emptyset$ with some $x \in X$, otherwise b could be added to X. That is there are elements a_1, $a_2 \in A$ such that $a_1 + b = a_2 + x$, or $b = a_2 - a_1 + x \in A - A + X$. This completes the proof.

□

Note that writing $-A$ in place of A we get the same statement with $|X| \leq |A - B|/|A|$ as an upper bound. This theorem states that whatever A and B are, B can be covered with a few translates of $A - A$. As a typical application of Theorem 1.7 we show

Corollary 1.8: If A is a finite subset of the real numbers then

$$\frac{|mA - nA|}{|A|} \leq \frac{|A - A|}{|A|} \left(\frac{|2A - 2A|}{|A|} \right)^{m+n-2}, \text{ for all } m, n \geq 1.$$

Proof: Apply Theorem 1.7 with $B = 2A - A$. We get a set $X \subset 2A - A$ such that $|X| \leq |2A - 2A|/|A|$ and

$$2A - A \subset A - A + X. \tag{1.8}$$

It is easy to see that $C \subset D$ implies $A + C \subset A + D$. Adding A to both sides of (1.8), and then applying (1.8) again we get

$$3A - A \subset 2A - A + X \subset A - A + 2X,$$

and repeating the argument several times we arrive at

$$mA - A \subset A - A + (m - 1)X. \tag{1.9}$$

From (1.8) we also have

$$A - 2A \subset A - A - X. \tag{1.10}$$

Subtracting A from both sides of (1.9) and applying (1.10) we get

$$mA - 2A \subset A - 2A + (m - 1)X \subset A - A + (m - 1)X - X.$$

Repeating the same step several times we arrive at

$$mA - nA \subset A - A + (m - 1)X - (n - 1)X.$$

Checking sizes this means

$$|mA - nA| \leq |A - A||(m - 1)X||(n - 1)X| \leq |A - A||X|^{m+n-2}$$

and the statement follows from $|X| \leq |2A - 2A|/|A|$.

□

Next we put the last brick to our building, by giving a simple, elementary proof of Plünnecke's theorem. We start with a lemma.

Lemma 1.9: Let A and B be finite subsets of the real numbers. If for a $c > 0$ we have

$$|A + B| = c|A|, \text{ while } |A' + B| \geq c|A'| \text{ for all } A' \subset A, \tag{1.11}$$

then for all finite subset C of the real numbers we have

$$|A + B + C| \leq c|A + C|. \tag{1.12}$$

Proof: We use induction on the size of C. If $|C| = 1$, that is $C = \{x\}$ then

$$|A + B + x| = |A + B| = c|A| = c|A + x|,$$

thus the statement is obvious. Suppose now that the statement (1.12) is true for a non–empty set C, and let $x \notin C$ be an arbitrary real number, finally let $C' = C \cup \{x\}$. We prove the statement for C' as well. Study the right hand side of (1.12) first. We define

$$A' = \{a \in A : \, a + x \in A + C\} = A \cap (A + C - x) \subset A.$$

We have $A' + x = (A + x) \cap (A + C)$ and $A + C' = (A + C) \cup (A + x)$ and so

$$|A + C'| = |A + C| + |A + x| - |A' + x| = |A + C| + |A| - |A'|.$$

As $A' + x \subset A + C$ implies also $A' + B + x \subset A + B + C$ and so

$$A + B + C' = (A + B + C) \cup (A + B + x) = (A + B + C) \cup [(A + B + x) \setminus (A' + B + x)],$$

moreover

$$|A+B+C'| \leq |A+B+C|+|A+B+x|-|A'+B+x| = |A+B+C|+|A+B|-|A'+B|.$$

We use the induction hypothesis (1,12) to estimate the first term on the right hand side, and the two conditions (1.11) to the other two terms, getting

$$|A + B + C'| \leq c|A + C| + c|A| - c|A'| = c|A + C'|,$$

as is required to complete the induction.

□

There are two important consequences of this lemma. We can prove the sum version of the triangle inequality, Theorem 1.4, and we can prove Plünnecke's inequality.

Theorem 1.10: Let A, B and C be finite subsets of the real numbers. We have

$$|A + C||B| \leq |A + B||B + C|.$$

Proof: Let $U \subset B$ a subset such that the ratio $|U + A|/|U| = c$ is minimal. We have for any $U' \subset U$ that $|U' + A|/|U'| \geq c$. The condition of Lemma 1.9 is satisfied for the choice

$$A \mapsto U, \quad B \mapsto A, \quad C \mapsto C,$$

getting

$$|U + A + C| \leq c|U + C|.$$

Obviously we have $|A + C| \leq |U + A + C|$, $|U + C| \leq |B + C|$ and $c \leq |A + B|/|B|$, as B itself is considered when the minimum is selected. The theorem follows from these last four inequalities.

□

Writing $C = A$ and $B = -A$ we get the companion of Corollary 1.5, more exactly that of (1.2), namely

$$\frac{|A + A|}{|A|} \leq \frac{|A - A|^2}{|A|^2}. \tag{1.13}$$

Writing $C = (h - 1)B$ to (1.12) we get $|A + hB| \leq c|A + (h - 1)B|$, and an induction gives the following result.

Corollary 1.11: Let A and B be finite subsets of the real numbers. If for a $c > 0$ we have

$$|A + B| = c|A|, \text{ while } |A' + B| \geq c|A'| \text{ for all } A' \subset A,$$

then for all positive integer h we have

$$|A + hB| \leq c^h|A|.$$

□

Theorem 1.12: (Plünnecke) Let B be a finite subset of the real numbers, satisfying $|B + B| \leq c|B|$ (or $|B - B| \leq c|B|$). For all non–negative integers m and n we have

$$|mB - nB| \leq c^{m+n}|B|.$$

Proof: Let $\emptyset \neq A \subset B$ (or $\emptyset \neq A \subset -B$) be one of the sets that minimizes the ratio $|A+B|/|A|$, and write $|A+B| = c'|A|$. Obviously $c' \leq c$. By this definition $|A'+B| \geq c'|A'|$ for all $A' \subset A$, thus the conditions of Corollary 1.11 are satisfied. We have

$$|hB| \leq |A + hB| \leq (c')^h|A| \leq c^h|B|,$$

proving the statement for the case of $m = 0$ or $n = 0$. For the general statement we appeal to Theorem 1.4. The roles are distributed as follows.

$$A \mapsto mB, \quad B \mapsto -A, \quad C \mapsto nB.$$

Theorem 1.4 and corollary 1.10 imply that

$$|mB - nB||A| \le |A + mB||A + nB| \le (c')^m|A|(c')^n|A| \le c^{m+n}|B||A|,$$

and the theorem follows by simplifying with $|A|$.

□

We close this chapter by another argument, that, for example, provides a weaker result than (1.13), but interesting anyway, and uses some important notions relevant in later chapters. First we introduce *representation functions*, whenever we speak about an operation with sets, like $A \pm B$, $A \cdot B$ or A/B, etc, we write

$$r_{A\pm B}(x) = \#\{(a,b) \in A \times B : \, x = a \pm b\},$$
$$r_{A\cdot B}(x) = \#\{(a,b) \in A \times B : \, x = ab\},$$
$$r_{A/B}(x) = \#\{(a,b) \in A \times B : \, x = a/b\}, \text{ etc.}$$

There are some fairly trivial observations about the first and second means of representation functions.

$$\sum_x r_{A+B}(x) = |A||B|,$$

and similarly with the other representation functions mentioned above.

$$\sum_x r_{A+B}(x)^2 = \#\{a_1 + b_1 = a_2 + b_2\} = \#\{a_1 - b_2 = a_2 - b_1\} =$$
$$= \#\{a_1 - a_2 = b_2 - b_1\} = \sum_x r_{A-B}(x)^2 = \sum_x r_{A-A}(x)\, r_{B-B}(x).$$

Theorem 1.13: Let A and B be finite subsets of the real numbers, and $x \in A + B$. We have

$$r_{A+B}(x) \le \frac{|A-B|^2}{|A+B|}.$$

Proof: It is very similar to the proof of Theorem 1.4. For any $s \in A + B$ we fix a representation $s = a_s + b_s$; $a_s \in A$, $b_s \in B$. We define an injective map from $R_x \times (A + B)$ to $(A - B) \times (A - B)$, where R_x denotes the set of representations of x as $x = a + b$: $a \in A$, $b \in B$, i.e. R_x is the set of such pairs (a, b). We map $((a, b), s) \in R_x \times (A + B)$ to $(a - b_s, a_s - b) \in (A - B) \times (A - B)$. This is indeed injective, if (u, v) is in the image then $s = x - u + v$ is determined, than a_s and b_s are determined, than $a = u + b_s$ and $b = a_s - v = x - a$ are determined. Any point in the image has a unique inverse, the statement of the theorem follows from comparing the size of the domain and the image.

□

Lemma 1.14 (Cauchy–Schwarz inequality): For any finite sets of real numbers x_j, y_j we have

$$\Big(\sum x_j y_j\Big)^2 \le \sum x_j^2 \sum y_j^2.$$

Proof: We claimed and we prove the statement for finite sums, but it is actually true (with identical proof) for infinite sums or integrals as well, provided the appropriate convergence. Let us abbreviate $S = \sum x_j y_j / \sum y_j^2$, and consider the inequality

$$0 \le \sum (x_j - y_j S)^2 = \sum x_j^2 - 2S \sum x_j y_j + S^2 \sum y_j^2 = \sum x_j^2 - \frac{\left(\sum x_j y_j\right)^2}{\sum y_j^2},$$

and the lemma follows.

□

Corollary 1.15: Let A and B be finite subsets of the real numbers. We have

$$|A + B| \le \frac{|A - B|^3}{|A||B|}, \tag{1.14}$$

especially

$$\frac{|A + A|}{|A|} \le \left(\frac{|A - A|}{|A|}\right)^3.$$

Proof: After this preparation we can express the argument in a chain of equations and inequalities. We start with Lemma 1.14 and in due course we use Theorem 1.13 as well.

$$\begin{aligned}(|A||B|)^2 = \left(\sum_x r_{A-B}(x)\right)^2 &\le |A - B| \sum_x r_{A-B}(x)^2 = \\ &= |A - B| \sum_x r_{A+B}(x)^2 \le |A - B| \max_y r_{A+B}(y) \sum_x r_{A+B}(x) = \\ &= |A||B||A - B| \max_y r_{A+B}(y) \le |A||B||A - B| \frac{|A - B|^2}{|A + B|},\end{aligned}$$

and the theorem follows from a little rearrangement.

□

CHAPTER 2

SIDON SETS

Recall that in Chapter 1 we named a set A of integers a Sidon set if the sums $a_i + a_j$ were different for all $i \le j$. Finite Sidon sets are examples for a maximal $A + A$, however infinite Sidon sets are also interesting objects to study. The main problem here is to find *dense* Sidon sets, that is either for all integers $N \ge 1$ find the largest n such that there is a Sidon set $A \subset \{1, \dots, N\}$, $|A| = n$, this is the finite case, or to find an infinite Sidon set A of positive integers such that its counting function

$$A(x) = \sum_{\substack{a \le x \\ a \in A}} 1$$

is large in some sense. Obviously any subset of a Sidon set is a Sidon set again, especially any finite segment of an infinite Sidon set is a Sidon set, however, there are better finite constructions. It is possible to prove, that for all $\epsilon > 0$ and all (large) N there is a Sidon set $A \subset \{1, \dots, N\}$ with $|A| > (1 - \epsilon)\sqrt{N}$, while any infinite Sidon set A satisfies

$$\liminf_{x \to \infty} \frac{A(x)}{\sqrt{x}} = 0.$$

The first useful observation is that $a_i + a_j = a_k + a_\ell$ iff $a_j - a_k = a_\ell - a_i$, in other words all sums are different iff all differences (except 0) are different. If A is a Sidon set in $\{1, \dots, N\}$ then all nonzero differences are among $\{1, \dots, N-1\}$. So $|A| = n$ must satisfy

$$\binom{n}{2} \le N - 1, \quad n \le \sqrt{2N} + 1.$$

This is obviously an upper bound for any infinite Sidon set as well, $A(x) \le \sqrt{2x} + 1$. We will improve these bounds soon.

The next useful observation is that we can use the greedy algorithm to construct a Sidon set. We pick the elements of A one by one. Let $a_1 = 1$, $a_2 = 2$ and suppose we have already picked $a_1, \dots, a_n$, where $n \ge 2$. When choosing a_{n+1} we have to avoid some integers, namely $a_i + a_j - a_k$. (Some of these are not positive, and we avoid them automatically.) Nevertheless we have to avoid at most

$$\binom{n+1}{2} n < n^3$$

integers, so picking the smallest possible integer, it will satisfy $a_{n+1} \le n^3$. If $8 \le n^3 \le x < (n+1)^3$ then the first $n+1$ elements of this Sidon set is not bigger than x, that is $A(x) \ge n + 1 > x^{1/3}$. This naturally provides a Sidon set up to N with at least $N^{1/3}$

elements. It maybe surprising that improving the lower bound for the infinite case is very hard, but for the finite case it is not so.

Let p be an odd prime, and choose

$$a_k = 2pk + s_k, \text{ where } s_k \equiv k^2 \pmod{p},\ 1 \le s_k \le p-1,\ k = 1, \dots, p-1.$$

Clearly this is a set of $p-1$ integers below $2p^2$. We show that they form a Sidon set. Suppose that $a_i + a_j = a_k + a_\ell$, that is

$$2pi + s_i + 2pj + s_j = 2pk + s_k + 2p\ell + s_\ell,$$

or

$$2p(i + j - k - \ell) = s_k + s_\ell - s_i - s_j.$$

The right hand side is at most $2(p-1)$ in absolute value, while the left hand side is at least $2p$, unless it is zero. The above equation can only hold if both sides are zero, that is

$$i + j = k + \ell, \quad i^2 + j^2 \equiv k^2 + \ell^2 \pmod{p}.$$

This is the same as

$$j - k = \ell - i, \quad (j-k)(j+k) = j^2 - k^2 \equiv \ell^2 - i^2 = (\ell - i)(\ell + i) \pmod{p}.$$

We conclude that either $j - k = \ell - i = 0$ or $j + k \equiv \ell + i \pmod{p}$. In the first case $j = k$ and $i = \ell$, while the second case together with $j - k = \ell - i$ imply $j \equiv \ell \pmod{p}$ and $k \equiv i \pmod{p}$ and thus $j = \ell$ and $k = i$. In any case $a_i + a_j = a_k + a_\ell$ imply that the pairs $\{i, j\}$ and $\{k, \ell\}$ are the same. We have found a Sidon set of $p-1$ elements up to $2p^2$. The Prime Number Theorem claims that denoting the number of positive primes $p \le x$ by $\pi(x)$ we have

$$\lim_{x \to \infty} \frac{\pi(x)}{x/\log x} = 1, \tag{2.1}$$

and it is an easy exercise to derive from here that for all $1 > \epsilon > 0$ and sufficiently large y there is a prime p satisfying $(1-\epsilon)y < p \le y$. With $y = \sqrt{N/2}$ the above construction provides a Sidon set of $(1-\epsilon)\sqrt{N/2}$ elements up to N.

Theorem 2.1: If $A \subset \{1, \dots, N\}$ is a Sidon set then

$$|A| \le N^{1/2} + N^{1/4} + 1.$$

Proof: Let us write $A = \{a_1 < a_2 < \cdots < a_n\}$. Remember that the differences $a_j - a_i$ are different for all $j > i$. Consider

$$S_1 =: \sum_{1 \le j \le n-1} (a_{j+1} - a_j) = a_n - a_1 \le N.$$

This sum is actually the sum of $n-1$ different integers, and as such, it is at least as big as the sum of the $n-1$ smallest positive integers. The upper bound coming from this argument does not go beyond our earlier knowledge. However, next we consider

$$S_2 =: \sum_{1\le j\le n-2} (a_{j+2}-a_j) = a_n + a_{n-1} - a_2 - a_1 \le 2N.$$

This sum is actually the sum of $n-2$ different integers, and they are also different from the previously considered $n-1$ differences. Since the sum of the two sets of differences are at least as big as the sum of the $2n-3$ smallest integers, we get

$$\frac{(2n-3)(2n-2)}{2} \le S_1 + S_2 \le 3N,$$

which leads to an improvement over the previous bound for n. Obviously we do not have to stop here. Consider

$$S_k =: \sum_{1\le j\le n-k} (a_{j+k}-a_j) = a_n + \cdots + a_{n-k+1} - a_k - \cdots - a_1 \le kN.$$

The number of differences in $S_1 + \cdots + S_k$ are exactly $(n-1) + \cdots + (n-k) = kn - k(k+1)/2 =: D$, and they are all different, their sum is at least as big as the sum of the D smallest integers. We have established the inequality

$$\frac{D(D+1)}{2} \le S_1 + \cdots + S_k \le (1 + \cdots + k)N,$$

and after further simplification

$$k(n - \frac{k+1}{2})^2 \le (k+1)N.$$

The optimal choice for k is given by the inequality

$$N^{1/4} \le k < N^{1/4} + 1,$$

and the last steps are

$$n \le \sqrt{(1+\frac{1}{k})N} + \frac{k+1}{2} \le \sqrt{N}(1+\frac{1}{2k}) + \frac{k+1}{2} \le N^{1/2} + N^{1/4} + 1.$$

□

This is obviously an upper bound for any finite segment of an infinite Sidon set,

Corollary 2.2: If A is an infinite Sidon set of positive integers, then

$$A(x) = \sum_{a\le x,\, a\in A} 1 \le x^{1/2} + x^{1/4} + 1.$$

□

Next we give a construction of a finite Sidon set of size near the upper bound in Theorem 2.1. There are several such constructions, they are all based on some properties of finite fields. We briefly summarize the necessary knowledge.
– For any prime power p^r there is a unique commutative field $\mathbb{F}_{p^r}$ of p^r elements.
– If $r|s$ then $\mathbb{F}_{p^r}$ is a subfield of $\mathbb{F}_{p^s}$.
– The multiplicative group of $\mathbb{F}_{p^r}$ is a cyclic group of order p^r-1.

$\mathbb{F}_p$ is easy to understand, this is $\mathbb{Z}_p$, the ring of residue classes mod p, which is actually a field. From the Fermat–Euler theorem it is well know that $x^p = x$ for all elements of $\mathbb{Z}_p$. In general, $\mathbb{F}_{p^r}$ can be considered as the set of all solutions of the equation $x^{p^r} - x = 0$ (in an algebraic closure of $\mathbb{F}_p$).

Theorem 2.3: For every prime p there is a Sidon set $A \subset \{1, \dots, p^2-1\}$ with $|A| = p$.

Proof: Let $\theta \in \mathbb{F}_{p^2}$ be a generator of the multiplicative group of $\mathbb{F}_{p^2}$. This means

$$\mathbb{F}_{p^2} = \{0, \theta, \theta^2, \dots, \theta^{p^2-1} = 1\}.$$

Realize that we use $0, 1$ here as the zero and the unit elements of the finite field $\mathbb{F}_{p^2}$. The exponents are residue classes mod p^2-1, but we represent the classes with their least positive elements, and we select our set A from these exponents. $\mathbb{F}_p = \{c_1, \dots, c_p\}$ is a subfield of $\mathbb{F}_{p^2}$. Trivially θ is not an element of $\mathbb{F}_p$ as it is a generator of a bigger group, so $\theta + c_j \neq 0$. We are ready to define A, for every $j = 1, \dots, p$ there is an exponent a_j such that

$$\theta^{a_j} = \theta + c_j,$$

and let $A = \{a_1, \dots, a_p\}$ be the set of these exponents. This provides that A is indeed in the given range. Now suppose that $a_i + a_j = a_k + a_\ell$. We have

$$(\theta + c_i)(\theta + c_j) = \theta^{a_i}\theta^{a_j} = \theta^{a_i + a_j} = \theta^{a_k + a_\ell} = \theta^{a_k}\theta^{a_\ell} = (\theta + c_k)(\theta + c_\ell),$$

and after a little rearrangement we arrive at

$$\theta(c_i + c_j) + c_i c_j = \theta(c_k + c_\ell) + c_k c_\ell.$$

As θ is not in $\mathbb{F}_p$, it cannot be a solution of a linear equation with coefficients in $\mathbb{F}_p$, the only possibility is

$$c_i + c_j = c_k + c_\ell, \quad c_i c_j = c_k c_\ell. \tag{2,2}$$

Denoting the common value of the first equation with b and the common value of the second equation with d we can see that both sets $\{c_i, c_j\}$ and $\{c_k, c_\ell\}$ are the two solutions of the quadratic equation $X^2 - bX + d = 0$, the two sets must be the same. This implies that also $\{a_i, a_j\}$ must be the same as $\{a_k, a_\ell\}$. This concludes the proof.

□

For any (large) integer N we can find a prime p such that $(1-\epsilon)\sqrt{N} < p \leq \sqrt{N}$ by (2.1).

Corollary 2.4: For any $1 > \epsilon > 0$ there is an N_0 such that for all $N \geq N_0$ there is a Sidon set $A \subset \{1, \dots, N\}$ with $|A| > (1-\epsilon)\sqrt{N}$.

□

We close this chapter (and subject) with showing that finite segments of an infinite Sidon set must sometimes be much smaller than it is in Corollary 2.4.

Theorem 2.5: (Erdős) Let A be an infinite Sidon set of positive integers. We have

$$\liminf_{x\to\infty} \frac{A(x)}{\sqrt{\frac{40x}{\log x}}} \le 1.$$

Proof: We prove the theorem in the following equivalent form. For all integer $N \ge 2$ we have

$$\tau(N) =: \inf_{n\ge N} \frac{A(n)}{\sqrt{\frac{40n}{\log n}}} \le 1. \tag{2.3}$$

A simple consequence of the definition is that for all $n \ge N$

$$A(n) \ge \tau(N)\sqrt{\frac{40n}{\log n}}. \tag{2.4}$$

We fix an $N \ge 2$ and consider the elements $a \in A$ in the intervals $(k-1)N < a \le kN$ for all $k = 1, \dots, N$. The number of elements of A in this interval is $d(k) = A(kN) - A((k-1)N)$. The difference between two elements of the same interval is $< N$, and being a Sidon set, the differences of the different elements of A are all different. This implies

$$\sum_{\substack{k=1\\ d(k)>1}}^{N} \binom{d(k)}{2} < N,$$

and so we have

$$\sum_{k=1}^{N} d(k)^2 \le \sum_{\substack{k=1\\ d(k)\le 1}}^{N} 1 + 2\sum_{\substack{k=1\\ d(k)>1}}^{N} d(k)(d(k)-1) < 5N.$$

From the Cauchy–Schwarz inequality, Lemma 1.14, we conclude

$$\left(\sum_{k=1}^{N} \frac{d(k)}{\sqrt{k}}\right)^2 \le \sum_{k=1}^{N} d(k)^2 \sum_{k=1}^{N} \frac{1}{k} < 5N \sum_{k=1}^{N} \frac{1}{k}. \tag{2.5}$$

Recall Abel summation, which is a simple rearrangement of a finite sum. For arbitrary complex numbers $b_0, b_1, \dots, b_N$ and $c_1, \dots, c_{N+1}$ we have

$$\sum_{k=1}^{N} (b_k - b_{k-1})c_k = \sum_{k=1}^{N} b_k(c_k - c_{k+1}) + b_N c_{N+1} - b_0 c_1.$$

In the following long but straightforward computation we use this identity twice together with (2.4) and the estimate

$$\sqrt{k}-\sqrt{k-1}=\frac{1}{\sqrt{k}+\sqrt{k-1}}>\frac{1}{2\sqrt{k}}.$$

We have

$$\begin{aligned}\sum_{k=1}^{N}\frac{d(k)}{\sqrt{k}}=\sum_{k=1}^{N}\frac{A(kN)-A((k-1)N)}{\sqrt{k}}=\\ =\sum_{k=1}^{N}A(kN)\left(\frac{1}{\sqrt{k}}-\frac{1}{\sqrt{k+1}}\right)+\frac{A(N^2)}{\sqrt{N+1}}\geq\\ \geq\sum_{k=1}^{N}\tau(N)\sqrt{\frac{40kN}{\log kN}}\left(\frac{1}{\sqrt{k}}-\frac{1}{\sqrt{k+1}}\right)+\frac{\tau(N)\sqrt{\frac{40N^2}{\log N^2}}}{\sqrt{N+1}}\geq\\ \geq\tau(N)\sqrt{\frac{20N}{\log N}}\left(\sum_{k=1}^{N}\sqrt{k}(\frac{1}{\sqrt{k}}-\frac{1}{\sqrt{k+1}})+\frac{\sqrt{N}}{\sqrt{N+1}}\right)=\\ =\tau(N)\sqrt{\frac{20N}{\log N}}\sum_{k=1}^{N}\frac{\sqrt{k}-\sqrt{k-1}}{\sqrt{k}}\geq\tau(N)\sqrt{\frac{5N}{\log N}}\sum_{k=1}^{N}\frac{1}{k}.\end{aligned}$$

Finally writing this into (2.5) we establish

$$\tau(N)\leq\frac{\log N}{\sum_{k=1}^{N}\frac{1}{k}},$$

and (2.3), as well as the theorem follows from

$$\sum_{k\leq x}\frac{1}{k}>\int_1^x\frac{dt}{t}=\log x.$$

□

CHAPTER 3

SUMS OF RESIDUE CLASSES

Remember that the proof of Theorem 1.1 used the ordering of the real numbers. That proof certainly do not generalizes to finite groups, for example, to the group of residue classes. In this chapter we study this question, however, we only concentrate on getting a lower bound for the size of $A+B$ if A and B are subsets of the (additive) group of residue classes mod m. To describe all extremal cases is an open question in general, but is solved for $\mathbb{Z}_p$, the group of residue classes mod p, p being a prime (Vosper's Theorem). Although the proof is not particularly complicated, we do not spend time for working out the details.

