Topics in Combinatorial Number Theory

Dissertation submitted to The Hungarian Academy of Sciences
for the degree Doctor of the HAS

Norbert Hegyvári

Eötvös Loránd University
Budapest

2017

Preface

In the present work we will discuss different issues from *Combinatorial Number Theory*. Some decade ago people called it "Erdős type" number theory. Recently the new name of combinatorial number theory is additive combinatorics. It is not too easy to distinguish combinatorial number theory from classical number theory, elementary number theory e.t.c. Trying to approach this question by looking at the tools that are used will not be very useful to answer the above.

As Ben Green wrote *"Well one might say that additive combinatorics is a marriage of number theory, harmonic analysis, combinatorics, and ideas from ergodic theory, which aims to understand very simple systems: the operations of addition and multiplication and how they interact."*

My dissertation contains five chapters from number theory in the topics mentioned above. Indeed; I tried to treat problems in combinatorial way, using probability, Fourier analysis and extremal set theory.

Acknowledgement: There are lots of people I should express my gratitude to. Instead of making a long list, do let me mention how lucky I feel to have had a chance to work with Paul Erdős, to whom I was introduced by Robert Freud. My work was also influenced by Imre Ruzsa and András Sárközy.

I would like to thank to my colleague, Francois Hennecart for many fruitful discussions, and the Jean Monnet University (in Saint Etienne, Lyon) for their recurring invitations and the peaceful environment to work.

I am grateful for Gergely Wintsche, and Tamás Héger for the technical help.

Last but not least, I would like to thank to my big family who – with their mere existence – encouraged me .

Notations

• Let G be any semigroup, $A, B \subseteq G$, and let

$$A + B = \{a + b : a \in A;\ b \in B\} \quad \text{similarly} \quad A \cdot B = \{a \cdot b : a \in A;\ b \in B\}.$$

• The counting function of $A \subseteq \mathbb{N}$ is

$$A(n) := \sum_{a \in A;\ a \leq n} 1$$

.

• We use $\mathbb{N}, \mathbb{N}^+, \mathbb{Z}, \mathbb{R}, \mathbb{R}^+, \mathbb{C}$ in the usual meaning.

• $[1, N] := \{1, 2, \ldots, N\}$

• We shall write $\mathcal{A} \sim \mathbb{N}$ to denote that a set of integers $\mathcal{A}$ contains all but finitely many positive integers

• For $A \subseteq \mathbb{N}$ let us define the lower density of A by

$$\underline{d}(A) = \liminf_{n \to \infty} \frac{|A \cap [1, n]|}{n},$$

the upper density by

$$\overline{d}(A) = \limsup_{n \to \infty} \frac{|A \cap [1, n]|}{n},$$

and the density by

$$d(A) = \lim_{n \to \infty} \frac{|A \cap [1, n]|}{n}$$

if the limit exists.

• Let p be prime number. Denote by $(\mathbb{F}_p, +, \cdot)$ – or briefly $\mathbb{F}_p$ – the p element primefield, $(\mathbb{F}_p^*, \cdot)$ – or shortly $\mathbb{F}_p^*$ – its multiplactive subgroup (sometimes just the set $\{1, 2, \ldots, p - 1\}$).

• Let $e_N(z) = e^{\frac{2\pi i z}{N}}$, and sometimes we leave the subscript.

• We will use the notation $|X| \ll |Y|$ (or $|X| = O(|Y|)$ to denote the estimate $|X| \leq C|Y|$ for some absolute constant $C > 0$. In some occasion we indicate that this constant C depends on a fix parameter K by subscript $|X| \ll_K |Y|$.

• $f \asymp g$, if $f \ll g$ and $g \ll f$.

• Let $X \subseteq \mathbb{F}_p^*$. $\langle X \rangle$ denotes the group generated by X, i.e $\langle X \rangle < \mathbb{F}_p^*$.

• Given a real number x we denote by $\langle x \rangle$ the fractional part of x. That is, $\langle x \rangle = x - \lfloor x \rfloor$.

• Given a subset A of $\mathbb{R}$, we write $\mu(A)$ for the outer Lebesgue measure of A.