Deriving the lower bound is especially simple when $|A|+|B|$ is large. Recall the general definition of representation functions, $r_{A+B}(z)$ denotes the number of pairs $(a,b) \in A \times B$ such that $z = a+b$.

Lemma 3.1: Let Z be a finite Abelian group, and let A and B be subsets of Z with $|A|+|B| \geq |Z|+t$. Then $r_{A+B}(z) \geq t$ for all $z \in Z$.

Proof: Consider the translates $z-B = \{z-b : b \in B\}$. Since

$$|Z| \geq |A \cup (z-B)| = |A| + |z-B| - |A \cap (z-B)| = |A| + |B| - |A \cap (z-B)|,$$

it follows that

$$|A \cap (z-B) \geq |A| + |B| - |Z| \geq t.$$

There exist (at least) t distinct elements $a \in A$ and t distinct elements $b \in B$ such that $a = z-b$, proving that $r_{A+B}(z) \geq t$. Since $z \in Z$ was arbitrary this concludes the lemma. □

The $t=1$ special case says that

Corollary 3.2: Let Z be a finite Abelian group, and let A and B be subsets of Z with $|A|+|B| > |Z|$. Then $A+B = Z$.

□

Next, following the footsteps of Freeman Dyson, the physicist, we introduce the *e–transform* of an ordered pair (A,B) of nonempty subsets of an Abelian group Z. Let $e \in Z$ be arbitrary, and define the e–transform of (A,B) as the pair $(A(e), B(e))$ of subsets of Z given by

$$A(e) = A \cup (B+e), \quad B(e) = B \cap (A-e).$$

We establish some simple properties of the e–transform.

Lemma 3.3: Let A and B be nonempty subsets of the (Abelian) group Z, and let e be any element of Z. We have

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

$$A(e) \setminus A = e + (B \setminus B(e)), \tag{3.2}$$

$$|A(e)| + |B(e)| = |A| + |B|, \text{ if } A \text{ and } B \text{ are finite.} \tag{3.3}$$

Proof: To accept (3.1) realize that

$$A(e) + B(e) = (A + B(e)) \cup (B + e + B(e)) \subset (A + B) \cup (B + e + A - e) \subset A + B.$$

To prove (3.2) we observe that

$$A(e) \setminus A = (B + e) \setminus A = \{b + e : b \in B,\, b + e \notin A\} = \\ = e + \{b \in B : b \notin A - e\} = e + \{b \in B : b \notin B(e)\} = e + (B \setminus B(e)).$$

Finally for (3.3) note that clearly $A \subset A(e)$ and $B(e) \subset B$. For finite sets we have (using (3.2) in due course)

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

□

Theorem 3.4: Let $m > 1$ be fixed. If A and B are nonempty subsets of $\mathbb{Z}_m$ such that $0 \in B$, and $\gcd(b, m) = 1$ for all $b \in B \setminus \{0\}$, then

$$|A + B| \geq \min\{m,\, |A| + |B| - 1\}.$$

Proof: If $|A| + |B| - 1 \geq m$ then the statement follows from Corollary 3.2, we can suppose $|A| + |B| \leq m$. We prove by induction, if $|B| = 1$ then $|A + B| = |A| = |A| + |B| - 1$, the statement is true. Let $|B| \geq 2$ and we are going to prove the statement for B if it is true for any $B' \subset \mathbb{Z}_m$ with $|B'| < |B|$. We can select a $b \in B$, $b \neq 0$. Now $A + b \neq A$, otherwise starting from an arbitrary $a_0 \in A$ we find that $a_1 = a_0 + b \in A$, and $a_2 = a_1 + b = a_0 + 2b \in A$ and so on, all $a_0 + jb \in A$. Using that $\gcd(b, m) = 1$ we find $\mathbb{Z}_m \subset A$ which is not possible by $|A| + |B| \leq m$, $|B| \geq 2$. We can take an $e \in A$ such that $e + b \notin A$. We want to use the induction hypothesis for the pair $(A(e), B(e))$. As $A(e) \supset A$ and $0 \in B \cap (A - e)$ they are not empty, also $b \notin A - e$ so $B(e) \subsetneqq B$, we conclude using (3.3) and (3.1) that

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

□

Theorem 3.5: (Cauchy–Davenport Theorem) Let p be a fixed prime . If A and B are nonempty subsets of $\mathbb{Z}_p$ then

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

Proof: Apply Theorem 3.4 with $B' = B - b$ for an arbitrary fixed element $b \in B$. The conditions are obviously satisfied.

□

Theorem 3.6: (Cauchy–Davenport Theorem) Let p be a fixed prime, and $h \geq 2$ be an integer. If $A_1, \ldots, A_h$ are nonempty subsets of $\mathbb{Z}_p$ then

$$|A_1 + \cdots + A_h| \geq \min\{p, |A_1| + \cdots + |A_h| - h + 1\}.$$

Proof: We prove by induction on h. The case $h = 2$ is precisely Theorem 3.5. Let $h \geq 3$, and suppose that the result is true for any $h-1$ subsets of $\mathbb{Z}_p$. Let $A_1 \ldots, A_h$ be nonempty subsets of $\mathbb{Z}_p$, and write $B = A_1 + \cdots + A_{h-1}$. By the induction hypothesis

$$|B| = |A_1 + \cdots + A_{h-1}| \geq \min\{p, |A_1| + \cdots + |A_{h-1}| - h + 2\},$$

and we have

$$\begin{aligned} |A_1 + \cdots + A_h| = |B + A_h| &\geq \min\{p, |B| + |A_h| - 1\} \geq \\ &\geq \min\{p, \min\{p, |A_1| + \cdots + |A_{h-1}| - h + 2\} + |A_h| - 1\} = \\ &= \min\{p, |A_1| + \cdots + |A_h| - h + 1\}. \end{aligned}$$

This completes the proof.

□

The Cauchy–Davenport theorem for $A = B$ claims that $|A + A| \geq \min\{p, 2|A| - 1\}$. It is a natural question to ask how this bound changes if we only add different elements of A. We (temporarily to the rest of this chapter) denote by $A \oplus A$ the set of sums $a + a'$ where $a, a' \in A$, $a \neq a'$. In case of Theorem 1.1, that is in the group of real numbers, one can quickly come up with the sharp lower bound $2|A| - 3$ by the very same method. It maybe surprising that the corresponding result for residue classes was conjectured by Erdős and Heilbronn in 1964 and was proved only 30 years later by Dias da Silva and Hamidoune, using rather complicated arguments. Here we include a different proof due to Alon, Nathanson and Ruzsa, we do it mainly to develop the polynomial method and to illustrate its use. The following lemma is the key step of the method. It is valid for polynomials over any field, but we claim and prove it only over the finite field of p elements, that is over the residue classes mod p.

Lemma 3.7: Let p be a prime, $h \geq 1$ be an integer, and $A_1, \ldots, A_h$ be nonempty subsets of $\mathbb{Z}_p$ with $|A_j| = k_j$. Let $f(x_1, \ldots, x_h)$ be a polynomial with coefficients in $\mathbb{Z}_p$ and of degree at most $k_j - 1$ in the variable x_j, for $j = 1, \ldots, h$. If $f(a_1, \ldots, a_h) = 0$ for all choices of $a_j \in A_j$, then $f(x_1, \ldots, x_h)$ is the zero polynomial.

Proof: We prove by induction on h. The case $h = 1$ is the *degree theorem*, that is a nonzero polynomial of degree n has at most n roots in a field. In the argument below $f = 0$ is understood in the given field, that is, it actually means $f \equiv 0 \pmod{p}$.

Let $h \geq 2$ and assume that the lemma is true for polynomials in less than h variables. We can write our polynomial according to the powers of x_h, that is

$$f(x_1, \ldots, x_h) = \sum_{j=0}^{k_h - 1} f_j(x_1, \ldots, x_{h-1}) x_h^j,$$

where for each $j = 0, \ldots, k_h - 1$ the polynomial $f_j(x_1, \ldots, x_{h-1})$ has degree at most $k_i - 1$ in the variable x_i. Fix an arbitrary

$$(a_1, \ldots, a_{h-1}) \in A_1 \times \cdots \times A_{h-1},$$

and write

$$g(x) = f(a_1, \ldots, a_{h-1}, x) = \sum_{j=0}^{k_h - 1} f_j(a_1, \ldots, a_{h-1}) x^j.$$

$g(x)$ is a polynomial of degree at most $k_h - 1$ such that $g(a_h) = 0$ for all $a_h \in A_h$. Since this counts k_h distinct roots, it follows that $g(x)$ is the zero polynomial, and so

$$f_j(a_1, \ldots, a_{h-1}) = 0$$

for all

$$(a_1, \ldots, a_{h-1}) \in A_1 \times \cdots \times A_{h-1}.$$

It follows from the induction hypothesis that $f_j(x_1, \ldots, x_{h-1})$ is the zero polynomial for all $j = 0, \ldots, k_h - 1$ and so $f(x_1, \ldots, x_h)$ is the zero polynomial as well.

□

Lemma 3.8: Let p be a prime and let A be a nonempty subset of $\mathbb{Z}_p$ with $|A| = k$. For every $m \geq 0$ there exist a polynomial $g_m(x)$ of degree at most $k - 1$ such that

$$g_m(a) = a^m, \text{ for all } a \in A.$$

Proof: This is a direct consequence of the usual Lagrange interpolation method. Another argument is the following. Consider the coefficients of $g_m(x) = u_0 + u_1 x + \cdots + u_{k-1} x^{k-1}$ as being variables. They must satisfy

$$\begin{aligned}
u_0 + u_1 a_1 + u_2 a_1^2 + \cdots + u_{k-1} a_1^{k-1} &= a_1^m, \\
u_0 + u_1 a_2 + u_2 a_2^2 + \cdots + u_{k-1} a_2^{k-1} &= a_2^m, \\
&\vdots \\
u_0 + u_1 a_k + u_2 a_k^2 + \cdots + u_{k-1} a_k^{k-1} &= a_k^m,
\end{aligned}$$

where $A = \{a_1, \ldots, a_k\}$. This is a system of k linear equations in the k unknowns $u_0, \ldots, u_{k-1}$, and it has a solution if the determinant of the coefficients is nonzero. The

lemma follows immediately from the observation that this determinant is a Vandermonde–determinant

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{k-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{k-1} \\ & & \vdots & & \\ 1 & a_k & a_k^2 & \dots & a_k^{k-1} \end{vmatrix} = \prod_{1 \le i < j \le k} (a_j - a_i) \neq 0.$$

□

Theorem 3.9: Let p be a prime number and let A and B be nonempty subsets of $\mathbb{Z}_p$ such that $|A| \neq |B|$. We have

$$|A \oplus B| \geq \min\{p,\, |A| + |B| - 2\}.$$

Proof: Let $|A| = k$ and $|B| = \ell$, and we can assume that $\ell < k \leq p$. We can also assume that $k + \ell - 2 \leq p$, because the case $k + \ell - 2 = p$ clearly implies the case $k + \ell - 2 > p$ by forgetting some elements of B, say. Writing $C = A \oplus B$, we have to prove $|C| \geq k + \ell - 2$. We prove by contradiction. Suppose $|C| \leq k + \ell - 3$, and let us choose $r \geq 0$ such that $r + |C| = k + \ell - 3$. We construct a polynomial in two variables, that is in $\mathbb{Z}_p[x, y]$, which vanishes on $A \times B$. We start by

$$f(x, y) = (x - y)(x + y)^r \prod_{c \in C} (x + y - c). \tag{3.5}$$

It has degree at most $1 + r + |C| = k + \ell - 2$ in both variables, as well as in total. That is, there exist coefficients $u_{m,n} \in \mathbb{Z}_p$ such that

$$f(x, y) = \sum_{m+n \leq k+\ell-2} u_{m,n} x^m y^n = (x - y)(x + y)^{k+\ell-3} + \text{ lower order terms.} \tag{3.6}$$

Using the binomial theorem we can read from here that the coefficient $u_{k-1,\ell-1}$ of the term $x^{k-1}y^{\ell-1}$ in $f(x, y)$ is actually

$$u_{k-1,\ell-1} = \binom{k+\ell-3}{k-2} - \binom{k+\ell-3}{k-1} = \frac{(k-\ell)(k+\ell-3)!}{(k-1)!(\ell-1)!} \neq 0, \tag{3.7}$$

because $1 \leq \ell < k \leq p$ and $1 \leq k + \ell - 3 < p$. By Lemma 3.8, for every $m \geq k$ there is a polynomial $g_m(x)$ of degree at most $k-1$ such that $g_m(a) = a^m$ for all $a \in A$, and similarly for every $n \geq \ell$ there is a polynomial $h_n(y)$ of degree at most $\ell - 1$ such that $h_n(b) = b^n$ for all $b \in B$. Writing these polynomials into (3.6) in place of x^m or y^n whenever $m \geq k$ or $n \geq \ell$ we arrive at a new polynomial $f^*(x, y)$ of degree at most $k-1$ in x and at most $\ell - 1$ in y. If $(a, b) \in A \times B$ then by the construction and by (3.5) we have $f^*(a, b) = f(a, b) = 0$. Lemma 3.7 says that $f^*(x, y)$ must be the zero polynomial. On the other hand the highest degree term, $x^{k-1}y^{\ell-1}$ can only come from terms $u_{m,n}x^m y^n$ of (3.6), where $m \geq k-1$ and $n \geq \ell - 1$. As $m + n \leq k + \ell - 2$ there is only one such term, namely $u_{k-1,\ell-1}x^{k-1}y^{\ell-1}$

itself, which is nonzero by (3.7), so $f^*(x, y)$ is not the zero polynomial. This contradiction proves the theorem.

□

Corollary 3.10: (The Erdős–Heilbronn conjecture) Let p be a prime and A be a nonempty subset of $\mathbb{Z}_p$. We have

$$|A \oplus A| \geq \min\{p, 2|A| - 3\}.$$

Proof: We can suppose that $|A| \geq 2$. Fix an arbitrary element of A, say a, and write $B = A \setminus \{a\}$. Obviously $A \oplus A = A \oplus B$ and Theorem 3.9 is applicable to the sets A, B, getting

$$|A \oplus A| = |A \oplus B| \geq \min\{p, |A| + |B| - 2\} = \min\{p, 2|A| - 3\}.$$

This completes the proof.

□

CHAPTER 4

ARITHMETIC PROGRESSIONS

Consider the following sets of integers.

$$X = \cup_{m=0}^{\infty} \{2^{2m} \leq n < 2^{2m+1}\},$$
$$Y = \cup_{m=1}^{\infty} \{2^{2m-1} \leq n < 2^{2m}\}.$$

Clearly $X \cup Y$ is a partition of the positive integers, that is, $X \cap Y = \emptyset$, $X \cup Y = \mathbb{N}$, where $\mathbb{N}$ denotes the set of positive integers. We can also think about a partition as a 2–coloring of $\mathbb{N}$, say X is the set of blue integers, Y is the set of red integers. Similarly partition into r disjoint sets is the same as r–coloring, the set of integers of the same color is a color class. As both X and Y contain arbitrarily long gaps, neither of them contains an infinite arithmetic progression. However, we have the following theorem.

Theorem 4.1: (van der Waerden, 1927) Let $k \geq 2$ and $r \geq 2$ be integers. For any r-coloring of the positive integers, at least one color class contains a k–term arithmetic progression.

We prove this theorem later in this chapter. Note that there are several other ways to state this theorem. For example, *let $r \geq 2$ be an integer. For any r–coloring of the positive integers, at least one color class contains arbitrarily long arithmetic progressions.* We can also finalize the statement. *Given integers $k \geq 2$ and $r \geq 2$, there exists a positive integer $N = N(k, r)$ such that for any r–coloring of the set $\{1, \ldots, N\}$, there is a monochromatic arithmetic progression of k–terms.* Here monochromatic means that all terms are in the same color class, they have the same color.

Edős and Turán conjectured that van der Waerden's theorem must be true for density reasons, that is any large subset of the integers should contain long arithmetic progressions. Their conjecture was proved by Szemerédi in 1975.

Theorem 4.2: (Szemerédi, 1975) Given an integer $k \geq 2$ and real number $1 > \delta > 0$, there is a positive integer $N(k, \delta)$ such that if $N \geq N(k, \delta)$ and $A \subset \{1, \ldots, N\}$, $|A| > \delta N$, then A contains a k–term arithmetic progression.

This is a deep theorem, one of the biggest achievement of mathematics. Although there are several different proofs by now, neither of them is suitable for this lecture note. We stated the statement here (without proof) for understanding some motivation of later results. Theorem 4.2 is also one of the important ingredients in the following recent result of Green and Tao.

Theorem 4.3: For any $k \geq 2$, the set of positive primes contains a k–term arithmetic progression.

Needless to say, we do not prove this statement either. Not only the proof rests on Szemerédi's theorem, it also uses deep results from analytic number theory. As is well

known, the set of primes is not dense enough for a direct application of Theorem 4.2. There is another conjecture of Erdős, which if true, it would imply Theorem 4.3.

Conjecture 4.4: If A is an infinite set of positive integers satisfying

$$\sum_{a\in A}\frac{1}{a}=\infty,$$

then A contains a k–term arithmetic progression for any $k\geq 2$.

At this point it is appropriate to discuss an early result, due to Schur. On one hand, Schur's theorem is interesting itself, it is not a consequence of van der Waerden's theorem, on the other hand, not only its proof is simple and elegant, but it gives some hint for the more complex argument coming after. Realize, that from density point of view the set of positive odd integers is a large set, it is half of the set of all integers. In spite of this, the equation $x+y=z$ cannot be solved in this large set. However, we have the following result.

Theorem 4.5: (Schur, 1916) Let $r\geq 1$ be an integer. For any r–coloring of $\mathbb{N}$ there is a monochromatic solution of $x+y=z$, that is a monochromatic triple $x,y,z\in\mathbb{N}$ such that $x+y=z$.

Proof: One can transform the problem to a somewhat more symmetric form by looking for a triple $a>b>c$ such that $x=a-b$, $y=b-c$, and $z=a-c$ are monochromatic. We prove the finalized version of this statement. *There is an $N(r)$ such that for all $N\geq N(r)$ and all r–coloring of the integers $\{1,\dots,N\}$, one can find integers $1\leq c<b<a\leq N$ such that $a-b$, $b-c$, and $a-c$ are monochromatic.*

Consider the complete graph on N vertices, and color the edge $\{a,b\}$ by the color of $|a-b|$. A monochromatic triangle $\{a,b,c\}$ is exactly what we want, if $a>b>c$ then $a-b$, $b-c$, and $a-c$ have the same color. The above statement than reads as follows. *There is an $N(r)$ such that for all $N\geq N(r)$ and all r–coloring of the complete graph on N vertices, one can find a monochromatic triangle.*

We prove by induction on r. The statement is trivial for $r=1$ with $N(1)=3$. Since the essential idea is already visible in the case of $r=2$, we work out that as well. Let $N\geq N(2)=2N(1)=6$. An arbitrary vertex a is the endpoint of $N-1\geq 5$ edges, they are 2–colored, so there are three edges of the same color, say red. Consider three vertices b,c,d attached to a with a red edge. If one of the edges $\{b,c\},\{c,d\},\{b,d\}$ is red, say $\{b,c\}$ is red, then $\{a,b,c\}$ is a red triangle, otherwise $\{b,c,d\}$ is a blue triangle.

Let $r\geq 2$ and suppose the statement is true for r with some $N(r)$. Let $N\geq N(r+1)=(r+1)(N(r)-1)+2$, and color the complete graph on N vertices by $r+1$ colors. An arbitrary vertex a is the endpoint of $N-1\geq(r+1)(N(r)-1)+1$ edges, there must be $N(r)$ of them in the same color, say red. Consider $N(r)$ vertices attached to a with a red edge. If there is a red edge between two of them, then a with that two vertices form a red triangle, otherwise the complete graph on that $N(r)$ vertices is colored with only r colors. By the induction hypothesis, it contains a monochromatic triangle. The proof is complete. It is recommended to compute $N(r)$ given by this proof, one gets a reasonable bound.

□

The arithmetic progression $\{a + jd : j = 0, \ldots, k-1\}$ is, in some sense, similar to an interval $\{1, \ldots, k\}$, indeed it is a dilate by d and translate by $a - d$. It is not hard to generalize van der Waerden's theorem into this direction. Theorem 4.1 follows from the theorem below by taking $S = \{1, \ldots, k\}$.

Theorem 4.6: (Grünwald) Let $r \geq 2$ be an integer and S be a finite nonempty set of positive integers. For any r–coloring of $\mathbb{N}$ there exist integers $a \geq 1$ and b such that $a \cdot S + b$ is monochromatic.

Proof: The proof is by induction on the size of S, but by no means can it be considered simple. For any finite set S of positive integers we state a claim, namely

A(S): for each $r \geq 2$ there is an $N(r)$ such that r–coloring of the integers, the set $\{1, \ldots, N\}$, $N \geq N(r)$ contains a monochromatic subset of the form $a \cdot S + b$ with some integers $a \geq 1$ and b.

The Theorem is proved if we prove A(S). We state another, rather different claim. For any finite set S of positive integers and an integer $n \geq 0$

B(S, n): for each $r \geq 2$ and $s \geq 1$ there is an $N(r, s)$ such that r–coloring of the integers, the set $\{1, \ldots, N\}$, $N \geq N(r, s)$ contains $n + 1$ monochromatic subsets, one for each $m = 0, \ldots, n$, of the form

$$T_m = \sum_{0 \leq j < m} a_j \cdot S + \left(\sum_{m \leq j \leq n} a_j \right) s + b,$$

with some integers $a_0 \geq 1, \ldots, a_n \geq 1$ and b.

Obviously, $N(r)$ in the statement A(S) depends on S as well, just like $N(r, s)$ in B(S, n) depends on S and n as well. The basic induction is that if we know A(S) then we want to know A($S \cup \{s\}$). The structure of the induction is as follows.
i. A(S) is trivially true for any one element set S, because any $a \cdot S + b$ is also one element, and thus monochromatic.
ii. We prove B(S, n) for every n by induction. The key step is that if we know A(S) and B(S, n) then we conclude B($S, n + 1$).
iii. B(S, n) for every n implies A($S \cup \{s\}$).