• Let $x_0, a_1 < a_2 < \cdots < a_d$ be any sequence of integers. The Hilbert cube is the set

$$H(x_0, a_1, a_2, \ldots, a_d) = \Big\{ x_0 + \sum_{1 \leq i \leq d} \varepsilon_i a_i \Big\} \quad \varepsilon_i \in \{0, 1\}.$$

We can define a Hilbert cube of order $r \geq 1$; $r \in \mathbb{N}$ extending the previous definition by

$$H_r(x_0, a_1, a_2, \ldots, a_d) = \Big\{ x_0 + \sum_{1 \leq i \leq d} \varepsilon_i a_i \Big\} \quad \varepsilon_i \in \{0, 1, \ldots, r\}.$$

When $r = 1$, we write shortly $H(x_0, a_1, a_2, \ldots, a_d) = H_1(x_0, a_1, a_2, \ldots, a_d)$.

We say that $\dim(H) := d$ is the dimension of H and $|H(x_0, a_1, a_2, \ldots, a_d)|$ is its size.

Let $\Delta, 0 < \Delta \leq 1$ be a real parameter. We say that a cube $H =: H_r(x_0, a_1, a_2, \ldots, a_d)$ is $\Delta-$degenerate, if $\frac{\log_{r+1} |H|}{d} = \Delta$.

$\log_{r+1} x$ means $\log x / \log(r + 1)$.

When $\Delta = 1$, then $|H| = (r + 1)^d$. In this case all terms of the cube are pairwise distinct and H is said to be non-degenerate.

• For a sequence of functions $f_1, f_2, \ldots, f_n$ and a real number $p \geq 1$, the p-norm is the mean

$$\Big(\sum_{i=1}^{n} |f_i|^p \Big)^{\frac{1}{p}}.$$

- For an arbitrary set $A \subseteq G$ its *additive* energy is defined by

$$E_+(A) := \{(a_1, a_2, a_3, a_4) \in A^4 : a_1 + a_2 = a_3 + a_4\}$$

and its *multiplicative* energy is defined by

$$E_\times(A) := \{(a_1, a_2, a_3, a_4) \in A^4 : a_1 \cdot a_2 = a_3 \cdot a_4\}.$$

- Let f be an arbitrary function from $\mathbb{F}_p^*$ to $\mathbb{C}$. Denote the Fourier transform (with respect to a multiplicative character) by

$$\widetilde{f(u)} := \sum_{x \in \mathbb{F}_p^*} f(x)\chi_u(x)$$

where $\chi_u(x)$ is the multiplicative (Dirichlet) character; $\chi_u(x) = e^{\frac{2\pi i ind x \cdot u}{p-1}}$ where *indx* is index of x (or it is sometimes said to be discrete logarithm). When $\chi \neq \chi_0$ is not the principal character, then let $\chi(0) = 0$.

- Recall (what we will use many times) that

$$\sum_{u \in \mathbb{F}_p^*} |\widetilde{f(u)}|^2 = (p-1) \sum_{x \in \mathbb{F}_p^*} |f(x)|^2$$

Let $g : \mathbb{F}_p \to \mathbb{C}$ and $x \in \mathbb{F}_p$. Denote the Fourier transform (with respect to an additive character) by

$$\widehat{g}(x) := \sum_{y \in \mathbb{F}_p} g(y) e_p(yx)$$

where $e_p(t) := \exp(2i\pi t/p)$.

Contents

Chapter 1

Introduction

In the present work I selected some of my results from 1993 (the year when I received my CSc) and there is a common feature of these works; I do not mean that the treatment of the problems are similar (I use combinatorial ideas, probabilistic-counting methods, Fourier analysis e.t.c) rather the topic.

The similarity is to show structures in various objects.

I devote CHAPTER 2 the investigation of different problems of Hilbert cubes. First I summarize known results from Hilbert to Szemerédi. Many authors worked in this area.