Next we work out ii. B($S, 0$) is trivial, we have to find a monochromatic $T_0 = a_0 s + b$ with some integers $a_0 \geq 1$ and b, but as T_0 is a single integer, it is always monochromatic. Although it is not necessary, let us check the claim B($S, 1$) as well. We have to find a monochromatic $T_0 = (a_0 + a_1)s + b$ and a monochromatic $T_1 = a_0 \cdot S + a_1 s + b$ simultaneously, with some integers $a_0, a_1 \geq 1$ and b. T_0 is still a single integer, and so monochromatic, we can concentrate on T_1. From A(S) we have a monochromatic $a' \cdot S + b'$, so we choose $a_0 = a'$, $a_1 = 1$ say, and $b = b' - s$. One can feel the difficulties, for $n = 2$ or more we have to point out the monochromatic property for three or more sets simultaneously, and only one of them is a singleton. This is the heart of the induction. Suppose that A(S) and B(S, n) are true and derive B($S, n + 1$). An S, s, and r–coloring of the integers are given. We refer to this coloring as the *original*. B(S, n) specifies an $N = N(r, s)$ such that

$\{1,\dots,N\}$ contains $n+1$ monochromatic subsets, one for each $m=0,\dots,n$, of the form

$$T_m=\sum_{0\le j<m} a_j\cdot S+\left(\sum_{m\le j\le n} a_j\right)s+b, \tag{2.1}$$

with some integers $a_0\ge 1,\dots,a_n\ge 1$ and b. We define an r'–coloring of the integers, namely the *new* color of u is the original color pattern of the interval $\{u+1,\dots,u+N\}$. Thus the number of new colors is $r'=r^N$, and u, v has the same new color if $u+j$, $v+j$ have the same original color for all $j=1,\dots,N$. Now $\mathrm{A}(S)$ specifies an $N'=N(r')$ such that $\{1,\dots,N'\}$ contains a (new) monochromatic $a'\cdot S+b'$ with some integers $a'\ge 1$ and b'. This means that for all $s'\in S$ the intervals $\{a's'+b'+1,\dots,a's'+b'+N\}$ have the same color pattern in the original r colors. For each $m=0,\dots,n$ that pattern has a monochromatic T_m described in (2.1), and thus it has the same color for all s', that is, we have a monochromatic $a'\cdot S+b'+T_m$ for all m. In more details, each line below is a monochromatic set, (but not all lines below have the same color).

$$\begin{gathered}
a'\cdot S+b'+(a_0+a_1+\dots+a_n)s+b,\\
a'\cdot S+b'+a_0\cdot S+(a_1+\dots+a_n)s+b,\\
a'\cdot S+b'+a_0\cdot S+a_1\cdot S+(a_2+\dots+a_n)s+b,\\
\vdots\\
a'\cdot S+b'+a_0\cdot S+a_1\cdot S+\dots+a_{n-1}\cdot S+a_ns+b.
\end{gathered}$$

Writing $a'_0=a'$, $a'_j=a_{j-1}$ for $j=1,\dots,n+1$, we got a monochromatic

$$T'_{m'}=\sum_{0\le j<m'} a'_j\cdot S+\left(\sum_{m'\le j\le n+1} a'_j\right)s+(b+b'),$$

for $m'=1,\dots,n+1$. Note that all these are subsets of $\{1,\dots,N'+N\}$. We have to add one more set for $m'=0$, which is the singleton

$$T'_0=\left(\sum_{0\le j\le n+1} a'_j\right)s+(b+b'),$$

and is trivially monochromatic, so $\mathrm{B}(S,n+1)$ follows. Note that this last subset is in $\{1,\dots,a's+b'+N\}$, and so $N(r,s)$ of $\mathrm{B}(S,n+1)$ satisfies $\le N+\max(N',a's+b')$, where $N=N(r,s)$ of $\mathrm{B}(S,n)$, $r'=r^N$, $N'=N(r')$ of $\mathrm{A}(S)$, and a',b' are the corresponding coefficients from $\mathrm{A}(S)$. This concludes ii.

Finally we work out iii. Suppose $\mathrm{B}(S,n)$ for all n. Especially for $n=r$ this says that there are $r+1$ monochromatic sets in the form of (2.1), and so there must be two of them

in the same color class, say $m = k, \ell$ with $k > \ell$. We can write the two sets as

$$T_\ell = \sum_{0 \le j < \ell} a_j \cdot S + \left(\sum_{\ell \le j < k} a_j \right) s + \left(\sum_{k \le j \le r} a_j \right) s + b,$$

$$T_k = \sum_{0 \le j < \ell} a_j \cdot S + \sum_{\ell \le j < k} a_j \cdot S + \left(\sum_{k \le j \le r} a_j \right) s + b.$$

That is $T_\ell \cup T_k$ is in one color class. Realize that this set is much bigger than what we actually need. Choosing an arbitrary $s_0 \in S$, one can easily see that

$$T_\ell \cup T_k \supset T = \left(\sum_{0 \le j < \ell} a_j \right) s_0 + \left(\sum_{\ell \le j < k} a_j \right) \cdot (S \cup \{s\}) + \left(\sum_{k \le j \le r} a_j \right) s + b$$

is monochromatic. Setting $a'' = a_\ell + \cdots + a_{k-1}$, and $b'' = (a_0 + \cdots + a_{\ell-1})s_0 + (a_k + \cdots + a_r)s + b$, we get that

$$T = a'' \cdot (S \cup \{s\}) + b''$$

is monochromatic, thus establishing $\mathrm{A}(S \cup \{s\})$. This completes the induction.

□

Theoretically it is possible to estimate $N(r)$ of $\mathrm{A}(S)$ from this proof, however it is not worth doing that, the result would be terribly weak. To get a reasonable bound for the finalized version of van der Waerden's theorem is a hard problem, and is still open.

CHAPTER 5

THE CIRCLE METHOD

In this chapter we introduce one of the most effective methods to attack additive problems. It was first developed by Hardy and Littlewood in the 1910's, and later elaborated by many mathematicians during the decades. We list some of the important milestones along this development.

- Waring problem, for each integer $k \geq 2$ there is an integer $s = s(k)$ such that every positive integer is a sum of at most s positive k–powers. (This was first proved by Hilbert, later Hardy and Littlewood developed the circle method to give a different proof.)
- Ternary Goldbach problem, every sufficiently large odd integer is a sum of three primes. (Vinogradov)
- There are infinitely many 3–term arithmetic progressions in primes. (van der Corput)
- Statistical binary Goldbach problem, almost all even integer is a sum of two primes. (Chudakov)
- $k = 3$ case of Szemerédi's theorem, a dense set of integers contains 3–term arithmetic progressions. (Roth)
- If A is dense, then $A - A$ contains infinitely many squares. (Sárközy)

We illustrate the method by proving Roth' result. Unfortunately we are forced to compute several estimates later on, so we introduce a notation. $X \ll Y$ abbreviates that there is an absolute constant $C > 0$ such that $|X| \leq CY$. It is indeed practical, for example, $2\pi X + \sqrt{X} \sin X \ll X$. Another notation we use (here and later) is an abbreviation for the exponential function.

$$e(x) =: e^{2\pi i x} = \cos 2\pi x + i \sin 2\pi x, \quad i = \sqrt{-1}.$$

This is a complex value function in a real variable, and is periodic mod 1. A little practice in the algebra of complex numbers is required, but we do not use any complex function theory. A last piece of notation is the distance of a real number from the nearest integer $\|x\| = \min\{|x - n| : n \in \mathbb{Z}\}$. We have the triangle inequality $\|x + y\| \leq \|x\| + \|y\|$, or for any positive integer $\|nx\| \leq n\|x\|$. Also, from the periodicity of the exponential function we have $e(x) = e(\pm\|x\|)$.

One can look for the background of the method in Fourier analysis, we instead reduce the explanation to the simple fact that the integral of the sin and cos functions over a full period is zero. This immediately implies for any integer n that

$$\int_0^1 e(n\alpha)\, d\alpha = \begin{cases} 1 \text{ if } n = 0, \\ 0 \text{ if } n \neq 0. \end{cases}$$

This integral is an excellent tool to test if an integer is zero or not, for example, if $a+c-2b$ is zero or not, that is if the integers a, b, c form a 3–term arithmetic progression or not.

The name of the method reflects that this integral is over the unit circle, if we consider it on the complex plane.

We need some more tools, Parseval identity and Dirichlet's approximation theorem.

Lemma 5.1: (Parseval identity) For any real or complex coefficients u_n we have

$$\int_0^1 |\sum_n u_n e(n\alpha)|^2 \, d\alpha = \sum_n |u_n|^2.$$

Proof: We prove it for a finite sum, however it is also true for infinite sums if the appropriate convergence is provided.

$$\int_0^1 |\sum_n u_n e(n\alpha)|^2 \, d\alpha = \int_0^1 \sum_n u_n e(n\alpha) \sum_m \overline{u_m} e(-m\alpha) \, d\alpha =$$
$$= \sum_n \sum_m u_n \overline{u_m} \int_0^1 e((n-m)\alpha) \, d\alpha = \sum_{n=m} u_n \overline{u_m} = \sum_n |u_n|^2.$$

□

Lemma 5.2: (Dirichlet's approximation theorem) Let $k \geq 1$ be an integer and $\alpha_1, \ldots, \alpha_k$ be real numbers. For any integer $Q \geq 1$ there is an integer $1 \leq q \leq Q^k$ such that

$$\|q\alpha_j\| \leq \frac{1}{Q}, \text{ for all } j = 1, \ldots, k.$$

Proof: The proof is the classical box principle. Split the k–dimensional unit cube into Q^k small cubes of side $1/Q$. These are the boxes. For $u = 0, \ldots, Q^k$ consider the vectors $(u\alpha_1, \ldots, u\alpha_k)$ mod 1, that is reduce all coordinates to the interval $[0, 1)$ by subtracting the largest not bigger integer from it. Each vector is in the unit cube and there are $Q^k + 1$ of them, at least one box should contain two vectors, they belong to $u > v$ say. $q = u - v$ will do it as there are integers u_j and v_j such that for all $j = 1, \ldots, k$ we have $|(u\alpha_j - u_j) - (v\alpha_j - v_j)| \leq 1/Q$.

□

In Chaper 7. we see a quite different proof of this lemma.

Theorem 5.3: (Roth, 1953) Let $\delta > 0$ be given. There is an $N(\delta)$ such that any set of integers $A \subset \{1, \ldots, N\}$, $|A| \geq \delta N$, $N \geq N(\delta)$ contains non trivial 3–term arithmetic progressions.

Proof: We can suppose that $|A| = \delta N$. Let us define

$$\hat{A}(\alpha) = \sum_{a \in A} e(a\alpha).$$

We start with the integral

$$S =: \int_0^1 \Big(\hat{A}(\alpha)\Big)^2 \hat{A}(-2\alpha)\, d\alpha = \int_0^1 \sum_{a_1\in A} e(a_1\alpha) \sum_{a_3\in A} e(a_3\alpha) \sum_{a_2\in A} e(-2a_2\alpha)\, d\alpha =$$

$$= \sum_{a_1\in A}\sum_{a_2\in A}\sum_{a_3\in A} \int_0^1 e((a_1+a_3-2a_2)\alpha)\, d\alpha = \sum_{a_1+a_3=2a_2,\, a_j\in A} 1.$$

S counts the 3–term arithmetic progressions in A, including the δN trivial progressions $a_1 = a_2 = a_3$. We have to compute this integral.

Note that proving, for example, the ternary Goldbach conjecture needs the study of the very similar integral with the function $\hat{P}(\alpha) = \sum_{p\le N,\, p \text{ prime}} e(p\alpha)$. There are just enough analytic information about $\hat{P}(\alpha)$ to compute an asymptotic estimate to the integral, for example, we have an asymptotic expansion of $\hat{P}(a/q)$ for small q, and an upper bound for all α not close to a rational number with small denominator. This makes the method so delicate, the integral is split into *major arcs*, where $\hat{P}(\alpha)$ is big, around rational numbers with small denominator, and into *minor arcs*, where it is small.

The present case is somewhat different, we know very little about the function $\hat{A}(\alpha)$. Nevertheless we follow the above pattern and we trace the arcs, where our function is big. It is little easier to do that in discrete points, so we choose a (large) integer X, and define cells

$$\mathcal{C}_j = \left\{ \frac{j}{X} - \frac{1}{2X} \le \alpha < \frac{j}{X} + \frac{1}{2X} \right\},\ j = 1, \dots, X.$$

They cover the interval $[1/2X, 1 + 1/2X)$ and the integral over this interval is equivalent to the original integral as $e(x)$ is periodic mod 1. The derivative of $\hat{A}(\alpha)$ can be estimated trivially by

$$|\hat{A}'(\alpha)| = |\sum_{a\in A} e(a\alpha) 2\pi i a| \le 2\pi \sum_{a\in A} a \le 2\pi\delta N^2,$$

and the mean–value theorem ensures

$$|\hat{A}(\alpha) - \hat{A}(\frac{j}{X})| \le |\alpha - \frac{j}{X}| 2\pi\delta N^2 \le \frac{\pi\delta N^2}{X}, \text{ if } \alpha \in \mathcal{C}_j. \tag{5.1}$$

We define the major arcs $\mathcal{M}$ and minor arcs m as follows.

$$\mathcal{M} = \cup \mathcal{C}_j \quad \text{for} \quad |\hat{A}(\frac{j}{X})| > \frac{2\pi\delta N^2}{X},$$

$$m = \cup \mathcal{C}_j \quad \text{for} \quad |\hat{A}(\frac{j}{X})| \le \frac{2\pi\delta N^2}{X}.$$

Coupling the definition with (5.1) we have the estimates

$$\text{If} \quad \alpha \in \mathcal{M} \quad \text{then} \quad |\hat{A}(\alpha)| > \frac{\pi\delta N^2}{X},$$

$$\text{If} \quad \alpha \in m \quad \text{then} \quad |\hat{A}(\alpha)| \le \frac{3\pi\delta N^2}{X}. \tag{5.2}$$

Parseval identity, Lemma 5.1, can help to estimate the size of the major arcs. Suppose $\mathcal{M}$ is the union of K cells. We have

$$K\frac{1}{X}\frac{\pi^2\delta^2N^4}{X^2} < \int_{\mathcal{M}} |\hat{A}(\alpha)|^2\,d\alpha \le \int_0^1 |\hat{A}(\alpha)|^2\,d\alpha = \delta N,$$

that is

$$K \le \frac{X^3}{\pi^2\delta N^3}. \tag{5.3}$$

(5.3) says that the major arcs are not really major, they occupy only a small fraction of the unit circle. However, they still provide the majority of the integral. It maybe surprising that to show this we actually need the small size of the major arcs. Note that in his original proof, Roth used a single arc around zero as major arc.

Next we construct a weight function, which is close to 1 on the major arcs, so writing it into the integral is harmless. Our basic help here is Dirichlet's approximation theorem, Lemma 5.2. Let $\{\alpha_1,\dots,\alpha_K\}$ be the centers of the major arcs, so they are the fractions j/X with $\mathcal{C}_j \subset \mathcal{M}$. For any integer Q we find another integer $1 \le q \le Q^K$ such that

$$\|\frac{qj}{X}\| \le \frac{1}{Q}, \text{ for all } j, \text{ with } \mathcal{C}_j \subset \mathcal{M}.$$

As any $\alpha \in \mathcal{M}$ is at most $1/2X$ apart from one of those fractions, for all $\alpha \in \mathcal{M}$ we have

$$\|q\alpha\| = \|\frac{qj}{X} + q(\alpha - \frac{j}{X})\| \le \|\frac{qj}{X}\| + q|\alpha - \frac{j}{X}| \le \frac{1}{Q} + \frac{Q^K}{2X}. \tag{5.4}$$

Observe that $e(\beta) - 1 \ll \|\beta\|$ for all β and so

$$W(\alpha) =: 1 - \frac{1}{4L+1}\sum_{\ell=-2L}^{2L} e(q\ell\alpha) \ll \frac{1}{L}\sum_{\ell=1}^{2L}\|q\ell\alpha\| \ll \frac{\|q\alpha\|}{L}\sum_{\ell=1}^{2L}\ell \ll L\|q\alpha\|. \tag{5.5}$$

We have an upper bound for $W(\alpha)\hat{A}(\alpha)$ for all α. On the major arcs we use (5.5) with (5.4) for $W(\alpha)$ coupled with the trivial bound $|\hat{A}(\alpha)| \le \delta N$, on the minor arcs we use (5.2) for $\hat{A}(\alpha)$ coupled with the trivial bound $|W(\alpha)| \le 2$. We arrive at

$$W(\alpha)\hat{A}(\alpha) \ll \delta N\left(\frac{N}{X} + \frac{L}{Q} + \frac{LQ^K}{X}\right).$$

Recall the definition of S from the beginning of the proof. Much the same way we can express the number of solutions of the equation $a_1 + a_3 = 2a_2 + 2q\ell$ where $a_j \in A$ and $|\ell| \le 2L$ by an integral using $W(\alpha)$ as well. Using Lemma 5.1 with the above bound we arrive at

$$\left|\sum_{\substack{a_1,a_2,a_3\in A\\ a_1+a_3=2a_2}} 1 - \frac{1}{4L+1}\sum_{\substack{a_1,a_2,a_3\in A,\ |\ell|\le 2L\\ a_1+a_3=2a_2+2q\ell}} 1\right| = \left|\int_0^1 \hat{A}^2(\alpha)W(-2\alpha)\hat{A}(-2\alpha)\,d\alpha\right| \le \tag{5.6}$$

$$\le \max_\beta |W(\beta)\hat{A}(\beta)| \int_0^1 |\hat{A}(\alpha)|^2\,d\alpha \ll \delta^2N^2\left(\frac{N}{X} + \frac{L}{Q} + \frac{LQ^K}{X}\right).$$

Let us analyze where we are. We can choose the parameters L, Q and X such that the right hand side is small. So we get a lower bound for S, our main object, if we get a lower bound for the second sum above, say S'.

$$S' =: \frac{1}{4L+1} \sum_{\substack{a_1,a_2,a_3\in A,\, |\ell|\le 2L \\ a_1+a_3=2a_2+2q\ell}} 1.$$

Writing it in a slightly different form reveals the subtle, but important difference to S. $a_1+a_3 = 2a_2+2q\ell$ means that $\frac{a_1+a_3}{2} \equiv a_2 \pmod q$ and $\frac{a_1+a_3}{2} - 2qL \le a_2 \le \frac{a_1+a_3}{2} + 2qL$, that is for many a_1 and a_3 we want to find an a_2, which is not necessary equal to, (as it should be in S), but it is in the same residue class, and not too far from $\frac{a_1+a_3}{2}$. This will not be too difficult if A is well distributed in segments of residue classes mod q of length $2qL$. In this point the obvious analytic flavor of the proof turns to a rather combinatorial flavor.

We define *blocks* of integers in the following way. For any integer u

$$\mathcal{B}_u = \{u + 2q\ell :\, \ell = 1, \dots, L\}.$$

The 2 in the definition is for our comfort only, this provides that all element in a single block has the same parity. As a block has L elements, we expect that it contains about δL integers from A. We call a block *good* or *bad* respectively if

$$|\mathcal{B}_u \cap A| > \frac{1}{9}\delta L, \text{ resp. } |\mathcal{B}_u \cap A| \le \frac{1}{9}\delta L.$$

First suppose that no blocks contain too many elements of A, this is the well distribution case. More precisely we suppose that

$$|\mathcal{B}_u \cap A| \le \frac{10}{9}\delta L, \text{ for all } u. \tag{5.7}$$

We can estimate the number of good and (non empty) bad blocks. Since every a is in exactly L blocks, on one hand we have

$$\sum_u |B_u \cap A| = \sum_{a\in A} \sum_{\substack{u \\ a\in B_u}} 1 = \sum_{a\in A} L = \delta N L.$$

Note that a block is obviously empty if $u \le -2qL$ or $u \ge N$, but the above argument is only complete if we consider blocks with negative subscript as well. Anyway, the number of non empty blocks with negative subscript is less than $2qL$. Let G and B be the number of good and bad blocks $\mathcal{B}_u$ with $0 \le u < N$. Note that $G + B = N$. We have

$$\sum_u |B_u \cap A| \le \frac{10}{9}\delta L(G + 2qL) + \frac{1}{9}\delta L(N - G) \le \delta LG + \frac{1}{8}\delta LN,$$

if we suppose that

$$160qL \le N. \tag{5.8}$$

Together with the lower bound we get

$$G \ge \frac{7}{8}N, \quad B \le \frac{1}{8}N. \tag{5.9}$$

We are going to show that for any $a_1 \in A$ there are many $a_3 \in A$ such that $\frac{a_1+a_3}{2}$ is in a good block, and so any element of A in that block gives a contribution to S'. We can fix a_1 and it is practical to look for a_3 in good blocks. In other words we start with the equation

$$a_1 + u = 2v, \quad u, v \ge 0, \quad \mathcal{B}_u, \mathcal{B}_v \text{ are good.} \tag{5.10}$$

There are at most $\frac{1}{2}N$ good blocks with opposite parity to a_1, so by (5.9) we have at least $\frac{3}{8}N$ good blocks $\mathcal{B}_u$ with $0 < \frac{a_1+u}{2} = v < N$ is an integer, and for at least $\frac{1}{4}N$ of them $\mathcal{B}_v$ must be a good block. As $\mathcal{B}_u$ and $\mathcal{B}_v$ are good blocks, there are at least $\frac{1}{9}\delta L$ choices of $a_3 = u + 2q\ell_3 \in A$ and at least $\frac{1}{9}\delta L$ choices of $a_2 = v + 2q\ell_2 \in A$ such that

$$\frac{a_1 + a_3}{2} = \frac{a_1 + u}{2} + q\ell_3 = v + q\ell_3 = a_2 + q(\ell_3 - 2\ell_2) = a_2 + q\ell,$$

where $|\ell| \le 2L$. All these triples counts for S', their number is at least $\ge \delta N \cdot \frac{1}{4}N \cdot \frac{1}{9}\delta L \cdot \frac{1}{9}\delta L$. The multiplicity of a triple (a_1, a_2, a_3) is at most L as a_1 is determined, a_3 determines at most L blocks $\mathcal{B}_u$, and v is determined by $a_1 + u = 2v$. We arrive at

$$S' \ge \frac{1}{4L+1} \cdot \frac{1}{L} \cdot \delta N \cdot \frac{1}{4}N \cdot \frac{1}{9}\delta L \cdot \frac{1}{9}\delta L \gg \delta^3 N^2.$$

We have to provide that the right hand side of (5.6) is at most half of this bound, therefore

$$S \gg \delta^3 N^2,$$

whenever

$$\frac{N}{X} + \frac{L}{Q} + \frac{LQ^K}{X} \le c\delta,$$

with a sufficiently small $c > 0$. The first term suggests the choice for X, the bound for K with a sufficiently large constant $C > 1$ comes from (5.3), the rest are comfortable choices.

$$X = \left\lceil \frac{4N}{c\delta} \right\rceil, \quad K \le \frac{C}{\delta^4}, \quad \frac{4}{c\delta} \le L = \left\lfloor N^{\frac{\delta^4}{3C}} \right\rfloor \le N^{\frac{1}{2K+1}}, \quad Q = L^2.$$

(5.8) is satisfied. Note that the choice of X, Q and L depend on N and δ alone, but not on K or A themselves. Since $\delta^3 N^2 > \delta N$ for sufficiently large N, A contains non trivial arithmetic progressions as well.

For the case of irregular distribution of A, we have found a block $|\mathcal{B}_u \cap A| > \frac{10}{9}\delta L$. Let $N_1 = L$, $\delta_1 = \frac{10}{9}\delta$ and $A_1 = \{1 \le \ell \le N_1 : u + 2q\ell \in A\}$. Observe that a non trivial 3–term arithmetic progression in A_1 naturally leads to a non trivial 3–term arithmetic progression in A. We have better chance when we repeat the whole argument with A_1, as δ_1 is substantially bigger than δ. If it is still not conclusive then we find an

$$A_2, \quad \delta_2 = \frac{10}{9}\delta_1 = (\frac{10}{9})^2\delta, \quad N_2 = \left[N_1^{\frac{\delta^4}{3C}}\right] \ge N_1^{\frac{\delta^4}{4C}} \ge N^{(\frac{\delta^4}{4C})^2},$$

and so on. After j steps we find an

$$A_j, \quad \delta_j = (\frac{10}{9})^j\delta, \quad N_j \ge N^{(\frac{\delta^4}{4C})^j}.$$

As the worst case, we continue repeating the argument until $j \ge \log \frac{7}{8\delta} / \log \frac{10}{9}$, that is we find an

$$A_j \subset \{1, \dots, N_j\}, \quad |A_j| \ge \delta_j N_j \ge \frac{7}{8} N_j.$$

However, exactly as we estimated the number of solutions to (5.10), we get that there are many 3–term arithmetic progressions in A_j, and so in A as well.

This concludes the long proof.

□

Note, that N_j is still a (small) power of N if δ is fixed, however tracing the speed of diminishing of N_j one can even let δ be a function of N. It is highly recommended to double check the last steps of the calculation to prove that *any set of integers $A \subset \{1, \dots, N\}$ contains non trivial 3–term arithmetic progressions if*

$$|A| \ge \frac{N}{e^{c\sqrt{\log \log N}}},$$

with a sufficiently small constant $c > 0$.

The best result in this direction is due to Sanders and proves the same result with

$$|A| \gg \frac{N(\log \log N)^C}{\log N},$$

with some explicit large constant C. On the other hand examples are known for a set $A \subset \{1, \dots, N\}$ with no 3–term arithmetic progression and size

$$|A| \ge \frac{N}{e^{c\sqrt{\log N}}}.$$

The first such example is due to Behrend and worked out in the homework problems. Recent constructions are improvements only on the numerical value of c.

CHAPTER 6

SUM–PRODUCT ESTIMATES

In Chapters 1 and 3 we studied questions about the additive structure of integers, real numbers, or residue classes, actually we only used the additive groups of these rings or fields. Most of the results can be translated to the multiplicative structures as well, sometimes it is simply the application of the very same statement to the set $B = \{\log a : a \in A\}$. The reader is invited to check when this argument works.

In this chapter, however, we want to use both operations, and study the connection between them. If A is a finite set of positive real numbers, then $A + A$ is minimal for arithmetic progressions, Corollary 1.3. This translates as $A \cdot A$ is minimal for geometric progressions. Arithmetic progressions or geometric progressions are quite different from each other. This observation led Erdős and Szemerédi to conjecture that

Conjecture 6.1: (Erdős and Szemerédi) If A is a finite set of positive real numbers, $\epsilon > 0$ is a fixed real number, and $|A| = n > n_0(\epsilon)$, then

$$\max\{|A + A|, |A \cdot A|\} \geq n^{2-\epsilon}.$$

They themselves proved that a weaker lower bound is indeed true, namely there is a small real number $c > 0$ such that

$$\max\{|A + A|, |A \cdot A|\} \gg n^{1+c}.$$

Explicit values of c were computed by several mathematicians. Elekes made the first breakthrough when using incidence geometry he proved $c = 1/4$. Even simpler approach was discovered by Solymosi proving $c = 1/3$, the best result so far.

Theorem 6.2: (Solymosi) If A is a finite set of positive real numbers, and $|A| = n$, then

$$|A + A|^2 |A \cdot A| \geq \frac{4n^4}{5\lceil \log_2 n \rceil}.$$

Notice that in the case when $A + A$ is close to its minimum, say $|A + A| \leq n^{1+\epsilon}$ we get the conjectural lower bound for $|A \cdot A|$. In general, bounding the left hand side with $\max\{|A + A|, |A \cdot A|\}^3$ we immediately get (after numerical simplification)

Corollary 6.3: If A is a finite set of positive real numbers, and $|A| = n$, then

$$\max\{|A + A|, |A \cdot A|\} \gg \frac{n^{4/3}}{\log^{1/3} n}.$$

Proof of Theorem 6.2: Recall the notion and elementary properties of representation functions $r_{A \cdot A}$ and $r_{A/A}$ from the end of Chapter 1. We have

$$\sum_y r_{A \cdot A}(y) = n^2,$$

and the Cauchy–Schwarz inequality (Lemma 1.14) implies that

$$n^4 = \left(\sum_y r_{A\cdot A}(y)\right)^2 \le \left(\sum_y 1\right)\left(\sum_y r_{A\cdot A}(y)^2\right) = |A\cdot A|\sum_y r_{A\cdot A}(y)^2,$$

as only $y \in A\cdot A$ gives a non zero contribution to the sums. Remember also that

$$\sum_y r_{A\cdot A}(y)^2 = \sum_x r_{A/A}(x)^2,$$

because they compute the number of solutions of $ab = a'b'$ or $a/b' = a'/b$ in $a, b, a', b' \in A$, and the two equations are the same. We arrive at

$$n^4 \le |A\cdot A|\sum_x r_{A/A}(x)^2. \tag{6.1}$$

The second half of the proof is more tricky, and has geometric motivation. We will estimate the sum on the right hand side of (6.1) by means of $A + A$. The argument works only for approximately equal $r_{A/A}(x)$, therefore we start with splitting the sum accordingly. Let K be defined by

$$2^{K-1} \le n < 2^K. \tag{6.2}$$

Since $r_{A/A}(x) \le n$, we have

$$\sum_x r_{A/A}(x)^2 = \sum_{k=1}^{K} \sum_{2^{k-1}\le r_{A/A}(x)<2^k} r_{A/A}(x)^2. \tag{6.3}$$

Let us fix a $1 \le k \le K$ for which the inner sum is non empty. We arrange the quotients x counted in the k-th sum by increasing order, that is $\{x \in A/A : 2^{k-1} \le r_{A/A}(x) < 2^k\} = \{x_1 < x_2 < \cdots < x_m\}$. Consider $A \times A$ on the plane, the integer vectors (a, b) with coordinates from A. They are inside the first quadrant. The line $b = x_j a$ covers exactly $r_{A/A}(x_j)$ of them. With a little abuse of notation we refer to this line, more precisely to the half line in the first quadrant, simply as the *ray* x_j. Pick a vector (a, b) on the ray x_j and another vector (a', b') on the next ray x_{j+1}. Realize that their sum $(a + a', b + b')$ is in the sector defined by the two rays, indeed one can quickly check that

$$x_j = \frac{b}{a} < \frac{b+b'}{a+a'} < \frac{b'}{a'} = x_{j+1}.$$

Not only the vector $(u, v) = (a + a', b + b')$ is an element of $(A + A) \times (A + A)$, but (u, v) determines the two initial vectors, as is clear from the parallelogram rule of adding vectors. This provides the next inequality.

$$r_{A/A}(x_j)\, r_{A/A}(x_{j+1}) \le \#\{(u, v) \in (A + A) \times (A + A) : x_j < \frac{v}{u} < x_{j+1}\}.$$

As for different j–s these sectors are disjoint, we also have

$$\sum_{j=1}^{m-1} r_{A/A}(x_j)\, r_{A/A}(x_{j+1}) \le \#\{(u,v) \in (A+A)\times(A+A) :\ x_1 < \frac{v}{u} < x_m\}. \tag{6.4}$$

(6.4) and (6.3) are enough to prove the Theorem with a weaker constant, so the rest of the proof is only cosmetics. It is recommended to make the shortcut from here.

We define two more half lines x_0 and x_∞ (roughly corresponds to 0 and ∞) to include more vectors, and to handle the $m = 1$ case. Let a_0 be the least element of A. The ray x_0 covers the vectors (a, a_0), where a runs over all the first coordinates of the vectors on the ray x_1. In other words, it is the vertical projection of the ray x_1 to the horizontal line $b = a_0$. Similarly the ray x_∞ is the horizontal projection of the ray x_m to the vertical line $a = a_0$. Obviously these rays contain $r_{A/A}(x_1)$ or $r_{A/A}(x_m)$ vectors from $A \times A$. Adding the vectors of the ray x_0 to the vectors of the ray x_1 will result vectors (u, v) of $(A+A)\times(A+A)$ with $v/u < x_1$. Similarly adding the vectors of the ray x_m to the vectors of the ray x_∞ will result vectors (u, v) of $(A + A) \times (A + A)$ with $v/u > x_m$. Just like beforehand the vector (u, v) determines its origins. Thus we can complete (6.4) to

$$r_{A/A}(x_1)^2 + \sum_{j=1}^{m-1} r_{A/A}(x_j)\, r_{A/A}(x_{j+1}) + r_{A/A}(x_m)^2 \le |A + A|^2. \tag{6.5}$$

We finish with a little elementary calculus. One can check that for $1/2 < z < 2$ we have $z + 1/z < 5/2$, that is with $z = r/s$ we have $r^2 + s^2 < 5/2 \cdot rs$. Therefore by (6.5)

$$\sum_{j=1}^{m} r_{A/A}(x_j)^2 = \frac{1}{2}\, r_{A/A}(x_1)^2 + \sum_{j=1}^{m-1} \frac{r_{A/A}(x_j)^2 + r_{A/A}(x_{j+1})^2}{2} + \frac{1}{2}\, r_{A/A}(x_m)^2 \le$$

$$\le \frac{1}{2}\, r_{A/A}(x_1)^2 + \frac{5}{4} \sum_{j=1}^{m-1} r_{A/A}(x_j)\, r_{A/A}(x_{j+1}) + \frac{1}{2}\, r_{A/A}(x_m)^2 \le \frac{5}{4}\, |A + A|^2.$$

Writing to (6.3) and (6.1) we arrive at

$$n^4 \le \frac{5}{4}\, K |A \cdot A||A + A|^2.$$

Finally the theorem follows from $K < \lceil \log_2 n \rceil$.

□

Next we turn to the similar problem in the finite field $\mathbb{Z}_p$, that is in the field of residue classes mod p, a prime number. As in Chapter 3, we sometimes write $=$ rather than $\equiv$ mod p. We start with a couple of useful lemmas. We basically develop the discrete analogue of the circle method. (Recall the exponential notation from Chapter 5.) Again, discrete Fourier analysis is simplified to the relation

$$\frac{1}{p} \sum_{x \in \mathbb{Z}_p} e(\frac{nx}{p}) = \begin{cases} 1, \text{ if } n = 0, \\ 0, \text{ if } 0 \ne n \in \mathbb{Z}_p. \end{cases} \tag{6.6}$$

The first statement is trivial and for $0 \neq n \in \mathbb{Z}_p$ xn runs over $\mathbb{Z}_p$ when x does the same. In this case the sum is the sum of p-th roots of unity, and it is zero. (Or this is a geometric sequence, and we know the sum formula.) We again have Parseval identity (compare to Lemma 5.1, and repeat the identical proof).

Lemma 6.4: (Parseval identity) For any (complex) coefficients u_n we have

$$\frac{1}{p}\sum_{x\in\mathbb{Z}_p}\left|\sum_{n\in\mathbb{Z}_p}u_n e(\frac{nx}{p})\right|^2 = \sum_{n\in\mathbb{Z}_p}|u_n|^2.$$

□

Lemma 6.5: For any(complex) coefficients u_m and v_n, and for any $0 \neq x \in \mathbb{Z}_p$ we have the bound

$$\left|\sum_{m\in\mathbb{Z}_p}\sum_{n\in\mathbb{Z}_p}u_m v_n e(\frac{mnx}{p})\right| \leq \left(p\sum_{m\in\mathbb{Z}_p}|u_m|^2\sum_{n\in\mathbb{Z}_p}|v_n|^2\right)^{1/2}.$$

Especially for any subsets A and B of $\mathbb{Z}_p$ we have

$$\left|\sum_{a\in A}\sum_{b\in B}e(\frac{abx}{p})\right| \leq \sqrt{p\,|A||B|}.$$

Proof: This follows from the Cauchy–Schwarz inequality (Lemma 1.14), and from Parseval identity (Lemma 6.4). Indeed

$$\left|\sum_{m\in\mathbb{Z}_p}\sum_{n\in\mathbb{Z}_p}u_m v_n e(\frac{mnx}{p})\right|^2 \leq \left(\sum_{m\in\mathbb{Z}_p}|u_m|\left|\sum_{n\in\mathbb{Z}_p}v_n e(\frac{mnx}{p})\right|\right)^2 \leq$$

$$\leq \sum_{m\in\mathbb{Z}_p}|u_m|^2\sum_{m\in\mathbb{Z}_p}\left|\sum_{n\in\mathbb{Z}_p}v_n e(\frac{mnx}{p})\right|^2 = p\sum_{m\in\mathbb{Z}_p}|u_m|^2\sum_{n\in\mathbb{Z}_p}|v_n|^2.$$

□

Theorem 6.6: (Garaev) For any non empty subset $0 \notin A$ of $\mathbb{Z}_p$ we have

$$|A+A||A\cdot A| \geq \min\{\frac{p}{2}|A|,\ \frac{1}{4p}|A|^4\}.$$

Corollary 6.7: For any non empty subset $0 \notin A$ of $\mathbb{Z}_p$ we have

$$\max\{|A+A|, |A\cdot A|\} \gg \min\{\sqrt{p|A|},\ \frac{|A|^2}{\sqrt{p}}\}.$$

Note that these results are non trivial only when $|A|$ is substantially larger than $\sqrt{p}$. To derive a result in the form $\max\{|A+A|, |A\cdot A|\} \gg |A|^{1+c}$ for smaller A is very deep already for a small $c>0$, beyond the level of this exposition.

Proof of Theorem 6.6: Let B be the set of the multiplicative inverses of A, that is $B=\{b=1/a:\ a\in A\}$, or in another language, $B=\{b\in\mathbb{Z}_p:\ ba\equiv 1 \pmod p \text{ for some } a\in A\}$. Obviously $|B|=|A|$. (The condition $0\notin A$ is required here.) We consider the number of solutions of

$$qb+a=s, \quad q\in A\cdot A,\ b\in B,\ a\in A,\ s\in A+A. \tag{6.7}$$

For any triple (a_1,a_2,a_3), $a_j\in A$ there is a solution to (6.7), namely $q=a_1a_2$, $b=1/a_1$, $a=a_3$, and $s=a_2+a_3$. These are all different solutions to (6.7) as a given solution (q,b,a,s) determines (a_1,a_2,a_3) by $a_1=1/b$, $a_2=s-a$, and $a_3=a$. This means that

$$|A|^3 \le \#\text{ solutions to (6.7)}. \tag{6.8}$$

Note that we do not have equation here, because there are solutions of different type as well. The theorem will follow from estimating the number of solutions to (6.7) by the circle method. By (6.6) we have

$$\frac{1}{p}\sum_{x\in\mathbb{Z}_p}\left(\sum_{q\in A\cdot A}\sum_{b\in B}e(\frac{qbx}{p})\right)\left(\sum_{a\in A}e(\frac{ax}{p})\right)\left(\sum_{s\in A+A}e(-\frac{sx}{p})\right)=$$

$$=\sum_{q\in A\cdot A}\sum_{b\in B}\sum_{a\in A}\sum_{s\in A+A}\frac{1}{p}\sum_{x\in\mathbb{Z}_p}e(\frac{(qb+a-s)x}{p})=\#\text{ solutions to (6.7)}.$$

This time we have a single "major arc", the point $x=0$, because for all other x the first function is small by Lemma 6.5. To estimate the rest of the sum over x, the "minor arc" contribution, we use the Cauchy–Schwarz inequality and the Parseval identity.

$$\#\text{ of solutions to (6.7)} = \tag{6.9}$$

$$=\frac{|A||B||A+A||A\cdot A|}{p}+\frac{1}{p}\sum_{\substack{x\in\mathbb{Z}_p\\ x\neq 0}}\left(\sum_{q\in A\cdot A}\sum_{b\in B}e(\frac{qbx}{p})\right)\left(\sum_{a\in A}e(\frac{ax}{p})\right)\left(\sum_{s\in A+A}e(-\frac{sx}{p})\right)\le$$

$$\le\frac{|A||B||A+A||A\cdot A|}{p}+\frac{1}{p}\sum_{\substack{x\in\mathbb{Z}_p\\ x\neq 0}}\left|\sum_{q\in A\cdot A}\sum_{b\in B}e(\frac{qbx}{p})\right|\left|\sum_{a\in A}e(\frac{ax}{p})\right|\left|\sum_{s\in A+A}e(\frac{sx}{p})\right|\le$$

$$\le\frac{|A||B||A+A||A\cdot A|}{p}+\frac{1}{p}\sqrt{p|A\cdot A||B|}\,\sqrt{p|A|}\,\sqrt{p|A+A|},$$

because in the last step after using Lemma 6.5 for the first factor we can extend the sum over x to the whole $\mathbb{Z}_p$ and use the Cauchy–Schwarz inequality as follows.

$$\left(\sum_{x\in\mathbb{Z}_p}\left|\sum_{a\in A}e(\frac{ax}{p})\right|\left|\sum_{s\in A+A}e(\frac{sx}{p})\right|\right)^2\le\sum_{x\in\mathbb{Z}_p}\left|\sum_{a\in A}e(\frac{ax}{p})\right|^2\sum_{x\in\mathbb{Z}_p}\left|\sum_{s\in A+A}e(\frac{sx}{p})\right|^2.$$

Recall that $|B| = |A|$, combining (6.8) and (6.9) we arrive at

$$|A|^2 \le \frac{1}{p}\,|A||A+A||A\cdot A| + \sqrt{p|A+A||A\cdot A|}.$$

Since at least one of the two terms on the right hand side must exceed half of the left hand side this concludes the proof of our theorem.

□

In the last part of this chapter we study the problem of representing $\mathbb{Z}_p$ as a sum like $B+\dots+B$ with a small number of terms. Without any special structure of the subset B one can not say too much. For example, if $B = \{1, 2, \dots, n\}$ then we need more than p/n terms. However, if B has some multiplicative structure, then the situation improves a lot. We consider the case of $B = A\cdot A$. We will use Theorem 6.6 to show the following

Theorem 6.8: Let $0 \notin A$ be a subset of $\mathbb{Z}_p$ with $|A| > \sqrt{2}\,p^{5/8}$. We have

$$(A+A)\cdot(A+A) + A\cdot A + A\cdot A = \mathbb{Z}_p.$$

Especially we have

$$A\cdot A + A\cdot A + A\cdot A + A\cdot A + A\cdot A + A\cdot A = 6(A\cdot A) = \mathbb{Z}_p.$$

Proof: We want to show, that for all $r \in \mathbb{Z}_p$ the equation

$$r = s_1 s_2 + q_1 + q_2, \quad s_i \in A+A,\ q_j \in A\cdot A, \tag{6.10}$$

can be solved. We use the circle method to express the number of solutions to (6.10). Indeed

$$\#\text{ of solutions to (6.10)} \;= \frac{1}{p}\sum_{x\in\mathbb{Z}_p}\left(\sum_{s_1\in A+A}\sum_{s_2\in A+A} e(\frac{s_1 s_2 x}{p})\right)\left(\sum_{q\in A\cdot A} e(\frac{qx}{p})\right)^2 e(-\frac{rx}{p}).$$

$x = 0$ is the major arc, and contributes

$$\frac{|A+A|^2|A\cdot A|^2}{p}.$$

The rest of x is the minor arc. We estimate the first factor, the sum over s_1 and s_2, by Lemma 6.5, after that we extend the sum over x to the whole $\mathbb{Z}_p$ and apply the Parseval identity, Lemma 6.4. The contribution of the minor arcs is than bounded by

$$\le \max_{y\neq 0}\left|\sum_{s_1\in A+A}\sum_{s_2\in A+A} e(\frac{s_1 s_2 y}{p})\right| \frac{1}{p}\sum_{x\in\mathbb{Z}_p}\left|\sum_{q\in A\cdot A} e(\frac{qx}{p})\right|^2 \le \sqrt{p}\,|A+A|\,|A\cdot A|.$$

Therefore we have

$$\#\text{ of solutions to (6.10)} \ \geq \frac{|A+A|^2|A\cdot A|^2}{p} - \sqrt{p}\,|A+A|\,|A\cdot A| > 0,$$

whenever

$$|A+A|\,|A\cdot A| > p^{3/2}.$$

By Theorem 6.6 it is enough to provide that

$$\min\{\frac{p}{2}\,|A|,\ \frac{1}{4p}|A|^4\} > p^{3/2},$$

what follows from $|A| > \sqrt{2}\,p^{5/8}$.

□

There are several similar results, they vary according to the number of terms, the number of factors in one term, and the condition on $|A|$. The reader is invited to mimic the above proof to show that $4\,(A\cdot A) = \mathbb{Z}_p$ whenever $|A| > (2p)^{2/3}$. The starting equation to represent any $r \in \mathbb{Z}_p$ in the required form should be $r = sa + q_1 + q_2$, where $s \in A+A$, $a \in A$, and $q_j \in A\cdot A$. A deep result of Bourgain implies that for all $\delta > 0$ there are (large) integers n and N such that $NA^n = \mathbb{Z}_p$, whenever $|A| > p^\delta$. The dependence of n and N on δ is in the center of investigation in additive combinatorics.

CHAPTER 7

TECHNICAL PREPARATION: LATTICES

Minkowski's lattice geometry is a wonderful tool in different problems in number theory, including additive representations, Diophantine approximation, theory of quadratic forms, and many more. However, we restrict ourselves to develop just enough for our application to the proof of Freiman's theorem in Chapter 10. The interested reader is encouraged to study this beautiful subject from other sources, for example, J.W.S.Cassels, An Introduction to the Geometry of Numbers, Springer–Verlag, Berlin, 1959.

Our presentation depends on a lot of linear algebra, therefore a certain level of knowledge is expected. We review many statements without proofs, in the same time we detail the less well known theorems, like Mikowski's convex body theorems.

We work in the d–dimensional real Euclidean space $\mathbb{R}^d$. Let $\mathbf{a}_1, \dots, \mathbf{a}_d$ be a basis, that is, be linearly independent. The abelian group Λ generated by these vectors is the set of vectors of the form

$$u_1\mathbf{a}_1 + \cdots + u_d\mathbf{a}_d, \quad u_1, \dots, u_d \in \mathbb{Z}.$$

This is a *lattice*, and the set of linearly independent vectors $\mathbf{a}_1, \dots, \mathbf{a}_d$ is a basis of the lattice. A lattice is a *discrete* subgroup of $\mathbb{R}^d$, and any discrete subgroup of $\mathbb{R}^d$, if it contains d linearly independent vectors, is a lattice.

For example, the standard basis of $\mathbb{R}^d$ is the set of vectors $\mathbf{e}_1, \dots, \mathbf{e}_d$, where

$$\begin{aligned} \mathbf{e}_1 &= (1, 0, 0, \dots, 0) \\ \mathbf{e}_2 &= (0, 1, 0, \dots, 0) \\ &\vdots \\ \mathbf{e}_d &= (0, 0, 0, \dots, 1) \end{aligned}$$

The integer lattice $\mathbb{Z}^d$ is the lattice with basis $\mathbf{e}_1, \dots, \mathbf{e}_d$. If $\mathbf{b}_1, \dots, \mathbf{b}_d$ are d linearly independent vectors in a lattice Λ, then there exists a basis $\mathbf{a}_1, \dots, \mathbf{a}_d$ for Λ such that each of the vectors $\mathbf{b}_j$ is of the form

$$\mathbf{b}_j = \sum_{i=1}^{j} v_{ij}\mathbf{a}_i, \quad v_{ij} \in \mathbb{Z}. \tag{7.1}$$

Notice that in this representation $\mathbf{b}_j$ is expressed from $\mathbf{a}_1, \dots, \mathbf{a}_j$ only.

The basis of a lattice is not uniquely determined by the lattice. For example, not only $(1, 0)$ and $(0, 1)$ generates $\mathbb{Z}^2$ but also $(7, 5)$ and $(4, 3)$, as one can easily check. The vectors $\mathbf{a}_1, \dots, \mathbf{a}_d$ and $\mathbf{b}_1, \dots, \mathbf{b}_d$ generate the same lattice iff there is a $d \times d$ unimodular matrix

$U = (u_{ij})$ (a matrix with integer entries and determinant $\det(U) = \pm 1$, the inverse of an unimodular matrix is also unimodular) such that

$$\mathbf{b}_j = \sum_{i=1}^{d} u_{ij}\mathbf{a}_i.$$

Let $\mathbf{a}_1, \ldots, \mathbf{a}_d$ be vectors of $\mathbb{R}^d$, and let

$$\mathbf{a}_j = \sum_{i=1}^{d} a_{ij}\mathbf{e}_i$$

for $j = 1, \ldots, d$, where $a_{ij} \in \mathbb{R}$ are the coordinates of $\mathbf{a}_j$ with respect to the standard basis of $\mathbb{R}^d$. The $d \times d$ matrix $A = (a_{ij})$ is the *matrix of the vectors* $\mathbf{a}_1, \ldots, \mathbf{a}_d$. They are linearly independent iff $\det(A) \neq 0$.

The *determinant of the lattice* Λ is defined by

$$\det(\Lambda) = |\det(A)|,$$

where A is the matrix of a basis of Λ. This is independent of the choice of the basis as the matrices of any two bases are connected by an unimodular matrix, and we refer to the product rule of determinants.