In section 2.1 I discuss some of my results on the dimension of dense and thin sets. The main difficulty lies the fact that we allow here *degenerate cubes* as well. Our approach is non-deterministic. This section based on the papers

N. Hegyvári, On the dimension of the Hilbert cubes. J. Number Theory 77 (1999), no. 2, 326–330.

N. Hegyvári, On Combinatorial Cubes, The Ramanujan Journal, 2004, Volume 8, Issue 3, pp 303307

In section 2.2 we discuss a result of Bergelson on the difference set $A - A$ with $\overline{d}(A) > 0$. The original proof used Fürstenberg Correspondence principle, (an ergodic theorem). We prove a more general (but in some sense weaker) version via combinatorial way and a stronger version (also due to Bergelson) using Følner theorem. We quote papers

N. Hegyvári, Additive Structure of Difference Sets, seminar Advanced

Courses in Mathematics CRM Barcelona, Thematic Seminars Chapter 4 p 253-265

N. Hegyvári, Note on difference sets in $\mathbb{Z}^n$ *Period. Math. Hungar. Vol 44 (2), 2002, pp. 183-185*

N. Hegyvári, I.Z. Ruzsa, Additive Structure of Difference Sets and a Theorem of Følner, Australasian J. of Combinatorics Volume 64(3) (2016), Pages 437-443

Recently many authors investigate character sums on certain structured sets. Let me just mention a recent work of Shparlinski, Petridis, Garaev, Konyagin and Shkredov. In section 2.3 I gave bounds to character sums on Hilbert cubes. The main tool is some estimation on the energy of the cubes; additive energy of multiplicative Hilbert cubes and multiplicative energy of additive Hilbert cubes.

We compare our result to other general bounds of other structures, (example of Montgomery). Results are from

N. Hegyvári, Note on character sums of Hilbert cubes, Journal of Number Theory Volume 160: pp. 526-535. (2016)

Section 2.4 The main part contains a joint work with A. Sárközy. The problem which was raised by Brown, Erdős and Freedman asked what the largest dimension of a Hilbert cube is contained in the first n squares and the first n primes respectively. We gave an improvement of an earlier result of Rivat-Sárközy-Stewart. Some related problems are also discussed. The section based on

N. Hegyvári, A. Sárközy, On Hilbert cubes in certain sets. Ramanujan J. 3 (1999), no. 3, 303–314.

Ramsey types question pops up in many places in additive combinatorics as well (Van der Waerden theorem, result of Schur, Quasi-progressions e.t.c).

In CHAPTER 3 we discuss the additive Ramsey type problems.

In 1968 Raimi proved, using topological tool the following theorem: There exists $E \subseteq \mathbb{N}$ such that, whenever $r \in \mathbb{N}$ and $\mathbb{N} = \bigcup_{i=1}^r D_i$ there exist $i \in \{1, 2, \dots, r\}$ and $k \in \mathbb{N}$ such that $(D_i + k) \cap E$ is infinite and $(D_i + k) \setminus E$ is infinite. In 1979 Hindman gave an elementary proof of Raimi's theorem.

In section 3.1 we give a far reaching generalization of Raimi-Hindman theorem. This result is connected to the previous chapter.

N. Hegyvári, On intersecting properties of partitions of integers, Combin. Probab. Comput. (14) 03, (2005), 319-323

In section 3.2 we give an answer to a problem of Sárközy; coloring the set of squares by two colours, then how many elements need to have a monochromatic representation of every sufficiently large numbers.

N. Hegyvári, F. Hennecart,On Monochromatic sums of squares and primes, Journal of Number Theory, Volume 124, Issue 2, 2007, Pages 314-324

I devote Chapter 4 to the topic *restricted addition*; i.e. sumsets, where the summands are pairwise distinct. We solve and improve problems and results of Erdős, Burr and Davenport.

N. Hegyvári, F. Hennecart and A. Plagne, Answer to the Burr-Erdős question on restricted addition and further results, Combinatorics, Probability and Computing, Volume 16, Issue 05, Sep 2007, pp 747-756,

N. Hegyvári, On the representation of integers as sums of distinct terms from a fixed set Acta Arith. 92.2 2000. 99-104

N. Hegyvári, On the completeness of an exponential type sequence. Acta Math. Hungar. 86 (2000), no. 1-2, 127–135

Chapter 5. Expanding polynomials. This topic is intensively investigated; it has a strong connection to computer science, and in the additive combinatorics to the "sum-product" problem. A polynomial in a prime field is said to be expander, if it blows up its domain. It is not too hard to construct a three-variable expanding polynomial. The first explicit two-variable expander is due to Bourgain. In this chapter we give an infinite class of explicit two-variable expanders. Furthermore we give explicit bounds to the expanding-measure. Further results are also considered.