The next important notion is the *fundamental parallelepiped* of the lattice Λ with respect to the basis $\mathbf{a}_1, \ldots, \mathbf{a}_d$, it is defined by

$$F(\Lambda) = F(\Lambda; \mathbf{a}_1, \ldots, \mathbf{a}_d) =: \left\{ \sum_{i=1}^{d} x_i\mathbf{a}_i : 0 \leq x_i < 1,\ i = 1, \ldots, d \right\} \subset \mathbb{R}^d.$$

The volume of a fundamental parallelepiped is

$$\mathrm{vol}(F(\Lambda)) = \det(\Lambda),$$

and such, it is independent of the choice of the basis (but not the fundamental parallelepiped itself). Every vector in $\mathbb{R}^d$ has a unique representation as the sum of a lattice vector and an element in the fundamental parallelepiped, $\mathbb{R}^d = \Lambda + F(\Lambda)$. A fundamental parallelepiped contains only one lattice point, the origin in one of the corners.

A set K in $\mathbb{R}^d$ is convex if, for every pairs of vectors $\mathbf{a}, \mathbf{b} \in K$, the line segment $(1-t)\mathbf{a} + t\mathbf{b}$, $0 \leq t \leq 1$ also belongs to K. A body in $\mathbb{R}^d$ is a bounded open set. We shall consider only nonempty convex bodies K with positive volume $\mathrm{vol}(K)$. For a real number $\lambda \geq 0$ we define (actually we have defined)

$$\lambda \cdot K = \{\lambda\,\mathbf{a} : \mathbf{a} \in K\}.$$

Notice that $\mathrm{vol}(\lambda \cdot K) = \lambda^d \mathrm{vol}(K)$. The set K is *symmetric* if $\mathbf{a} \in K$ implies that $-\mathbf{a} \in K$. A symmetric convex body contains the origin.

Lemma 7.1: (Blichfeldt) Let Λ be a lattice in $\mathbb{R}^d$ and K be a body in $\mathbb{R}^d$ with volume $\mathrm{vol}(K) > \det(\Lambda)$. Then there exist vectors $\mathbf{a} \neq \mathbf{b}$ in K such that $\mathbf{a} - \mathbf{b} \in \Lambda$.

Proof: We fix a fundamental parallelepiped F of the lattice Λ. We have $\mathrm{vol}(F) = \det(\Lambda) < \mathrm{vol}(K)$. Since K is bounded and Λ is discrete, there exist only finitely many lattice points $\mathbf{u} \in \Lambda$ with $K \cap (\mathbf{u} + F) \neq \emptyset$. Since $\mathbb{R}^d = \cup_{\mathbf{u}\in\Lambda}(\mathbf{u} + F)$, it follows that $K = \cup_{\mathbf{u}\in\Lambda}(K \cap (\mathbf{u} + F))$, and we have

$$\mathrm{vol}(F) < \mathrm{vol}(K) = \sum_{\mathbf{u}\in\Lambda} \mathrm{vol}(K \cap (\mathbf{u} + F)) = \sum_{\mathbf{u}\in\Lambda} \mathrm{vol}((K - \mathbf{u}) \cap F).$$

Since obviously $(K - \mathbf{u}) \cap F \subset F$, it follows that the sets $(K - \mathbf{u}) \cap F$ cannot be pairwise disjoint, and so there exist distinct lattice points $\mathbf{u}_1 \neq \mathbf{u}_2$ such that $(K-\mathbf{u}_1)\cap(K-\mathbf{u}_2) \neq \emptyset$ That means there exist distict vectors $\mathbf{a} \neq \mathbf{b}$ in K such that $\mathbf{a} - \mathbf{u}_1 = \mathbf{b} - \mathbf{u}_2$, and this conclude the proof as $\mathbf{a} - \mathbf{b} = \mathbf{u}_1 - \mathbf{u}_2 \in \Lambda$.

□

Theorem 7.2: (Minkowski) Let Λ be a lattice in $\mathbb{R}^d$, and let K be a symmetric convex body in $\mathbb{R}^d$ with volume $\mathrm{vol}(K) > 2^d \det(\Lambda)$. Then K contains a lattice point different from the origin.

Proof: Let $K' = (1/2) \cdot K$ be the shrunk to half of K. Then K' is a symmetric convex body with

$$\mathrm{vol}(K') = \frac{\mathrm{vol}(K)}{2^d} > \det(\Lambda).$$

It follows from Lemma 7.1 that there exist vectors $\mathbf{a}' \neq \mathbf{b}'$ in K' such that $\mathbf{a}' - \mathbf{b}' \in \Lambda$. Obviously $\mathbf{a} = 2\mathbf{a}'$ and $\mathbf{b} = 2\mathbf{b}'$ are in K, moreover $-\mathbf{b}$ is in K and

$$\frac{1}{2}\,\mathbf{a} + \frac{1}{2}\,(-\mathbf{b}) = \mathbf{a}' - \mathbf{b}' \in K \cap \Lambda$$

since K is symmetric and convex.

□

The *successive minima* of the convex body $\mathbf{0} \in K$ with respect to the lattice $\Lambda \subset \mathbb{R}^d$ are a sequence of real numbers $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_d$, and defined by

$$\lambda_k = \inf\{\lambda > 0 :\ \lambda \cdot K \text{ contains } k \text{ linearly independent vectors of } \Lambda\}.$$

Because K is an open set, it follows that $\lambda_k \cdot K$ contains at most $k-1$ linearly independent vectors of Λ, while $\lambda_k \cdot \overline{K}$ contains at least k linearly independent vectors of Λ. Not exactly $k-1$ and k because the boundary of $\lambda_k \cdot K$ can contain several linearly independent lattice vectors. (Here $\overline{K}$ is the closure of K.)

Theorem 7.2 can be reformulated as: *Let Λ be a lattice in $\mathbb{R}^d$, and let K be a symmetric convex body in $\mathbb{R}^d$ with volume* $\mathrm{vol}(K) \geq 2^d \det(\Lambda)$. *Then $\overline{K}$ contains a lattice point different from the origin.*

Theorem 7.3: (Minkowski's second theorem) Let K be a symmetric convex body in $\mathbb{R}^d$, and let Λ be a lattice in $\mathbb{R}^d$. Let $\lambda_1, \ldots, \lambda_d$ be the successive minima of K with respect to the lattice Λ. We have

$$\lambda_1 \cdots \lambda_d \operatorname{vol}(K) \leq 2^d \det(\Lambda).$$

Proof: A rigorous proof is somewhat complex and lengthy, and therefore we omit. Here we present a short, heuristic argument, why the theorem holds. (%) will indicate that a claim needs more explanation, or even some more precision. Some of those claims are studied in the homework problems in more details.

An alternative, maybe more practical way of seeing the successive minima, which also identify a natural basis of $\mathbb{R}^d$, is the following. For a sufficiently small λ, $\lambda \cdot \overline{K}$ contains only one lattice vector, the origin, because Λ is a discrete group. By increasing λ we magnify $\overline{K}$ until it hits another lattice vector. This defines λ_1 and we choose one such vector as $\mathbf{b}_1$. We magnify $\overline{K}$ further, until it hits a lattice vector, linearly independent from $\mathbf{b}_1$. This defines λ_2 and we choose one such vector as $\mathbf{b}_2$. We keep doing this, λ_d is the minimal factor of magnification to hit a lattice vector, linearly independent of $\mathbf{b}_1, \ldots, \mathbf{b}_{d-1}$ and we choose one such vector as $\mathbf{b}_d$. Note that magnifying further sometimes means no more magnification, in one step we can hit several linearly independent lattice vectors. The set of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_d$ is a basis of $\mathbb{R}^d$ as being d linearly independent vectors, but not necessary a basis of Λ, they may span a *sublattice* of Λ (%).

Now magnify K by a factor λ_j in the direction of $\mathbf{b}_j$ for all $j = 1, \ldots, d$. In other words we define

$$K^* = \{\sum_{j=1}^{d} \lambda_j x_j \mathbf{b}_j : \sum_{j=1}^{d} x_j \mathbf{b}_j \in K\}.$$

K^* is a symmetric convex body (%) with volume $\operatorname{vol}(K^*) = \lambda_1 \cdots \lambda_d \operatorname{vol}(K)$ (%). Moreover K^* contains no lattice point other than the origin (but there are d linearly independent lattice points, and some other lattice points as well, on the boundary of K^*) (%). Theorem 7.2 implies that $\operatorname{vol}(K^*) \leq 2^d \operatorname{vol}(\Lambda)$, thus proving the theorem.

□

Remark: Theorem 7.3 implies Theorem 7.2. Indeed, we have $\lambda_1^d \operatorname{vol}(K) \leq \lambda_1 \ldots \lambda_d \operatorname{vol}(K) \leq 2^d \operatorname{vol}(\Lambda)$ by Theorem 7.3. Together with $\operatorname{vol}(K) > 2^d \operatorname{vol}(\Lambda)$ this implies $\lambda_1 < 1$, which is equivalent to the fact that there is already a lattice vector, other than the origin, in K.

As a typical application of Minkowski's theorems we give another proof of Dirichlet's approximation theorem, Lemma 5.2. Recall the statement: *Let $k \geq 1$ be an integer and $\alpha_1, \ldots, \alpha_k$ be real numbers. For any integer $Q \geq 1$ there is an integer $1 \leq q \leq Q^k$ such that*

$$\|q\alpha_j\| \leq \frac{1}{Q}, \text{ for all } j = 1, \ldots, k.$$

We define a lattice $\Lambda \subset \mathbb{R}^{k+1}$ by the following matrix.

$$\Lambda = \begin{pmatrix} 1 & \alpha_1 & \dots & \alpha_k \\ & 1 & & \\ & & \vdots & \\ & & & 1 \end{pmatrix} \mathbb{Z}^{k+1} = \begin{pmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_k \end{pmatrix} \mathbb{Z}^{k+1},$$

and we define a symmetric convex body K, actually K is a rectangle, by

$$K = \{(x_0, \dots x_k) : |x_0| < Q^k + 1, \, |x_j| < \frac{1}{Q}, \, j = 1, \dots, k\}.$$

Obviously $\mathrm{vol}(\Lambda) = 1$ and $\mathrm{vol}(K) = 2^{k+1}(Q^k + 1)/Q^k > 2^{k+1}\mathrm{vol}(\Lambda)$, and according to Theorem 7.2, K contains a non–trivial lattice point, say it belongs to the integer vector $(q, -a_1, \dots, -a_k)$, that is $q\mathbf{u}_0 - a_1\mathbf{u}_1 - \dots - a_k\mathbf{u}_k \in K$. Being in K determines an inequality for each coordinates, namely $|q| < Q^k + 1$ and $|q\alpha_j - a_j| < 1/Q$ for all $j = 1, \dots, k$. Note that $q = 0$ is impossible, and without loss of generality we can suppose $1 \le q \le Q^k$, this proves the statement.

□

CHAPTER 8

TECHNICAL PREPARATION: DIRECTED GRAPHS

This chapter discusses some graph results. They are used in the next chapter to provide an alternative, actually the original proof of Plünnecke's theorem, Theorem 1.12. Although the graph theoretic method gives some extra knowledge over Theorem 1.12, we do not need that, so these two chapters can be skipped at the first study of this lecture notes.

A *directed graph* $G = (V, E)$ consists of a finite set V of vertices and a set E of edges, where each edge $e \in E$ is an ordered pair $(u, v) \in V \times V$, and we say e is an edge from u to v. We can write $v \in G$ or $e \in G$ for a vertex or an edge of G. In this lecture note we exclude the possibility of $u = v$, loops from a vertex to itself will not play any role. In case of we want to emphasis that V or E are the vertices or edges of G, we can use the notation $V(G)$ or $E(G)$. A *path from vertex a to vertex b* in G is a finite sequence of vertices $a = v_0, v_1, \dots, v_k = b$ such that (v_{j-1}, v_j) are edges of G for $j = 1, \dots, k$. Equivalently a path can be identified as a sequence of its edges $(v_0, v_1), (v_1, v_2), \dots, (v_{k-1}, v_k)$. Two paths $a = v_0, v_1, \dots, v_k = b$ and $a = w_0, w_1, \dots, w_\ell = b$ from a to b are disjoint if $v_i \neq w_j$ for all *intermediate vertices* $i = 1, \dots, k-1$ and $j = 1, \dots, \ell - 1$. A set S of vertices *separates a from b* if every path from a to b has an intermediate vertex in S. If there are m pairwise disjoint paths from a to b then S must contain an element from each of these paths, and so $|S| \geq m$. It is an important statement that the minimal $|S|$ equals to the maximal m.

Theorem 8.1: (Menger) Let a and b be different vertices of a directed graph $G = (V, E)$ such that $(a, b) \notin E$. Suppose S separates a from b such that $|S|$ is minimal. Let m be the maximal number of pairwise disjoint paths from a to b. Then $|S| = m$.

Proof: We prove by induction on ℓ that *if the minimal size of a set S that separates a from b is ℓ then there are ℓ disjoint paths from a to b.* This proves that $\ell \leq m$, and we are done.

There is nothing to prove for $\ell = 0$ or $\ell = 1$, as we simply claim that there is at least 0 or at least 1 (pairwise disjoint) path from a to b, and indeed, if $\ell = 1$ then the empty set does not separate a from b, so there must be a path from a to b.

The next case $\ell = 2$ is not at all this simple, we move to the general case.

We prove the induction step by contradiction. Let $\ell \geq 2$ and suppose that our claim is true for any directed graph and for any two vertices if they can be separated by a set of size $\leq \ell - 1$, but it is false for ℓ. This means that there is a graph G with vertices a and b such that $(a, b) \notin E(G)$ and there are at most $\ell - 1$ pairwise disjoint path from a to b, but the smallest set S that separates a from b has cardinality exactly ℓ.

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 $G = (V, E)$. Do not be surprised if we find very peculiar properties of G, after all we want to get a

contradiction showing that it does not exist. First we prove a series of claims. The minimal properties of G play a crucial role in the sequel.

(8.1) *Let $e \in E$ be any edge, $G' = (V, E \setminus \{e\})$, and $S(e)$ be a set of vertices of minimum cardinality that separates a from b in G'. Then $|S(e)| = \ell - 1$.*

Note that $e \neq (a, b) \notin E$. Since S separates a from b in G' as well, we have $|S(e)| \leq \ell$, but actually $|S(e)| \leq \ell - 1$ otherwise G' could act as G with fewer edges. If $e = (u, v)$ and $u \neq a$ then $S(e) \cup \{u\}$ separates a from b in G as every path from a to b either intersects $S(e)$ or includes the edge e. Similarly if $v \neq b$ then $S(e) \cup \{v\}$ separates a from b in G. In any case $|S(e)| + 1 \geq \ell$ and the claim is proved.

(8.2) *There is no path of the form a, v, b in G.*

If a, v, b is a path in G then $v \in S$ for any separating set S. Therefore $S' = S \setminus \{v\}$ separates a from b in $G' = G \setminus \{v\}$. (The graph that we get from G by deleting the vertex v with all connected edges.) Similarly, if S' separates a from b in G' then $S' \cup \{v\}$ separates a from b in G. Therefore, $|S'| \geq \ell - 1$ and $|S'| = \ell - 1$ for some separating set in G'. By the induction hypothesis, the graph G' contains $\ell - 1$ pairwise disjoint paths from a to b, they together with the path a, v, b give ℓ pairwise disjoint paths from a to b in G, which is false.

Let S be a separating set in G with minimal cardinality $|S| = \ell$. Further let $P(a, S)$ denote the set of all path from a to a vertex of S with no intermediate vertex in S, and let $I(a, S)$ be the intermediate vertices of paths in $P(a, S)$. Similarly let $P(S, b)$ denote the set of all path from a vertex of S to b with no intermediate vertex in S, and let $I(S, b)$ be the intermediate vertices of paths in $P(S, b)$. Since S separates a from b every path from a to b contains a vertex $s \in S$, 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 $|S|$ is minimal, there is a path from a to s in $P(a, S)$ and a path from s to b in $P(S, b)$ for all $s \in S$.

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

If $v \in I(a, S) \cap I(S, b)$ then there is a path from a to v and a path from v to b which avoid S. This is not possible.

(8.4) *Either $I(a, S) = \emptyset$ or $I(S, b) = \emptyset$.*

Suppose that both sets are nonempty. We define a graph G' by

$$V(G') = \{a\} \cup \{v \in V(G) : v \text{ is a vertex of a path in } P(S, b)\},$$
$$E(G') = \{(a, s) : s \in S\} \cup \{(u, v) \in E(G) : (u, v) \text{ is an edge of a path in } P(S, b)\}.$$

Since $I(a, S) \cap V(G') = \emptyset$ by the claim (8.3) and $I(a, S) \neq \emptyset$ we have $|V(G')| < |V(G)|$. Let S' be a set that separates a form b in G'. The terminal segment of any path from a to b in G is a path from a vertex $s \in S$ to b in G', and the edge (a, s) is also in G', there is a path from a to b in G' with identical terminal segment. Thus S' also separates a from b in G, consequently $|S'| \geq \ell$, there is no smaller separating set than $|S| = \ell$. Since the graph G' has strictly fewer vertices than the graph G, by the minimality property, there are ℓ pairwise disjoint path from a to b in G'. This means that we can connect each vertex

$s \in S$ to b with a path in G such that these ℓ paths are pairwise disjoint. Similarly, we can connect a to each vertex $s \in S$ with a path in G such that these ℓ paths are pairwise disjoint. However, connecting the appropriate (half) paths at the vertices of S we get ℓ pairwise disjoint paths from a to b in G. This is impossible, proving that our initial assumption, that both sets $I(a, S)$ and $I(S, b)$ were nonempty, was false.

(8.5) *If S separates a from b and has cardinality $|S| = \ell$, then either $(a, s) \in E(G)$ for all $s \in S$ or $(s, b) \in E(G)$ for all $s \in S$.*

This is just a reformulation of claim (8.4).

We are ready to finish the proof of our theorem. Let a, u, v be the first three vertices of a shortest path from a to b in G. Note that $(a, v) \notin E$, otherwise we could drop u from the path. By claim (8.2) we have $b \neq v$ and $(u, b) \notin E$. Let $e = (u, v) \in E$ and $S(e)$ as in claim (8.1), thus we have $|S(e)| = \ell - 1 \geq 1$, and so $S(e) \cup \{u\}$ is a separating set in G of minimal cardinality. It follows from claim (8.5) that $(a, s) \in G$ for all $s \in S(e)$, since $(u, b) \notin G$. However, similarly $S(e) \cup \{v\}$ is a separating set of minimal cardinality, and it follows from claim (8.5) that $(s, b) \in G$ for all $s \in S(e)$, since $(a, v) \notin G$. This contradicts to claim (8.2), since $S(e) \neq \emptyset$. The induction is completed and the theorem concludes.

□

Let $G = (V, E)$ be a directed graph, and let X and Y be nonempty, disjoint sets of vertices of G. A *path from the set X to the set Y* in G is a finite sequence of vertices $v_0, v_1, \ldots, v_k$ such that $v_0 \in X$, $v_k \in Y$ and and $(v_{i-1}, v_i) \in E$ for $i = 1, \ldots, k$. Two such paths are *totally disjoint* if they have no common vertex. A set S of vertices *separates the set X from the set Y* if every path from X to Y contains an element of S. Observe that S can intersect X or Y this time.

Theorem 8.2: Let X and Y be nonempty, disjoint sets of vertices of a directed graph G. Let m be the maximal number of pairwise totally disjoint paths from X to Y, and let ℓ be the cardinality of the smallest set S of vertices that separates the set X from the set Y. Then $\ell = m$.

Proof: Let a and b be two new vertices, not belonging to $V(G)$, and define the graph G^* by

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

Then $a = v_0, v_1, \ldots, v_k, v_{k+1} = b$ is a path from a to b in G^* iff $v_1, \ldots, v_k$ is a path from X to Y in G, and two such paths are disjoint in G^* iff the corresponding shortened paths are totally disjoint in G. It follows that m is also the maximal number of pairwise disjoint paths from a to b in G^*. Theorem 8.1 implies that there is a set S such that S separates a from b in G^* and $|S| = m$. By definition $a, b \notin S$, and S separates X from Y in G. This completes the proof.

□

A directed graph $G = (V, E)$ is a *layered graph of level h* if the vertex set V is a disjoint union of $h+1$ nonempty sets $V = V_0 \cup V_1 \cup \cdots \cup V_h$ and if every edge of G connects

a vertex of a layer V_{i-1} to a vertex of the next layer V_i, that is

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

A layered graph G is *commutative* (also called a *Plünnecke graph*) if it satisfies the following two conditions:

(8.6) *if $u \in V_{i-1}$, $v \in V_i$, and $w_1, \ldots, w_k \in V_{i+1}$ are different vertices of G such that u, v, w_j are k different paths in G, then there exist different vertices $v_1, \ldots, v_k$ such that u, v_j, w_j are also k different paths in G.*

(8.7) *if $u_1, \ldots, u_k \in V_{i-1}$, $v \in V_i$, and $w \in V_{i+1}$ are different vertices of G such that u_j, v, w are k different paths in G, then there exist different vertices $v_1, \ldots, v_k$ such that u_j, v_j, w are also k different paths in G.*

The following example shows the importance of commutative layered graphs to our purposes. Let A and B be nonempty finite subsets of the integers (of any additive abelian group), $V_0 = A$, $V_1 = A + B$, ... ,$V_h = A + hB$ and $E = \{(v, v + b) : v \in V_{i-1}, b \in B\}$. Note that the sets V_i and V_j are different copies of integers, even if the sets $A + iB$ and $A + jB$ are not disjoint. We have defined a layered graph of level h this way, and this is a commutative graph, as $u, u+b, u+b+b_j$ are paths in $G = (V, E)$ iff $u, u+b_j, u+b_j+b$ are paths in G. Similarly $u_j, u_j+b_j = v, v+b = w$ are paths in G iff $u_j, u_j+b = v_j, v_j+b_j = w$ are paths. These follow from the commutativity of addition, hence the name. Later on we refer to this graph as the *addition graph of level h over $A + hB$*.

Let X and Y be nonempty subsets of $V(G)$. The *image* of X in Y, denoted by $\mathrm{im}(X, Y)$ 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

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

For a layered graph G of level h with vertex set $V = \cup_{i=0}^{h} V_i$ the magnification ratio of V_0 in V_i is denoted by $\mu_i = \mu_i(G) = D(V_0, V_i)$, $i = 1, \ldots, h$.

We define the *product graph* $G \times G'$ of two layered graphs of level h. Let $V(G) = \cup_{i=0}^{h} V_i$ and $V(G') = \cup_{i=0}^{h} V_i'$. The vertex set of $G \times G'$ is the set

$$V(G \times G') = \bigcup_{i=0}^{h} (V_i \times V_i'),$$

and there is an edge from vertex $(u, u') \in V_{i-1} \times V_{i-1}'$ to the vertex $(v, v') \in V_i \times V_i'$ whenever both $(u, v) \in E(G)$ and $(u', v') \in E(G')$. We can extend this definition to the product of more than two factors. The product of commutative graphs is commutative.

We prove an elegant result about the magnification ratios of graphs of level h.

Lemma 8.3 Let G and G' be layered graphs of level h. We have

$$\mu_i(G \times G') = \mu_i(G)\mu_i(G'), \text{ for } i = 1, \ldots, h.$$

Proof: It is enough to prove the theorem for $h = 1$. Indeed, for a fixed $i = 1, \dots, h$ we can define a layered graph on the vertex sets $V_0(G) \cup V_i(G)$ such that any path from $u \in V_0(G)$ to $v \in V_i(G)$ is simplified to an edge. Obviously $\mathrm{im}(Z, V_i(G))$ is the same in the original and in the simplified graph for any $\emptyset \neq Z \subset V_0(G)$, and thus the magnification ratios agree as well. We do the same with G' and $G \times G'$, and we also notice that the simplification of the product is the product of the simplified graphs.

We consider $G = (V_0 \cup V_1, E)$ and $G' = (V_0' \cup V_1', E')$ and study $\mu(G) = D(V_0, V_1)$ and $\mu(G') = D(V_0', V_1')$. Let Z and Z' be nonempty subsets of V_0 and V_0', respectively such that

$$\mu(G) = \frac{|\mathrm{im}(Z, V_1)|}{|Z|}, \quad \mu(G') = \frac{|\mathrm{im}(Z', V_1')|}{|Z'|}.$$

Obviously

$$Z \times Z' \subset V_0 \times V_0' = V(G \times G')$$

and

$$\mathrm{im}(Z \times Z', V_1 \times V_1') = \mathrm{im}(Z, V_1) \times \mathrm{im}(Z', V_1').$$

It follows that

$$\mu(G \times G') \leq \frac{|\mathrm{im}(Z \times Z', V_1 \times V_1')|}{|Z \times Z'|} = \frac{|\mathrm{im}(Z, V_1)|}{|Z|} \, \frac{|\mathrm{im}(Z', V_1')|}{|Z'|} = \mu(G)\, \mu(G').$$

To complete the proof of the theorem we have to prove the inverse inequality

$$\mu(G \times G') \geq \mu(G)\, \mu(G'). \tag{8.8}$$

Strangely enough we transform the situation to the study of a layered graph of level 2. We construct a graph H with vertices $V(H) = W_0 \cup W_1 \cup W_2$, where

$$\begin{aligned} W_0 &= V_0 \times V_0' = V_0(G \times G'), \\ W_1 &= V_1 \times V_0', \\ W_2 &= V_1 \times V_1' = V_1(G \times G'), \end{aligned}$$

and edges from W_0 to W_1 will consists of all pairs $((a, c), (b, c))$ whenever (a, b) is an edge in G, $c \in V_0'$, and the edges from W_1 to W_2 will consists of all pairs $((b, c), (b, d))$ whenever (c, d) is an edge in G', $b \in V_1$. There is a path from $(v, v') \in W_0$ to $(u, u') \in W_2$ iff there is a path from $(v, v') \in V_0(G \times G')$ to $(u, u') \in V_1(G \times G')$ in $G \times G'$. Therefore

$$\mu_2(H) = D(W_0, W_2) = \mu(G \times G'). \tag{8.9}$$

Let Z be a nonempty subset of $W_0 = V_0 \times V_0'$, and partition Z according to the second coordinates. That is

$$Z = \bigcup_{t \in V_0'} (Z_t \times \{t\}),$$

for $Z_t = \{v \in V_0 : (v,t) \in Z$. Obviously

$$\mathrm{im}(Z, W_1) = \bigcup_{t \in V_0'} (\mathrm{im}(Z_t, V_1) \times \{t\}).$$

Therefore

$$|\mathrm{im}(Z, W_1)| = \sum_{t \in V_0'} |\mathrm{im}(Z_t, V_1)| \geq \mu(G) \sum_{t \in V_0'} |Z_t| = \mu(G)|Z|,$$

and so we have

$$D(W_0, W_1) = \min \left\{ \frac{|\mathrm{im}(Z, W_1)|}{|Z|} : \emptyset \neq Z \subset W_0 \right\} \geq \mu(G). \tag{8.10}$$

Much the same way we also have

$$D(W_1, W_2) \geq \mu(G'). \tag{8.11}$$

If $\mathrm{im}(Z, W_1) = \emptyset$ for some nonempty $Z \subset W_0$, then $\mu(G) = \mu(G \times G') = 0$ and there is nothing to prove. We can suppose that $\mathrm{im}(Z, W_1) \neq \emptyset$ for all $\emptyset \neq Z \subset W_0$. It follows from (8.9), (8.10), and (8.11) that

$$\begin{aligned}
\mu(G \times G') = \mu_2(H) = D(W_0, W_2) = \\
&= \min \left\{ \frac{|\mathrm{im}(Z, W_2)|}{|Z|} : \emptyset \neq Z \subset W_0 \right\} = \\
&= \min \left\{ \frac{|\mathrm{im}(Z, W_1)|}{|Z|} \, \frac{|\mathrm{im}(Z, W_2)|}{|\mathrm{im}(Z, W_1)|} : \emptyset \neq Z \subset W_0 \right\} = \\
&= \min \left\{ \frac{|\mathrm{im}(Z, W_1)|}{|Z|} \, \frac{|\mathrm{im}(\mathrm{im}(Z, W_1), W_2)|}{|\mathrm{im}(Z, W_1)|} : \emptyset \neq Z \subset W_0 \right\} \geq \\
&\geq \min \left\{ \frac{|\mathrm{im}(Z, W_1)|}{|Z|} : \emptyset \neq Z \subset W_0 \right\} \min \left\{ \frac{|\mathrm{im}(\mathrm{im}(Z, W_1), W_2)|}{|\mathrm{im}(Z, W_1)|} : \emptyset \neq Z \subset W_0 \right\} \geq \\
&\geq \min \left\{ \frac{|\mathrm{im}(Z, W_1)|}{|Z|} : \emptyset \neq Z \subset W_0 \right\} \min \left\{ \frac{|\mathrm{im}(Z', W_2)|}{|Z'|} : \emptyset \neq Z' \subset W_1 \right\} = \\
&= D(W_0, W_1)\, D(W_1, W_2) \geq \mu(G)\, \mu(G').
\end{aligned}$$

This completes the proof.

□

Corollary 8.4: Let $G_1, \ldots, G_r$ be layered graphs of level h, $h \geq 1$, $r \geq 2$. We have for all $i = 1, \ldots, h$ that

$$\mu_i(G_1 \times \cdots \times G_r) = \mu_i(G_1) \cdots \mu_i(G_r).$$

□

CHAPTER 9

PLÜNNECKE'S INEQUALITY

In this chapter we develop further the graph theoretic machinery of Chapter 8, and apply these results to estimate $|A+hB|$, where A and B are finite subsets of the integers, say. We use the notions and notations of the previous chapter, especially G is a commutative graph on $h \geq 2$ layers with vertex set $V_0 \cup \cdots \cup V_h$ and magnification ratios μ_i for $i = 1, \dots h$.

In any directed graph $G = (V, E)$ the *out–degree* $d^+(v)$ is the number of edges starting from the vertex $v \in V$, while the *in–degree* $d^-(v)$ is the number of edges ending at the vertex $v \in V$.

Lemma 9.1: Let G be a commutative graph on $h \geq 2$ layers. For all edge (u, v) of G we have

$$d^+(u) \geq d^+(v), \quad d^-(u) \leq d^-(v).$$

Proof: If $d^+(v) = m$ then there are exactly m edges $(v, w_1), \dots, (v, w_m)$ starting from the vertex v, and by the commutativity of G there are vertices $v_1, \dots, v_m$ such that the paths $u, v_1, w_1, \dots, u, v_m, w_m$ all belong to G. This implies $d^+(u) \geq m$, and the first inequality follows. The second is very similar.

□

Lemma 9.2: Let G be a commutative graph on $h \geq 2$ layers with vertex set $V_0 \cup \cdots \cup V_h$. If $\mu_h \geq 1$ then there are $|V_0|$ totally disjoint paths from V_0 to V_h.

Proof: Let m be the maximal number of pairwise totally disjoint paths from V_0 to V_h. By Theorem 8.2 there is a set S of cardinality m that separates V_0 from V_h. We have to prove that $|S| = |V_0|$.

For $v \in V$, let $i(v)$ be the unique integer i such that $v \in V_i$. Choose a separating set S with $|S| = m$ such that

$$\sum_{s \in S} i(s) \tag{9.1}$$

is minimal. We have to prove that $S = V_0$. As a first step towards this goal we show by contradiction that

$$S \subset V_0 \cup V_h. \tag{9.2}$$

Suppose not, that is $S \cap V_\ell = \{s_1, \dots, s_k\} \neq \emptyset$ for some $1 \leq \ell \leq h-1$ and $1 \leq k \leq m$. We also choose m totally disjoint paths from V_0 to V_h through S, as is guaranteed by Theorem 8.2, we refer to these paths as π_i. k of them, say $\pi_1, \dots, \pi_k$ cross the vertices $s_1, \dots, s_k$, let $r_1, \dots, r_k$ be their predecessors and $t_1, \dots, t_k$ be their successors on the paths

$\pi_1, \dots, \pi_k$. If we change the elements $s_1, \dots, s_k$ to the elements $r_1, \dots, r_k$ then we decrease the amount (9.1) so it is no longer a separating set (due to the minimality of S). There is a path, π_0 crossing one of the vertices $s_1, \dots, s_k$ but non of the vertices $r_1, \dots, r_k$. We can suppose this path contains the edge (r_0, s_1), where r_0 is different from $r_1, \dots, r_k$.

Consider the following graph H on 2 layers

$$W_0 = \{r_0, \dots, r_k\}, \quad W_2 = \{t_1, \dots, t_k\},$$

and $W_1 \subset V_\ell$ consists of all vertices that are midpoints of paths from W_0 to W_2 in G. The edges of H are exactly the edges of G from W_0 to W_1 and from W_1 to W_2. Due to the paths r_i, s_i, t_i we have $\{s_1, \dots, s_k\} \subset W_1$. We show that actually $\{s_1, \dots, s_k\} = W_1$. Indeed, the initial segments, that connects V_0 to W_0 of the paths $\pi_0, \dots, \pi_k$ as well as the terminal segments, that connects W_2 to V_h of the paths $\pi_1, \dots, \pi_k$ do not intersect S, and so any path r_i, v, t_j could result a path from V_0 to V_h which avoids S unless $v \in S$.

As H contains all paths from W_0 to W_2 in G, and the definition of commutativity are based on paths (not on edges simply), H is commutative. We can use Lemma 9.1 to the edges (r_i, s_i) and (s_i, t_i) getting (we work in H now)

$$d^+(r_i) \geq d^+(s_i), \quad d^-(t_i) \geq d^-(s_i).$$

Since the number of edges leaving W_0 equals to the number of edges going into W_1 and also the number of edges leaving W_1 equals to the number of edges going into W_2 we have

$$\sum_{i=1}^{k} d^+(r_i) \geq \sum_{i=1}^{k} d^+(s_i) = \sum_{i=1}^{k} d^-(t_i) \geq \sum_{i=1}^{k} d^-(s_i) = \sum_{i=0}^{k} d^+(r_i) \geq 1 + \sum_{i=1}^{k} d^+(r_i),$$

which is a contradiction eventually, so (9.2) is proved.

Next we show that $|S| \geq |V_0|$. This is trivial if $V_0 \setminus S = \emptyset$. If $V_0 \setminus S$ is nonempty, then by (9.2), every path in G from $V_0 \setminus S$ must end in $V_h \cap S$. This is the time to use the condition $\mu_h \geq 1$.

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

and so

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

Therefore

$$|S| = |V_0 \cap S| + |V_h \cap S| \geq |V_0 \cap S| + |V_0 \setminus S| = |V_0|.$$

As V_0 is trivially a separating set, we are done.

□

Lemma 9.3: Let G be a commutative graph on $h \geq 2$ layers. If $\mu_h \geq 1$ then $\mu_i \geq 1$ for all $i = 1, \dots, h$.

Proof: Since $\mu_h \geq 1$ it follows from Lemma 9.2 that there are $|V_0|$ totally disjoint paths from V_0 to V_h, and so $|\mathrm{im}(Z, V_i)| \geq |Z|$ for all nonempty subset Z of V_0.

□

Theorem 9.4: (Plünnecke) Let G be a commutative graph on $h \geq 2$ layers. Then

$$\mu_1 \geq \mu_2^{1/2} \geq \cdots \geq \mu_h^{1/h}.$$

Proof: It is enough to prove that

$$\mu_{h-1} \geq \mu_h^{\frac{h-1}{h}}, \tag{9.3}$$

because cutting the tail of a commutative graph (deleting the last few layers with all connected edges) we get a commutative graph. (9.3) with $h = 2$ proves the first inequality of Theorem 9.4, with $h = 3$ proves the second, and so on. (9.3) is obvious for $\mu_h = 0$, and follows from Lemma 9.3 for $\mu_h = 1$.

For all positive integers n let H_n be the following graph on h layers, $A = \{0\}$, $B = \{1, h+1, (h+1)^2, \ldots, (h+1)^{n-1}\}$, $V_0 = A$, $V_1 = A + B$, $\ldots$,$V_h = A + hB$ and $E = \{(v, v+b) : v \in V_{i-1},\, b \in B\}$. H_n is the addition graph of level h over $A + hB$, as defined in Chapter 8, it is commutative, and as all i–fold sums of elements of B are different, we have $|V_i| = \binom{n+i-1}{i}$. We denote the magnification ratios of the graph H_n by $\mu_i(H_n)$. As $|V_0| = 1$ we have $\mu_i(H_n) = \binom{n+i-1}{i}$.

Suppose that $O < \mu_h < 1$. Let r be any positive integer, and let n be an integer satisfying

$$(h!\mu_h^{-r})^{1/h} < n \leq 2(h!\mu_h^{-r})^{1/h}. \tag{9.4}$$

Consider the product graph of r factors of G and one factor of H_n, that is the graph $G^r \times H_n$. For the magnification ratios of this graph, Corollary 8.4 implies that

$$\mu_h(G^r \times H_n) = \mu_h(G)^r\, \mu_h(H_n) = \mu_h^r \binom{n+h-1}{h} \geq \frac{\mu_h^r n^h}{h!} \geq 1.$$

Next from Lemma 9.3 we have

$$1 \leq \mu_{h-1}(G^r \times H_n) = \mu_{h-1}^r\, \mu_{h-1}(H_n) = \mu_{h-1}^r \binom{n+h-2}{h-1} \leq \mu_{h-1}^r\, n^{h-1}. \tag{9.5}$$

(The last inequality follows from $(n+i)/(i+1) \leq n$.) Combining (9.4) and (9.5) we arrive at

$$\mu_{h-1} \geq n^{-(h-1)/r} \geq (2(h!\mu_h^{-r})^{1/h})^{-(h-1)/r} = (2(h!)^{1/h})^{-(h-1)/r} \mu_h^{(h-1)/h}.$$

This is an inequality for positive real numbers and true for all r, if we let r tend to ∞ we get (9.3) in this case.

The case of $\mu_h > 1$ is very similar in spirit. Let r be any positive integer, and let n be an integer satisfying

$$\frac{1}{2}\, \mu_h^{r/h} < n \leq \mu_h^{r/h}. \tag{9.6}$$

Let H_n^* be the directed graph we get by inverting the direction of each edge in H_n. This is also a commutative graph on h layers, and we have the trivial bounds for its magnification ratios.

$$\mu_h(H_n^*) = \frac{1}{\binom{n+h-1}{h}} \geq n^{-h},$$
$$\mu_{h-1}(H_n^*) \leq \frac{|B|}{|hB|} = \frac{n}{\binom{n+h-1}{h}} \leq h! n^{-h+1}.$$

Again from Corollary 8.4 we have

$$\mu_h(G^r \times H_n^*) = \mu_h^r \, \mu_h(H_n^*) \geq \mu_h^r \, n^{-h} \geq 1,$$

and from Lemma 9.3 we have

$$1 \leq \mu_{h-1}(G^r \times H_n^*) = \mu_{h-1}^r \, \mu_{h-1}(H_n^*) \leq \mu_{h-1}^r \, h! n^{-h+1}.$$

Combining this inequality with (9.6) we get

$$\mu_{h-1} \geq (h!)^{-1/r} \, n^{(h-1)/r} \geq (h!)^{-1/r} \big(\frac{1}{2} \, \mu_h^{r/h}\big)^{(h-1)/r} = (2^{h-1} h!)^{-1/r} \mu_h^{(h-1)/h}.$$

This is an inequality for positive real numbers and true for all r, if we let r tend to ∞ we get (9.3) in the remaining case.

□

Theorem 9.5: (Plünnecke) Let A and B be finite subsets of the real numbers (of any abelian group), and let $1 \leq i \leq h$ be integers. If $|A + iB| \leq c|A|$ then there is a subset $Z \subset A$ such that

$$|Z + hB| \leq c^{h/i}|Z|.$$

Proof: Let G be the addition graph of level h over $A + hB$. By definition, there is a subset $\emptyset \neq Z \subset A$ such that

$$\mu_h = \frac{|Z + hB|}{|Z|}.$$

Theorem 9.4 implies that

$$|Z + hB| = |Z|\mu_h \leq |Z|\mu_i^{h/i} \leq |Z|\big(\frac{|A + iB|}{|A|}\big)^{h/i} \leq c^{h/i}|Z|.$$

□

Theorem 9.6: Let A and B be finite sets of the real numbers (of any abelian group). For all $h \geq 2$ we have

$$\frac{|hB|}{|A|} \leq \left(\frac{|A + B|}{|A|}\right)^h.$$

Proof: Let $c = |A+B|/|A|$ and we apply Theorem 9.5 with $i = 1$. Then there is a nonempty $Z \subset A$ such that

$$|hB| \le |Z + hB| \le c^h |Z| \le c^h |A|.$$

□

Corollary 9.7: Let B be a finite set of the real numbers (of any abelian group). For all $h \ge 2$ we have

$$\frac{|hB|}{|B|} \le \left(\frac{|B \pm B|}{|B|}\right)^h.$$

□

Corollary 9.8: (already proved in (1.13)) Let B be a finite set of the real numbers (of any abelian group). We have

$$\frac{|B+B|}{|B|} \le \left(\frac{|B-B|}{|B|}\right)^2.$$

□

Theorem 1.12: Let B be a finite set of the real numbers (of any abelian group). We have

$$\frac{|mB-nB|}{|B|} \le \left(\frac{|2B|}{|B|}\right)^{m+n}$$

for all $m, n \ge 1$.

Proof: Let $c = |2B|/|B|$. We can suppose that $m \le n$ because $|mB-nB| = |nB-mB|$. We apply Theorem 9.5 with $A = B$ and $i = 1$. Then there is a nonempty $Z \subset B$ such that

$$|Z + mB| \le c^m |Z|. \tag{9.7}$$

Now apply Theorem 9.5 again with $A = Z$ and $i = m$. Then $|A + iB| \le c^m |A|$ and there is a nonempty $Y \subset Z$ such that

$$|Y + nB| \le (c^m)^{n/m} |Y| = c^n |Y|. \tag{9.8}$$

Recall Theorem 1.4, the triangle inequality, for any finite sets $|U||V-W| \le |V-U||U-W|$. Let $U = -Y$, $V = mB$ and $W = nB$, we get from (9.7) and (9.8) that

$$|Y|||mB-nB| \le |Y+mB||Y+nB| \le |Z+mB||Y+nB| \le c^m|Z|c^n|Y| \le c^{m+n}|B||Y|,$$

and the theorem follows.

□

CHAPTER 10

FREIMAN'S THEOREM

In this chapter we prove Freiman's theorem, in the 1950's he recognized, if $A + A$ is small, then A has a special structure, it is a subset of a not much bigger generalized arithmetic progression of small dimension. The original proof was long and complicated, but in the 1990's Ruzsa and Bilu independently found more accessible proofs. Our presentation is based on Ruzsa's work.

Remember the definition of a generalized arithmetic progression, given in Chapter 1.

$$Q = \{a_0 + a_1 n_1 + a_2 n_2 + \cdots + a_d n_d : 0 \le n_j < N_j \text{ for } j = 1, \dots, d\}, \tag{10.1}$$

where a_0, a_1, ... ,a_d are fixed real numbers (or integers, or residue classes, or ...) and N_1, N_2, ... ,N_d are integers ≥ 1. This is a *generalized arithmetic progression* of *dimension* $\dim Q = d$ and *volume* $\mathrm{vol}(Q) = N_1 \cdots N_d$. Obviously $|Q| \le \mathrm{vol}(Q)$. Q is *proper* if all the above elements are different, in which case $|Q| = \mathrm{vol}(Q) = N_1 \cdots N_d$. Realize that $Q + Q$ is another generalized arithmetic progression, namely

$$Q + Q = \{2a_0 + a_1 n_1 + a_2 n_2 + \cdots + a_d n_d : 0 \le n_j < 2N_j - 1 \text{ for } j = 1, \dots, d\}$$

(this is not necessary proper, even if Q is proper). Nevertheless $|Q + Q| \le (2N_1 - 1) \cdots (2N_d - 1)$ and if Q is proper then $|Q + Q| < 2^d |Q|$. Simply speaking a generalized arithmetic progression has small doubling, this result is ready to reverse.

Theorem 10.1: (Freiman, Ruzsa) Let A be a finite subset of the integers, and $C \ge 2$. If $|A + A| \le C|A|$ then there is a generalized arithmetic progression Q such that

$$A \subset Q, \quad \mathrm{vol}(Q) \le F(C)|A|, \quad \dim Q \le d(C), \tag{10.2}$$

with explicitly calculable functions $F(C)$ and $d(C)$.

Remark 10.2: The best proved bounds are $d(C) \le C$ and $F(C) \le e^{O(C^2 \log C)}$. Following closely our proof one can derive explicit bounds, but they are worse with one exponential, that is $d(C) \le e^{O(C^{48})}$ and $F(C) \le e^{e^{O(C^{48})}}$.

We return later to the proof of Theorem 10.1 as it is split into smaller steps, first we prove these steps one by one. All the steps, as well as the completion of the proof require much ingenuity, our main ingredients are, among others, the circle method (Chapter 5), geometry of numbers (Chapter 7) and Plünnecke's inequality (Theorem 1.12 or Chapter 9).

Theorem 10.3: (Bogolyubov, Ruzsa) Let N be a (large) prime, $0 < \alpha < 1$ be a real number, and $A \subset \mathbb{Z}_N$, $|A| = \alpha N$. Then $2A - 2A$ contains a proper generalized arithmetic progression of $\dim \le \alpha^{-2}$ and $\mathrm{vol} \ge (\alpha^2/4)^{\alpha^{-2}} N$.

Proof: Let us define (the Fourier coefficients)

$$\hat{A}(r) = \sum_{a\in A} e\big(\frac{ar}{N}\big),$$

and let $\mathcal{R}$ be the set where $\hat{A}(r)$ is big.

$$\mathcal{R} =: \{r \in \mathbb{Z}_N : |\hat{A}(r)| \geq \alpha^{3/2}N\}.$$

Parseval identity (Lemma 6.4) gives the opportunity to estimate the size of $\mathcal{R}$, indeed we have

$$\alpha N^2 = N|A| = \sum_{r\in\mathbb{Z}_N} |\hat{A}(r)|^2 \geq \sum_{r\in\mathcal{R}} |\hat{A}(r)|^2 \geq |\mathcal{R}|\alpha^3 N^2,$$

and so

$$|\mathcal{R}| \leq \frac{1}{\alpha^2}. \tag{10.3}$$

We are going to study the representation function $r_{2A-2A}(x)$ and we will see that there is a generalized arithmetic progression (with parameters described above) on which $r_{2A-2A}(x) > 0$. By the circle method (actually by (6.6)) we have

$$r_{2A-2A}(x) = \frac{1}{N}\sum_{r\in\mathbb{Z}_N} |\hat{A}(r)|^4 e\big(\frac{rx}{N}\big) = \frac{1}{N}\sum_{r\in\mathbb{Z}_N} |\hat{A}(r)|^4 \cos\big(2\pi\frac{rx}{N}\big), \tag{10.4}$$

where we took the real part, as the left hand side is real. We restrict ourselves to special residue classes x, namely those with

$$\Big\|\frac{rx}{N}\Big\| \leq \frac{1}{4}, \tag{10.5}$$

for all $r \in \mathcal{R}$. (Recall that $\|y\|$ denotes the distance of y to the nearest integer.) As $\cos(2\pi\frac{rx}{N}) \geq 0$ for all such x and all $r \in \mathcal{R}$, we have

$$r_{2A-2A}(x) \geq \frac{1}{N}|\hat{A}(0)|^4 - \frac{1}{N}\sum_{r\notin\mathcal{R}} |\hat{A}(r)|^4 >$$
$$> \frac{1}{N}|\hat{A}(0)|^4 - \frac{1}{N}\big(\alpha^{3/2}N\big)^2 \sum_{r\in\mathbb{Z}_N} |\hat{A}(r)|^2 = \alpha^4 N^3 - \alpha^3 N\alpha N^2 = 0.$$

In other words, $2A-2A$ contains all $x \in \mathbb{Z}_N$ for which (10.5) holds with every $r \in \mathcal{R}$. We are going to show that they contain a generalized arithmetic progression. (Note that Lemma 5.2 provide a true arithmetic progression of length $N^{1/|\mathcal{R}|} \geq N^{\alpha^2}$ in this set of x.)