N. Hegyvári, F. Hennecart, Explicit Constructions of Extractors and Expanders Acta Arith. 140 (2009), 233-249.

N. Hegyvári, Some Remarks on Multilinear Exponential Sums with an Application, Journal of Number Theory Volume 132, Issue 1, January 2012, Pages 94-102

N. Hegyvári, On sum-product bases, Ramanujan J. (2009) 19:p 1-8

Chapter 6. Lately new results pop up on expansion of Lie-type simple groups. Helfgott proved that for $A \subset SL_n(\mathbb{Z}_p)$, $|A \cdot A \cdot A| > |A|^{1+\varepsilon}$ (where $\varepsilon > 0$ is an absolute constant) unless A is contained in a proper subgroup. Or a nice and deep result (called "Convolution bound") of Babai-Nikolov-Pyber, which ensures that if $A \subset SL_2(\mathbb{Z}_p)$, $|A| \sim p^{5/2}$ then $|A^2|$ covers at least one third of the group.

Nevertheless, it is very less known on the structure of (k-fold) product sets in this non-abelian groups. In this chapter we show some structure theorem in Heisenberg groups. The method of my paper (On sum-product bases) is well applicable.

N. Hegyvari and F. Hennecart, A structure result for bricks in Heisenberg groups, Journal of Number Theory 133(9) (2013): 29993006.

In *some section* I include further results as well.

Chapter 2

On Hilbert cubes

In 1892 D. Hilbert published a paper in [Hil] on irreducibility of k-variable polynomials with integral coefficients. His theorem has many nice applications; e.g. if $f(x) \in \mathbb{Z}[x]$ and for $x > x_0$, the values of f are always square number, then $f(x)$ itself a square of some polynomial over $\mathbb{Z}$. A special, 2-variable case can be written as follows (note; the original version may be written differently):

Theorem 2.1 (Hilbert)**.** Let $f(x, y) \in \mathbb{Z}[x, y]$ be irreducible. Then there is an infinite set Y, such that for every $y^* \in Y$ $f(x, y^*)$ is irreducible in $\mathbb{Z}[x]$.

To prove this, Hilbert showed the first Ramsey type theorem (25 years older than the famous "$x + y = z$" problem of I. Schur).

Theorem 2.2 (Hilbert)**.** Let m and r be positive integers. For every $r-$colouring of $\mathbb{N}$ there exists a monochromatic affine cube $H(a_0, x_1, x_2, \dots, x_m)$.

(Of course it is a modern terminology of the theorem).

Hilbert cubes have many applications. The effective version was an important tool in the celebrated Szemerédi's theorem:

Theorem 2.3 (Szemerédi, 1969)**.** *Let $A \subseteq \mathbb{N}$ with $\eta := \underline{d}(A) > 0$. Then there exists a $\beta > 0$ real number such that for $n > n_0(\eta)$ the set $A \cap [1, n]$ contains a Hilbert cube with dimension at least $\beta \log\log n$.*

Definition 2.4. Let A be an infinite increasing sequence of integers. Let

$$H_A(n) = \max\{m : A \cap [1,n] \text{ contains a Hilbert cube } H(a_0, x_1, x_2, \ldots, x_m)\}$$

Recall that a Hilbert cube is non-degenerate if $|H(a_0, x_1, x_2, \ldots, x_m)| = 2^m$ (i.e. there is no coincidence in the "vertices"), otherwise let us call degenerate.

2.1 On the dimension of Hilbert cubes

2.1.1 Hilbert cubes in dense sets

In this section we allow the degenerate cube as well. We prove the following theorem:

Theorem 2.5 (Hegyvári, [H97])**.** There exists an infinite sequence of positive integers with $\underline{d}(A) > 0$ and

$$H(n) \leq c\sqrt{\log n \log\log n}$$

where $c = 4(\log(4/3))^{-1/2}$.

Proof of Theorem 2.5. We start by an easy lemma:

Lemma 2.6. Let $B = \{b_1 < b_2 < \cdots < b_k\}$ be a sequence of integers. Then

$$\binom{k+1}{2} + 1 \leq |FS(B)| \leq 2^k$$

The proof is simple or see [He96].