This is the place where the geometry of numbers enters the picture. Let $d = |\mathcal{R}|$ and let $r_1, \dots, r_d$ be a fixed set of representatives of the residue classes $\mathcal{R}$. Let $\mathbf{r} = (r_1, \dots, r_d)$

be the integer vector composed from the elements of $\mathcal{R}$. Coordinates here and below are understood in the standard orthogonal basis of $\mathbb{R}^d$. Finally let

$$\Lambda = N \cdot \mathbb{Z}^d + \mathbf{r} \cdot \mathbb{Z} = \{(a_1 N + br_1, \dots, a_d N + br_d) : a_j, b \in \mathbb{Z}\} \subset \mathbb{Z}^d \subset \mathbb{R}^d.$$

Λ is a lattice, and $\det(\Lambda) = N^{d-1}$. To see this, note that Λ is a subgroup, and that $N \cdot \mathbb{Z}^d$ is a subgroup of Λ. Obviously $\det(N \cdot \mathbb{Z}^d) = N^d$, and a fundamental parallelepiped of $N \cdot \mathbb{Z}^d$ is the cube with side N. It contains exactly N points of Λ, namely $0 \cdot \mathbf{r}, 1 \cdot \mathbf{r}, \dots, (N-1) \cdot \mathbf{r}$ after each coordinates are reduced to the least nonnegative residue mod N. Thus N translates of a fundamental parallelepiped of Λ fill the N sided cube completely (mod N, that is you may have to cut and paste pieces).

Next let $K = \{(u_1, \dots, u_d) : |u_j| < 1\}$ be a symmetric, convex body with $\mathrm{vol}(K) = 2^d$. Let $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_d$ be the sequence of successive minima of K with respect to Λ. As in the sketch of the proof of Theorem 7.3 (Minkowski's second theorem), there are linearly independent vectors $\mathbf{b}_1, \dots, \mathbf{b}_d$ of Λ such that $\mathbf{b}_j$ is on the boundary of $\lambda_j \cdot K$, that is all coordinates of $\mathbf{b}_j$ is bounded in absolute value by λ_j. But $\mathbf{b}_j \in \Lambda$ as well, so $\mathbf{b}_j = s_j \cdot \mathbf{r} + \mathbf{v}_j$ where $0 \le s_j < N$ is an integer, while $\mathbf{v}_j \in N \cdot \mathbb{Z}^d$ is an integer vector, all its coordinates are multiples of N. Moreover

$$\left\| \frac{s_j \mathbf{r}}{N} \right\| = \left\| \frac{\mathbf{b}_j}{N} \right\| \le \frac{\lambda_j}{N}, \tag{10.6}$$

where the above relations are understood coordinatewise.

We are in the position to define our generalized arithmetic progression P.

$$P = \{a_1 s_1 + \cdots + a_d s_d : |a_j| \le \frac{N}{4d\lambda_j}\}.$$

If $x \in P$ then $x = a_1 s_1 + \cdots + a_d s_d$ and

$$\left\| \frac{x\mathbf{r}}{N} \right\| \le \sum_{j=1}^{d} |a_j| \left\| \frac{s_j \mathbf{r}}{N} \right\| \le \sum_{j=1}^{d} \frac{N}{4d\lambda_j} \frac{\lambda_j}{N} = \frac{1}{4},$$

and so $x \in 2A - 2A$, indeed $P \subset 2A - 2A$. We show that P is proper. Suppose that $x = x'$ where $x = a_1 s_1 + \cdots + a_d s_d \in P$ and $x' = a'_1 s_1 + \cdots + a'_d s_d \in P$, that is $0 = (a_1 - a'_1)s_1 + \cdots + (a_d - a'_d)s_d$. Thus (relations are understood coordinatewise) we have

$$\mathbf{0} = (\sum (a_j - a'_j) s_j) \cdot \mathbf{r} = \sum (a_j - a'_j)(\mathbf{b}_j - \mathbf{v}_j) \equiv \sum (a_j - a'_j)\mathbf{b}_j \pmod{N}.$$

Since all coordinates of $\mathbf{b}_j$ is bounded by λ_j, and $|a_j - a'_j| \le N/(2d\lambda_j)$ we have that all coordinates of $\sum (a_j - a'_j)\mathbf{b}_j$ is bounded by $N/2$, and as all coordinates are divisible by N, they are equal to 0. That is $\sum (a_j - a'_j)\mathbf{b}_j = \mathbf{0}$, and from the linear independence of the vectors $\mathbf{b}_j$ we conclude $a_j - a'_j = 0$ for all j.

Finally we estimate the volume of P. Since Theorem 7.3 implies $\lambda_1 \dots \lambda_d \operatorname{vol}(K) \le 2^d \det(\Lambda)$, and since any interval of the form $(-y, y)$ contains at least $y/2$ integers, we have that

$$\operatorname{vol}(P) \ge \prod_{j=1}^{d} \frac{N}{4d\lambda_j} \ge \left(\frac{N}{4d}\right)^d \frac{\operatorname{vol}(K)}{2^d \det(\Lambda)} = \left(\frac{1}{4d}\right)^d N.$$

With (10.3) this concludes the proof.

□

We want to establish connection between $\mathbb{Z}$ and $\mathbb{Z}_N$. To this end we introduce the notion of a k–isomorphism (also called Freiman isomorphism).

Definition 10.4: Let $k \ge 2$ be an integer, and $A \subset Z_1$ and $B \subset Z_2$, where Z_1 and Z_2 are two (additive) abelian groups. A bijection $\phi : A \to B$ is a k–isomorphism if

$$a_1 + \dots + a_k = a'_1 + \dots + a'_k \text{ iff } \phi(a_1) + \dots + \phi(a_k) = \phi(a'_1) + \dots + \phi(a'_k).$$

Note that ϕ is not necessarily defined on the whole group Z_1, not alone on $a_1 + \dots + a_k$, when it is not in A. A k–isomorphism is an ℓ–isomorphism as well, for all $\ell < k$. Obviously the inverse of a k–isomorphism is a k–isomorphism, and $|A| = |B|$. The 2–isomorphic image of a (proper) generalized arithmetic progression is a (proper) generalized arithmetic progression of the same dimension and volume. This follows from the repeated application of the fact that $a_1 - a_2 = a_3 - a_4$ iff $a_1 + a_4 = a_2 + a_3$ iff $\phi(a_1) + \phi(a_4) = \phi(a_2) + \phi(a_3)$ iff $\phi(a_1) - \phi(a_2) = \phi(a_3) - \phi(a_4)$. If B is a $k(m+n)$–isomorphic image of A then $mB - nB$ is a k–isomorphic image of $mA - nA$. Indeed, the map $\phi : A \to B$ induces a map $\phi^* : mA - nA \to mB - nB$ by

$$\phi^*(a_1 + \dots + a_m - a_{m+1} - \dots - a_{m+n}) = \phi(a_1) + \dots + \phi(a_m) - \phi(a_{m+1}) - \dots - \phi(a_{m+n}),$$

since ϕ is an $(m+n)$–isomorhism, this is well defined, and since ϕ is a $k(m+n)$–isomorhism, this is a k–isomorphism.

Theorem 10.5: (Ruzsa's embedding theorem) Let $k \ge 2$ be an integer and let $A \subset \mathbb{Z}$ be a finite set of integers. If $N > 2|kA - kA|$ is a prime, then there is a subset $A' \subset A$ and a k–isomorphism $\phi : A' \to B \subset \mathbb{Z}_N$, where $|A'| \ge |A|/k$.

Proof: With simple terms, any subset A of the integers contains a rather "big" subset A', such that A' has a k–isomorphic image B in $\mathbb{Z}_N$, where N is a well chosen not too large prime. We construct this isomorphism in four steps.

First step. We choose a prime $p > k(\max A - \min A)$, and map the elements of A into the residue classes of $\mathbb{Z}_p$, they represent. This is a k–isomorphism, because $L =: k \min A \le a_1 + \dots + a_k \le k \max A < L + p$ and two such integers are equal iff they are congruent mod p.

Second step. For an integer $1 \le q \le p-1$ we map $a \in \mathbb{Z}_p$ to $qa \in \mathbb{Z}_p$. This is a k–isomorphism, obviously. Our aim with this step is to shuffle the image of A in $\mathbb{Z}_p$ to a more uniform position. Let us denote the image of A after these two isomorphisms by $A_q \subset \mathbb{Z}_p$.

Third step. We map $A_q \subset \mathbb{Z}_p$ to their least non negative residues in $\mathbb{Z}$. This is not a k–isomorphism, obviously. However, there is an interval of length p/k containing at least $|A|/k$ images of A_q, we denote these images by $A_q'' \subset \mathbb{Z}$, and we denote the pre–image of A_q'' by $A_q^* \subset A_q \subset \mathbb{Z}_p$. Note that $M =: \min A_q'' \le \max A_q'' < M + \frac{p}{k}$, our map is a k–isomorphism between A_q^* and A_q'' for the same reason as in the first step.

Fourth step. We choose a prime $N > 2|kA - kA|$, and map the elements of A_q'' into B_q, the reside classes of $\mathbb{Z}_N$, they represent. This is not a k–isomorphism usually, however we show that there is a q for which this is. To this end we compute the number of bad choices for the role of q. The inverses of the first three isomorphisms define $A_q' \subset A$, the isomorphic image of A_q'' in A. We have $|A_q'| > |A|/k$. We denote the composition of all four maps by ϕ, that is $\phi : A_q' \to B_q \subset \mathbb{Z}_N$. q is a bad choice if there are

$$0 \neq c = a_1 + \cdots + a_k - a_1' - \cdots - a_k' \in kA_q' - kA_q' \subset kA - kA,$$

such that

$$\phi(a_1) + \cdots + \phi(a_k) - \phi(a_1') - \cdots - \phi(a_k') \equiv 0 \pmod N$$

in B_q. Especially, the number of different choices for c is bounded by $\le |kA - kA|$. The images of $a_1, \ldots, a_k'$ in $A_q'' \subset \mathbb{Z}$ after the first three isomorphisms are $c_1, \ldots, c_k'$ respectively. They satisfy

$$r =: c_1 + \cdots + c_k - c_1' - \cdots - c_k' \equiv qa_1 + \cdots + qa_k - qa_1' - \cdots - qa_k' = qc \not\equiv 0 \pmod p,$$

the last because of A_q^* is isomorphic to A_q', and we also conclude $0 < |r| < p$, while

$$r = c_1 + \cdots + c_k - c_1' - \cdots - c_k' \equiv \phi(a_1) + \cdots + \phi(a_k) - \phi(a_1') - \cdots - \phi(a_k') \equiv 0 \pmod N.$$

From this last congruence (and the bound for the size of r), the number of choices for r is at most $\le 2(p-1)/N$. From the previous congruence $q \equiv r/c \pmod p$, and so the number of different choices for a bad q is bounded by

$$\le \frac{2(p-1)}{N} |kA - kA| < p - 1,$$

there must be a q for which ϕ is a k–isomorphism. Let $A' = A_q'$ and $B = B_q$ with such a q. This concludes the proof of our theorem.

□

Proof of Theorem 10.1: Let A be a nonempty finite subset of the integers satisfying $|A + A| \le C|A|$. Theorem 1.12 (Plünnecke's inequality) implies that $|8A - 8A| \le C^{16}|A|$. We fix a prime N satisfying $2C^{16}|A| < N \le 4C^{16}|A|$. Theorem 10.5 implies that there is an $A' \subset A$ and an 8–isomorphism $\phi : A' \to B \subset \mathbb{Z}_N$, where $|A'| \ge |A|/8$. Note that $|B| = |A'| \ge |A|/8 \ge N/(32C^{16}) =: \alpha N$. Theorem 10.3 (Bogolyubov) implies that there is a proper generalized arithmetic progression $P \subset 2B - 2B$ with $\dim P \le \alpha^{-2}$ and $\mathrm{vol}(P) \ge (\alpha^2/4)^{\alpha^{-2}} N$. The (inverse) isomorphic image of P is a proper generalized arithmetic progression $P' \subset 2A' - 2A' \subset 2A - 2A$ with the same dim and vol. We write

$\dim P' = \delta \le \alpha^{-2}$ and $\mathrm{vol}(P') = \gamma|A| \ge (\alpha^2/4)^{\alpha^{-2}} N \ge 2C^{16}(\alpha^2/4)^{\alpha^{-2}}|A|$. Note also that $\gamma|A| = \mathrm{vol}(P') \le |2A - 2A| \le C^4|A|$.

Let $S \subset A$ be maximal to the property that the sets $s + P'$ are pairwise disjoint for all $s \in S$. Obviously $|S||P'| = |S + P'|$, moreover $S + P' \subset A + P' \subset 3A - 2A$ and by Plünnecke's inequality again, we have

$$|S| \le \frac{|3A - 2A|}{|P'|} \le \frac{C^5|A|}{\mathrm{vol}(P')} = \frac{C^5}{\gamma}.$$

From the maximality of S, for all $a \in A$ there is an $s \in S$ such that $(a+P') \cap (s+P') \neq \emptyset$. This means that $a + t_1 = s + t_2$ where $t_1, t_2 \in P'$, that is $a = t_2 - t_1 + s$ and so

$$A \subset P' - P' + S \subset Q,$$

where, if $P' = \{a_0 + a_1 n_1 + \cdots + a_\delta n_\delta : 0 \le n_j < N_j\}$ then

$$Q = \{\sum_{j=1}^{\delta} a_j n_j + \sum_{s \in S} \epsilon_s s : |n_j| < N_j,\, 0 \le \epsilon_s < 2\},$$

is a generalized arithmetic progression with $\dim Q = \delta + |S|$ and $\mathrm{vol}(Q) \le 2^\delta \gamma |A| 2^{|S|}$. This completes the proof.

□

It is recommended to follow closely the proof in order to compute how the parameters α, δ, γ , and $|S|$ depends solely on C, finally to find an expression for $d(C)$ and $F(C)$.

These bounds can be significantly improved by the following observation, due to Chang (2002). The poor bounds in the above argument come from the enormous size of S, we will improve it by replacing S and Q above to a much smaller S' and a somewhat bigger Q'. We use an iterative method. Let $S_0 = S$ and $Q_0 = Q$ be as in the last paragraph of the proof, and let $m > C$ be an integer chosen later. If S_j and Q_j are constructed then do the following. If $|S_j| \le m$ then we stop and let this j is denoted by J. If $|S_j| > m$ then select any subset $T_j \subset S_j$ with $|T_j| = m$ and let $Q_{j+1} = T_j + Q_j$. (Ignore the fact that these are not anymore generalized arithmetic progressions.) Nevertheless we select $S_{j+1} \subset A$ in the same way, that is a maximal set with $s + Q_{j+1}$ are pairwise disjoint. We have $|S_{j+1}||Q_{j+1}| = |S_{j+1} + Q_{j+1}|$. Note that this also imply (from the previous step) that $m|Q_j| = |T_j||Q_j| = |T_j + Q_j| = |Q_{j+1}|$, and by induction $|Q_j| = m^j|Q|$. On the other hand $Q_{j+1} \subset S_j + Q_j \subset A + Q_j \subset \cdots \subset (j+1)A + Q \subset (j+3)A - 2A$. By Plünnecke's inequality $|S_j| \le |S_j + Q_j|/|Q_j| \le (C/m)^j\, C^5|A|/|Q|$, since this is decreasing to 0 as j increases, we have to stop at some J. From the same computation we have $m^J|Q| = |Q_J| \le C^{J+4}|A|$, that is $(m/C)^J \le C^4/\gamma$.

From the construction

$$A \subset S_J + Q_J - Q_J \subset S_J + \sum_{j=0}^{J-1} (T_j - T_j) + Q - Q,$$

which is a generalized arithmetic progression of $\dim = |S_J| + mJ + \delta \le (m+1)J + \alpha^{-2}$ and $\mathrm{vol} \le 2^{|S_J|} 3^{mJ} \gamma |A|$. The reader is invited to fill up the details in this argument, to find the optimal choice of m, and to trace through for getting the bounds of $d(C)$ and $F(C)$.

CHAPTER 11

A STATISTICAL RESULT

Let Z be an additive abelian group ($\mathbb{Z}$, for example), $A \subset Z$ and $B \subset Z$ are finite subsets. Suppose there are "many" additive quadruples in the form $a_1 + b_1 = a_2 + b_2$, that is there are many solutions to this equation, where $a_1, a_2 \in A$ and $b_1, b_2 \in B$. Is this a severe condition on A and B?

Alternatively let $G \subset A \times B$, and $S = \{a + b : (a, b) \in G\}$. The best way to view G is that G is the set of the edges of a bipartity graph on the two vertex sets A and B, as disjoint vertex sets, (even if they are not disjoint as subsets of Z). Than S is the set of the sums of the two vertices along each edges of G. Suppose that $|S|$ is "small" while $|G|$ is "large". Is this a severe condition on A and B?

In this chapter we prove that they indeed imply that A and B have to have some "weak" structure. There is a family of statements of this kind. We have already expressed the basic condition in two forms, and the results also vary according to the precise meaning of the words "many", "small", "large" and "weak" above. The results follow the scheme, that there are "large" subsets $A' \subset A$ and $B' \subset B$ such that the complete sum set $A' + B' = \{a + b : a \in A', b \in B'\}$ is "small". Roughly speaking, partial information on the sum set imply the existence of subsets with complete information. One can now appeal to the structure theorem of Freiman (Chapter 10) to describe A' and B' more precisely.

Theorem 11.1: Let Z be an additive abelian group, $A \subset Z$, $B \subset Z$, $G \subset A \times B$, and $S = \{a + b : (a, b) \in G\}$. If $|A|, |B|, |S| \leq N$ and $|G| = \alpha N^2$ then there are $A' \subset A$ and $B' \subset B$ such that

$$\text{(11.1)} \qquad \text{i.) } |A' + B'| \leq \frac{2^{15}}{\alpha^5} N, \quad \text{ii.) } |A'|, |B'| \geq \frac{\alpha}{8} N, \quad \text{iii.) } |(A' \times B') \cap G| \geq \frac{\alpha^2}{32} N^2.$$

Theorem 11.2 Let Z be an additive abelian group, $A \subset Z$, $B \subset Z$, and $|A|, |B| \leq N$. If there are $\geq 2\alpha N^3$ solutions to the equation $a_1 + b_1 = a_2 + b_2$, in $a_1, a_2 \in A$ and $b_1, b_2 \in B$, then there are $A' \subset A$ and $B' \subset B$ such that

$$\text{i.) } |A' + B'| \leq \frac{2^{15}}{\alpha^5} N, \quad \text{ii.) } |A'|, |B'| \geq \frac{\alpha}{8} N.$$

First we show that Theorem 11.2 follows from Theorem 11.1 by the use of a simple averaging argument. Suppose that the condition of Theorem 11.2 holds and remember the definition from Chapter 1, $r_{A+B}(n) = \#\{n = a + b : a \in A, b \in B\} \leq N$, moreover let $S = \{n : r_{A+B}(n) \geq \alpha N\}$. Note that $\sum r_{A+B}(n) = |A||B| \leq N^2$. We have

$$2\alpha N^3 \leq \sum_{a_1+b_1=a_2+b_2} 1 = \sum_n r_{A+B}(n)^2 = \sum_{n \in S} r_{A+B}(n)^2 + \sum_{n \notin S} r_{A+B}(n)^2 \leq$$
$$\leq N \sum_{n \in S} r_{A+B}(n) + \alpha N \sum_{n \notin S} r_{A+B}(n) \leq N \sum_{n \in S} r_{A+B}(n) + \alpha N^3.$$

If S is so abundant that $|S| > N$ then we can drop a few elements to reach $|S| = N$ and still maintain that $\sum_{n\in S} r_{A+B}(n) \geq \alpha N^2$. Finally let $G = \{(a,b) : a+b \in S\}$, $|G| = \sum_{n\in S} r_{A+B}(n)$ and the conditions of Theorem 11.1 are satisfied.

□

Actually, the whole proof of Theorem 11.1 is no more than the repeated use of this same averaging argument, so we state it here independently.

Lemma 11.3: Let S be any set with $|S| \leq N$ and $f : S \to \mathbb{R}$ be any function with $f(s) \leq M$.

$$\text{If} \quad \sum_{s\in S} f(s) \geq \alpha MN, \quad \text{and} \quad S' = \{s \in S : f(s) \geq \frac{\alpha}{2}M\}$$

$$\text{then} \quad \sum_{s\in S'} f(s) \geq \frac{\alpha}{2}MN \quad \text{and} \quad |S'| \geq \frac{\alpha}{2}N. \tag{11.2}$$

Proof: One can verify (11.2) in a few lines just copying the above argument.

$$\alpha MN \leq \sum_{s\in S} f(s) = \sum_{s\in S'} f(s) + \sum_{s\notin S'} f(s) \leq \sum_{s\in S'} f(s) + \frac{\alpha}{2}M|S|,$$

$$\frac{\alpha}{2}MN \leq \sum_{s\in S'} f(s) \leq M|S'|.$$

□

Proof of Theorem 11.1: Any a resp. b with or without subscript or superscript will always denote elements of A resp. B. Our plan is the following. If we can write

$$a + b = (a + b') - (a' + b') + (a' + b) \tag{11.3}$$

such that all three terms on the RHS are sums from S, and in addition we have $\gg N$ such choices of a' and $\gg N$ such choices of b' then we have $\gg N^2$ representations

$$a + b = s_1 - s_2 + s_3.$$

They can be the same for different pairs a, b, but they are certainly different for different values of $a + b$. As there are no more than N^3 different three–tuples of elements of S, this prevents us to have too many different values of $a + b$. Of course, we have to restrict our choice of a and b to provide the numerous representations of the required manner, but this is the only thing we have to concentrate on.

We define a bipartity graph G_1 as is described earlier. That is the left hand side vertex set A_1 is identified with A, the right hand side vertex set B_1 is identified with B and a, b are connected with an edge if $(a, b) \in G$. Our strategy is to reduce this graph step by step by dropping vertices and/or edges to gain better properties. The subsequent graphs, as well as its set of edges will be denoted by G_n, and the corresponding vertex sets by A_n and B_n. We do not go further than G_4.

The most important constituents of our graphs are R_a resp. R_b, the set of vertices on the other side connected to a resp. b, we call them the "shadow" of a resp. the shadow of b. Note that R_a and R_b depend of the actual graph, we have to follow their changes with some care.

Note that $|R_a| \leq N$ and we have that

$$\sum_{a \in A} |R_a| = |G_1| = |G| = \alpha N^2,$$

define $A_2 = \{a \in A : |R_a| \geq \frac{\alpha}{2} N\}$, the set of vertices with big shadows. We do not touch the other side, $B_2 = B_1 = B$. (11.2) implies that

$$|G_2| = \sum_{a \in A_2} |R_a| \geq \frac{\alpha}{2} N^2, \quad |R_a| \geq \frac{\alpha}{2} N \text{ in } G_2. \tag{11.4}$$

The next reduction, still on the A side, is the most delicate step of the whole process. Remember our goal, (11.3). We want that the three terms on the right hand side correspond to edges in G_2. Now, if a and b are fixed and we picked an $a' \in R_b$, then any choice of $b' \in R_a \cap R_{a'}$ will work. This justify the following definition, let $\gamma > 0$ be a real number, chosen later, and

$$Y = \{(a, a') \in A_2 \times A_2 : |R_a \cap R_{a'}| \leq \gamma N\} \tag{11.5}$$

be the set of "unfriendly" pairs with poorly intersecting shadows. We expect that $|Y|$ is small if γ is small enough. On one hand

$$\sum_{(a,a') \in Y} |R_a \cap R_{a'}| \leq \gamma N |Y| \leq \gamma N^3, \tag{11.6}$$

on the other hand (note that $b \in R_a \Leftrightarrow a \in R_b \Leftrightarrow (a, b) \in G_2$) we have

$$\sum_{(a,a') \in Y} |R_a \cap R_{a'}| = \sum_{(a,a') \in Y} \sum_{b \in R_a \cap R_{a'}} 1 = \sum_{b \in B} |(R_b \times R_b) \cap Y|. \tag{11.7}$$

(11.6) and (11.7) together say that on average there are not too many unfriendly pairs in shadows. A good candidate for A_3 is a well chosen R_b, but we have to provide several things simultaneously.

We have from (11.4) that

$$\sum_{b \in B_2} |R_b| = |G_2| \geq \frac{\alpha}{2} N^2.$$

Writing $B^* = \{b \in B : |R_b| \geq \frac{\alpha}{4} N\}$, we get from our standard averaging argument (11.2) that

$$\sum_{b \in B^*} |R_b| \geq \frac{\alpha}{4} N^2. \tag{11.8}$$

Now suppose that for every $b \in B^*$ we have

$$|(R_b \times R_b) \cap Y| > \frac{4\gamma}{\alpha} N|R_b|. \tag{11.9}$$

From (11.6), (11.7), (11.8) and (11.9) we get a contradiction, indeed

$$\gamma N^3 \le N \sum_{b \in B^*} \frac{4\gamma}{\alpha} |R_b| < \sum_{b \in B^*} |(R_b \times R_b) \cap Y| \le \gamma N^3.$$

At least once the contrary of (11.9) should be true, there is a $b \in B$ such that choosing $A_3 = R_b$, $B_3 = B_2 = B$ we have

$$|(A_3 \times A_3) \cap Y| \le \frac{4\gamma}{\alpha} N|A_3|, \quad |A_3| \ge \frac{\alpha}{4} N, \quad |G_3| \ge \frac{\alpha}{2} N|A_3|. \tag{11.10}$$

The last estimate is a consequence of the fact that all R_a is big in A_3 by (11.4). For simplifying notation we write $|A_3| = \rho N$, so (11.10) is actually

$$|(A_3 \times A_3) \cap Y| \le \frac{4\gamma\rho}{\alpha} N^2, \quad |A_3| = \rho N, \quad |G_3| \ge \frac{\alpha\rho}{2} N^2, \quad \frac{\alpha}{4} \le \rho \le 1. \tag{11.11}$$

Note that the (later) choice of γ cannot depend on ρ as the choice of A_3 depends on γ.

A_3 is still not perfect, some elements can have many unfriendly counterparts. Let $\delta > 0$ be chosen soon, and let $A_4 = \{a \in A_3 : \#\{a' \in A_3 : (a, a') \in Y\} \le \delta N\}$ be the set of friendly elements, elements with only a few unfriendly counterparts. Trivially

$$\frac{4\gamma\rho}{\alpha} N^2 \ge |(A_3 \times A_3) \cap Y| = \sum_{a \in A_3} \sum_{\substack{a' \in A_3 \\ (a,a') \in Y}} 1 \ge \sum_{a \notin A_4} \sum_{\substack{a' \in A_3 \\ (a,a') \in Y}} 1 \ge \delta N |A_3 \setminus A_4|,$$

$$|A_4| = |A_3| - |A_3 \setminus A_4| \ge \rho N - \frac{4\gamma\rho}{\alpha\delta} N = \frac{\rho}{2} N, \quad \text{if} \quad \delta = \frac{8\gamma}{\alpha}.$$

Still keeping $B_4 = B_3 = B$, for all $a \in A_4$ we have from (11.4) that

$$\#\{a' \in A_4 : (a, a') \in Y\} \le \frac{8\gamma}{\alpha} N, \quad |A_4| \ge \frac{\rho}{2} N, \quad |G_4| \ge \frac{\alpha\rho}{4} N^2. \tag{11.12}$$

There is no more reduction, we choose $A' = A_4$ and $B' = \{b \in B : |R_b| \ge \frac{\alpha\rho}{8} N\}$. Note that $|R_b| \le |A_4| \le |A_3| = \rho N$ and that

$$\sum_{b \in B} |R_b| = |G_4| \ge \frac{\alpha\rho}{4} N^2,$$

so our averaging argument (11.2) (this time with $M = \rho N$) implies that

$$\sum_{b \in B'} |R_b| \ge \frac{\alpha\rho}{8} N^2, \quad |B'| \ge \frac{\alpha}{8} N. \tag{13}$$

The first is a bound for the number of edges connecting A' and B', they are all edges of the original G, so we have by the last inequality of (11.11) that

$$|(A' \times B') \cap G| \geq \frac{\alpha\rho}{8} N^2 \geq \frac{\alpha^2}{32} N^2,$$

which is (11.1.iii.). Also (11.12) and (11.13) prove (11.1.ii.). Now we can complete the plan as is described in the beginning of the proof. Note that we are working in G_4 so R_a and R_b refer to this stage.

Fix any pair of elements $a \in A'$ and $b \in B'$. By the choice of B', the shadow of b is big, $|R_b| \geq \frac{\alpha\rho}{8} N$, while the number of unfriendly counterparts of a is small, $\#\{a' \in A_3 : (a, a') \in Y\} \leq \frac{8\gamma}{\alpha} N$. This is less than half of the shadow if

$$\frac{8\gamma}{\alpha} \leq \frac{1}{2}\frac{\alpha\rho}{8} \quad \Leftrightarrow \quad \gamma \leq \frac{\alpha^2\rho}{128} \quad \Leftarrow \quad \gamma = \frac{\alpha^3}{512}.$$

For all of those friendly a' we have a choice of $|R_a \cap R_{a'}| \geq \gamma N = \frac{\alpha^3}{512} N$ different $b' \in B$. Collecting these bounds we get that the number of different edge–triples (a, b'), (a', b'), and (a', b) in G_4 is at least

$$\#\{a'\}\#\{b'\} \geq \frac{\alpha\rho}{16} N\gamma N \geq \frac{\alpha^5}{2^{15}} N^2. \tag{11.14}$$

Thus for every pair a, b the number of representations

$$a + b = (a + b') - (a' + b') + (a' + b) = s_1 - s_2 + s_3, \tag{11.15}$$

with $s_1, s_2, s_3 \in S$ is by (11.14) at least

$$\#\{a'\}\#\{b'\} \geq \frac{\alpha^5}{2^{15}} N^2.$$

They can be the same for different pairs a, b, but they are certainly different for different values of $a + b$. Since there are no more than $|S|^3 \leq N^3$ different three–tuples of s_1, s_2, s_3, we arrive at (11.1.i.), that is

$$|A' + B'| \leq \frac{\#\text{ of three–tuples in } S}{\#\text{ of representations of } a + b \text{ in (11.15)}} \leq \frac{N^3}{\frac{\alpha^5}{2^{15}} N^2} = \frac{2^{15}}{\alpha^5} N.$$

□

These type of results play a crutial role in the recent argument of Gowers to prove a quantitative Szemerédi's theorem. We illustrate some of the diverse applications in the next chapter. As we have already mentioned, there are many more variations of Theorem 11.1. They basically differ by the type of conclusions. One can ask if partial information about sums imply full information about differences in large subsets, (in contrast to Theorem 11.1, where the full information is derived from sums to sums). In a wide selection of

applications $A = B$, and we want to conclude an $A' \subset A$ with $A' + A'$ is small, (in contrast to Theorem 11.1, where we conclude two, possibly different sets $A', B' \subset A$ with $A' + B'$ is small). Little more generally one can ask information from $A + B$ to $A' \pm A'$. Plünnecke's inequality makes possible to connect the sizes of $A' \pm B', A' \pm A'$, however, an application blows the exponents. It is sometimes more economical to modify the above proof to the new situations. It is a recommended exercise to improve the quantitative aspects of the following corollaries, by adopting the proof of Theorem 11.1. Doing so one can run into an interesting parity problem. For example, $a_1 - a_2 = (a_1 + b) - (a + b) + (a + b') - (a_2 + b')$, however, $a_1 + a_2$ cannot be expressed as a combination of sums from $A + B$. Some estimates are considerable more difficult than others (similarly to the case of Theorem 1.4 and Theorem 1.10).

Corollary 11.4: Let Z be an additive abelian group, $A \subset Z$, $B \subset Z$, $G \subset A \times B$, and $S = \{a + b : (a, b) \in G\}$. If $|A|, |B|, |S| \leq N$ and $|G| = \alpha N^2$ then there is an $A' \subset A$ such that

$$\text{i.) } |A' \pm A'| \leq \frac{2^{33}}{\alpha^{11}} N, \quad \text{ii.) } |A'| \geq \frac{\alpha}{8} N.$$

Proof: An application of Theorem 11.1 provides an $A' \subset A$ and $B' \subset B$ such that

$$\text{i.) } |A' + B'| \leq \frac{2^{15}}{\alpha^5} N, \quad \text{ii.) } |A'|, |B'| \geq \frac{\alpha}{8} N.$$

Either of Corollary 1.5 and Theorem 9.6 than imply that

$$|A' \pm A'| \leq \frac{|A' + B'|^2}{|B'|} \leq \big(\frac{2^{15}}{\alpha^5} N\big)^2 \big(\frac{\alpha}{8} N\big)^{-1} = \frac{2^{33}}{\alpha^{11}} N.$$

□

Corollary 11.5: Let Z be an additive abelian group, $A \subset Z$, $B \subset Z$, $G \subset A \times B$, and $S = \{a + b : (a, b) \in G\}$. If $|A|, |B|, |S| \leq N$ and $|G| = \alpha N^2$ then there is an $A' \subset A$ and a $B' \subset B$ such that

$$\text{i.) } |A' - B'| \leq \frac{2^{51}}{\alpha^{17}} N, \quad \text{ii.) } |A'|, |B'| \geq \frac{\alpha}{8} N, \quad \text{iii.) } |(A' \times B') \cap G| \geq \frac{\alpha^2}{32} N^2.$$

Proof: An application of Theorem 11.1 provides an $A' \subset A$ and $B' \subset B$ such that

$$\text{i.) } |A' + B'| \leq \frac{2^{15}}{\alpha^5} N, \quad \text{ii.) } |A'|, |B'| \geq \frac{\alpha}{8} N, \quad \text{iii.) } |(A' \times B') \cap G| \geq \frac{\alpha^2}{32} N^2.$$

As in the proof of Corollary 11.4, we also have

$$|A' + A'| \leq \frac{2^{33}}{\alpha^{11}} N.$$

Writing $A = A'$, $B = -A'$, and $C = B'$ in Theorem 1.4 we conclude that

$$|A' - B'| \leq \frac{|A' + A'||A' + B'|}{|A'|} \leq \big(\frac{2^{33}}{\alpha^{11}} N\big)\big(\frac{2^{15}}{\alpha^5} N\big)\big(\frac{\alpha}{8} N\big)^{-1} = \frac{2^{51}}{\alpha^{17}} N.$$

□

CHAPTER 12

SOME APPLICATIONS

Erdős conjectured that a set of integers that contains many three–term arithmetic progressions must contain a long arithmetic progression as well. In the first half of this section, we apply Corollary 11.4 (the statistical result), Theorem 10.1 (Freiman's theorem), and Theorem 4.2 (Szemerédi's theorem) to prove this statement. Note that Theorem 4.2 is stated without proof in Chapter 4 of this lecture notes. A full proof is beyond the scope of this course.

Theorem 12.1: Let $\alpha > 0$ and $k \geq 3$. If A is a set of $N \geq N_0(\alpha, k)$ integers that contains at least αN^2 distinct three–term arithmetic progressions, then A contains a k–term arithmetic progression.

Proof: Let A be a set of integers with $|A| = N$. We define a set $G \subset A \times A$. For each non–trivial three–term arithmetic progression a, b, c in A, we put the pair (a, c) into G. We have $|G| \geq \alpha N^2$. Let $S = \{a + c : (a, c) \in G\}$. A three–term arithmetic progression a, b, c satisfies $a + c = 2b$, that is $a + c \in 2 \cdot A$, hence $|S| \leq |2 \cdot A| = N$. Corollary 11.4 (with $A = B$) yields to an $A' \subset A$ with

$$|A'| \geq \frac{\alpha}{8} N, \quad |A' + A'| \leq \frac{2^{33}}{\alpha^{11}} N \leq \frac{2^{36}}{\alpha^{12}} |A'|.$$

Next Theorem 10.1 yields to a generalized arithmetic progression Q such that

$$A' \subset Q = \{a_0 + a_1 n_1 + \cdots + a_d n_d : 0 \leq n_j < N_j \text{ for } j = 1, \ldots, d\},$$
$$N_1 \cdots N_d = \mathrm{vol}(Q) \leq F(C)|A'|, \quad d = \dim Q \leq d(C),$$

where $C = 2^{36}/\alpha^{12}$. We can assume without loss of generality that

$$\max(N_1, \ldots, N_d) = N_d, \quad N_d \geq |A'|^{1/d},$$

since $A' \subset Q$ implies $|A'| \leq N_1 \cdots N_d \leq N_d^d$. Let $\mathcal{V}$ be the set of lattice points

$$\mathbf{v} = (n_1, \ldots, n_{d-1}) \in \mathbb{Z}^{d-1}, \quad 0 \leq n_j < N_j \text{ for } j = 1, \ldots, d-1.$$

Obviously $|\mathcal{V}| = N_1 \cdots N_{d-1}$. For each $\mathbf{v} \in \mathcal{V}$, the set

$$\mathcal{L}(\mathbf{v}) = \{a_0 + a_1 n_1 + \ldots a_{d-1} n_{d-1} + a_d n_d : 0 \leq n_d < N_d\}$$

is an arithmetic progression of length N_d, and $Q = \cup_{\mathbf{v} \in \mathcal{V}} \mathcal{L}(\mathbf{v})$. Thus $A' \subset Q$ means $A' = \cup_{\mathbf{v} \in \mathcal{V}} (\mathcal{L}(\mathbf{v}) \cap A')$, and so

$$\frac{1}{F(C)} N_1 \cdots N_d \leq |A'| \leq \sum_{\mathbf{v} \in \mathcal{V}} |\mathcal{L}(\mathbf{v}) \cap A'|.$$

There is at least one $\mathbf{v} \in \mathcal{V}$ with

$$|\mathcal{L}(\mathbf{v}) \cap A'| \geq \frac{1}{F(C)} N_d.$$

Let us fix one such $\mathbf{v}$, and define

$$B = \{0 < n \leq N_d : a(n) =: a_0 + a_1 n_1 + \dots a_{d-1} n_{d-1} + a_d(n-1) \in A'\}.$$

Since $|B| \geq N_d/F(C)$, by Szemerédi's theorem B contains an arithmetic progression of length k whenever $N_d \geq N(k, \delta)$ with $\delta = 1/F(C)$, and this follows from

$$\frac{\alpha}{8} N \geq N(k, \delta)^d,$$

where the parameters C, δ, d are all functions of α in Theorem 12.1. If $b_1, \dots, b_k$ is a non–trivial arithmetic progression in B then $a(b_1), \dots, a(b_k)$ is a non–trivial arithmetic progression in $A' \subset A$, and we are done.

□

Next we apply Corollary 11.5 to the Kakeya problem. This problem can be studied both in the real Euclidean space or in finite fields; for simplicity we treat with the finite field case only.

Let p be an odd prime, and $\mathbb{Z}_p$ be the finite field of p elements, the field of residue classes mod p. For a fixed integer $n \geq 2$ we consider $\mathbb{Z}_p^n$, the n– dimensional vector space over $\mathbb{Z}_p$. A *line* in $\mathbb{Z}_p^n$ is any set $\mathcal{L}$ of the form

$$\mathcal{L} = \{\mathbf{x} + t\mathbf{v} : t \in \mathbb{Z}_p\},$$

where $\mathbf{x} \in \mathbb{Z}_p^n$ and $\mathbf{v} \in \mathbb{Z}_p^n \setminus \{0\}$; we refer to $\mathbf{v}$ as the *direction* of $\mathcal{L}$, though this direction is only defined up to scalar multiplication by non–zero residue classes. There are $(p^n - 1)/(p-1) > p^{n-1}$ directions. A set $E \subset \mathbb{Z}_p^n$ which contains a line in every direction is called a *Besicovitch set*. (The Euclidean Besicovitch set contains a unit interval in every direction. There are Besicovitch sets of zero measure, and the Kakeya conjecture claims that any Besicovitch set has *Hausdorff dimension* n; it can have zero measure, but it is big.)

The corresponding finite field Kakeya conjecture asserted that

$$|E| \geq c(n, \epsilon) p^{n-\epsilon},$$

for every $\epsilon > 0$ and some $c(n, \epsilon) > 0$. In other words, Besicovitch sets are conjectured to fill up practically all of $\mathbb{Z}_p^n$. The conjecture is recently proved in the stronger form $|E| \gg p^n$, using the polynomial method introduced in Chapter 3. We give the proof later in this chapter. Unfortunately this proof cannot be transferred to the Euclidean case. We start with weaker results that can be transferred (although we skip the transference).

First we prove an easy estimate, which motivates the deeper developments. Fix a Besicovich set E. For every $u \in \mathbb{Z}_p$, consider the *horizontal plane* $\mathbb{Z}_p^{n-1} \times \{u\}$, and let E_u denote the *horizontal slice*

$$E_u =: E \cap (\mathbb{Z}_p^{n-1} \times \{u\}).$$

We obviously have

$$\sum_{u\in\mathbb{Z}_p} |E_u| = |E|,$$

so the average size of $|E_u|$ is $|E|/p$. There must be at least two values of u for which $|E_u| \le 2|E|/p$. By an affine transformation, we may assume those values are 0 and 1, that is $|E_0|, |E_1| \le 2|E|/p$. Now consider the lines in E. For every $\mathbf{w} \in \mathbb{Z}_p^{n-1}$, there must be a line, $\mathcal{L}_{\mathbf{w}}$ with direction $(\mathbf{w}, 1)$ that lies in E. This line intersects E_0 at some point $a(\mathbf{w}) \in E_0$, and intersects E_1 at another point $b(\mathbf{w}) \in E_1$. Clearly $b(\mathbf{w}) - a(\mathbf{w}) = t(\mathbf{w}, 1) = (\mathbf{w}, 1)$ by comparing the last coordinates. Thus to each $\mathbf{w} \in \mathbb{Z}_p^{n-1}$ we can associate a pair $(a(\mathbf{w}), b(\mathbf{w})) \in E_0 \times E_1$ of points, so that the line across them lies in E and has direction $(\mathbf{w}, 1)$. Since two different points determine exactly one line, all these pairs are distinct, so their number is p^{n-1}. On the other hand, their number is bounded by

$$p^{n-1} \le |E_0 \times E_1| = |E_0||E_1| \le \frac{4|E|^2}{p^2}.$$

This yield to the inequality

$$|E| \ge \frac{1}{2}\, p^{\frac{n+1}{2}}.$$

Anything better than this bound is non–trivial. We are going to use Corollary 11.5 to improve upon this (for large dimensions n). Note that the exponents in Corollary 11.5 are not the best obtainable, the result below can be improved with the same method, and can be further improved with a different method.

Theorem 12.2: Let p be an odd prime and $n \ge 2$ be an integer. Every Besicovitch set E in $\mathbb{Z}_p^n$ has size

$$|E| \ge \frac{1}{25}\, p^{\frac{19n+18}{37}}.$$

Proof: We fix a Besicovitch set E, and again, we consider the slices E_u. We have

$$\frac{8|E|}{p}\, |\{u \in \mathbb{Z}_p : |E_u| > \frac{8|E|}{p}\}| < \sum_{u\in\mathbb{Z}_p} |E_u| = |E|,$$

and so

$$|\{u \in \mathbb{Z}_p : |E_u| \le \frac{8|E|}{p}\}| > \frac{7}{8}\, p.$$

In particular, we can find a non–trivial arithmetic progression u, $u + r$, $u + 2r$ such that $|E_u|, |E_{u+r}|, |E_{u+2r}| \le 8|E|/p$. By an affine transformation we may suppose that $u = 0, r = 1/2$, thus we have $|E_0|, |E_{1/2}|, |E_1| \le N =: 8|E|/p$.

As before, we consider the lines in E. For every $\mathbf{w} \in \mathbb{Z}_p^{n-1}$, there must be a line, $\mathcal{L}_{\mathbf{w}}$ with direction $(\mathbf{w}, 1)$ that lies in E. This line intersects E_0 at some point $a(\mathbf{w}) \in E_0$, and intersects E_1 at another point $b(\mathbf{w}) \in E_1$. Clearly $b(\mathbf{w}) - a(\mathbf{w}) = t(\mathbf{w}, 1) = (\mathbf{w}, 1)$ by comparing the last coordinates, while the midpoint $(a(\mathbf{w}) + b(\mathbf{w}))/2$ lies in $E_{1/2}$. Thus to each $\mathbf{w} \in \mathbb{Z}_p^{n-1}$ we can associate a pair $(a(\mathbf{w}), b(\mathbf{w})) \in E_0 \times E_1$ of points, so

that the line across them lies in E and has direction $(\mathbf{w}, 1)$. Since two different points determine exactly one line, all these pairs are distinct, so their number is p^{n-1}. We set $G = \{(a(\mathbf{w}), b(\mathbf{w})); \mathbf{w} \in \mathbb{Z}_p^{n-1}\}$, and in fact, since all differences $b - a$ for $(a, b) \in G$ are different, we have

$$|\{b - a : (a, b) \in G\}| = p^{n-1}, \text{ while } |\{a + b : (a, b) \in G\}| = |E_{1/2}| \leq N.$$

Write $\alpha =: p^{n-1}/N^2 = p^{n+1}/(64|E|^2)$, so that $|G| = \alpha N^2$. We can apply Corollary 12.5, and we get $A' \subset E_0$ and $B' \subset E_1$ such that

$$|A' - B'| \leq \frac{2^{51}}{\alpha^{17}} N.$$

However, since all the differences in G are different, using iii.) from Corollary 11.5 we arrive at

$$\frac{\alpha^2}{32} N^2 \leq |G \cap (A' \times B')| \leq |A' - B'|,$$

and the two inequalities together yield (with some numerical simplification) that

$$\alpha^{19} N \leq 2^{56}, \quad |E| \geq \frac{1}{25} p^{\frac{19n+18}{37}}.$$

□

Theorem 12.3: Let p be an odd prime and $n \geq 2$ be an integer. Every Besicovitch set E in $\mathbb{Z}_p^n$ has size

$$|E| \geq \binom{n+p-1}{n} \geq \frac{p^n}{n!}.$$

Proof: We prove by contradiction. Suppose E is a Besicovitch set satisfying $|E| < \binom{n+p-1}{n}$. There is a non–zero polynomial $F(\mathbf{x}) \in \mathbb{Z}_p[x_1, \ldots, x_n]$ of degree at most $p - 1$ such that $F(\mathbf{x}) = 0$ for all $\mathbf{x} \in E$. This follows from solving a system of linear equations, where the unknowns are the coefficients of $F(\mathbf{x})$, and there is an equation for each points $\mathbf{x} \in E$. The number of unknowns is exactly the number of monomials $x_1^{k_1} \cdots x_n^{k_n}$ with $0 \leq k_1 + \cdots + k_n \leq p-1$, and this is exactly $\binom{n+p-1}{n}$, bigger than the number of equations.

Write $F(\mathbf{x}) = \sum_{i=0}^{p-1} F_i(\mathbf{x})$, where F_i is the homogeneous part of degree i of F. Fix an arbitrary direction $\mathbf{v} \in \mathbb{Z}_p^n$, there is a point $\mathbf{b}$ such that the line $\{\mathbf{b} + t\mathbf{v}\} \in E$, which means $F(\mathbf{b} + t\mathbf{v}) = 0$ for all $t \in \mathbb{Z}_p$. For a fixed $\mathbf{b}$ and $\mathbf{v}$ the left hand side is a polynomial in t of degree at most $p-1$ with p zeros, it must be the zero polynomial. All its coefficients are zero, in particular the coefficient of the highest degree term t^{p-1} is zero. It is easy to see that this coefficient is $F_{p-1}(\mathbf{v})$. Since $\mathbf{v}$ was arbitrary, $F_{p-1}(\mathbf{v}) = 0$ for all $\mathbf{v}$, it is the zero polynomial by Lemma 3.7. In other words $F(\mathbf{x})$ is of degree $p - 2$ at most, and we can repeat the above argument. Again, for an arbitrary direction $\mathbf{v} \in \mathbb{Z}_p^n$, there is a point $\mathbf{b}$ such that $F(\mathbf{b} + t\mathbf{v}) = 0$ for all $t \in \mathbb{Z}_p$. The coefficient of highest degree term t^{p-2} is zero, which is $F_{p-2}(\mathbf{v}) = 0$ for all $\mathbf{v}$. This must be the zero polynomial, $F(\mathbf{x})$ is of degree $p - 3$ at most, and so on, $F(\mathbf{x})$ itself is the zero polynomial. This is a contradiction.

□