Lemma 2.7. We have

$$T := |\{A \subseteq [1,n] : |A| = k \text{ and } |FS(A)| < k^3\}| < n^{3\log_2 k} \cdot 3^{k^2}.$$

Proof of Lemma 2.7. Let $U = \{A \subseteq [1,n] : |A| = k \text{ and } |FS(A)| < k^3\}$. Let $R = \lfloor 3\log_2 k \rfloor$ and assume $A = \{a_1 < a_2 < \cdots < a_k\} \in U$.

An element a_j is said to be *doubler* if

$$FS(a_1 < a_2 < \cdots < a_{j-1}) \cap \{a_j + FS(a_1 < a_2 < \cdots < a_{j-1})\} = \emptyset \quad (2.1)$$

Since

$$FS(a_1 < a_2 < \cdots < a_j) = \{0, a_j\} + FS(a_1 < a_2 < \cdots < a_{j-1})$$

thus if a_j is a doubler then

$$|FS(a_1 < a_2 < \cdots < a_j)| = 2|FS(a_1 < a_2 < \cdots < a_{j-1})| \tag{2.2}$$

This yields that

$$|FS(a_1 < a_2 < \cdots < a_k)| \geq 2^H \tag{2.3}$$

where Hdenotes the number of doublers in A.

H is at most R since in the opposite case $2^H \geq 2^{R+1} > 2^{3\log_2 3} = k^3$, which by (2.3) contradicts the fact $A \in U$.

Now if a_j is not a doubler then we must have

$$a_j \in \{x - x' : x, x' \in FS(a_1 < a_2 < \cdots < a_{j-1})\},$$

which easily implies that we can write a_j in the form

$$a_j = \sum_{i \neq j} \delta_i a_i;\ \ \delta_i \in \{1, +1, -1\}, \tag{2.4}$$

which yields that the number of non-doubler elements is at most 3^k.

Now we get an upper estimation for T:

We can select $\binom{k}{R}$ subscripts j where a_j is a doubler, the number of possible values of the doublers being at most n^R. Finally, the number of non-doublers is at most $(3^k)^{k-R}$.

Thus we have

$$T \leq \binom{k}{R} \cdot n^R \cdot (3^k)^{k-R} \leq$$

$$\leq k^R \cdot n^R \cdot 3^{k^2 - kR} \leq n^R \cdot 3^{k^2}$$

using the inequality $k^R < 3^{kR}$.

$\square$

Now we turn to the proof of the Theorem.

Let X be a random sequence of integers with $Pr(x \in X) = \frac{1}{16}$. Clearly with probability 1 we have $\underline{d}(X) > 0$. Let H_n be the event

$$H_X(n) > c\sqrt{\log n \log\log n}$$

where $c = 4(\log(4/3))^{-1/2}$. We are going to show

$$Pr(H_n) < \frac{1}{n^2}. \tag{2.5}$$

We have

$$Pr(H_n) \le \sum_{\substack{1 \le a \le n \\ 1 \le x_1, \dots, x_k \le n}} \left(\frac{1}{16}\right)^{|FS(x_1 < x_2 < \dots < x_k)|} =$$

$$= \sum_{\substack{1 \le a \le n \\ 1 \le x_1, \dots, x_k \le n \\ |FS(x_1 < \dots < x_k)| < k^3}} \left(\frac{1}{16}\right)^{|FS(x_1 < \dots < x_k)|} + \sum_{\substack{1 \le a \le n \\ 1 \le x_1, \dots, x_k \le n \\ |FS(x_1 < \dots < x_k)| \ge k^3}} \left(\frac{1}{16}\right)^{|FS(x_1 < \dots < x_k)|}. \tag{2.6}$$

By Lemma 2.1 and Lemma 2.2 we have

$$\sum_{\substack{1 \le a \le n \\ 1 \le x_1, \dots, x_k \le n \\ |FS(x_1 < \dots < x_k)| < k^3}} \left(\frac{1}{16}\right)^{|FS(x_1 < \dots < x_k)|} \le \sum_{1 \le a \le n} n^{3 \log_2 k} \cdot 3^{k^2} \left(\frac{1}{16}\right)^{k^2/2} =$$

$$= n \cdot n^{3 \log_2 k} \left(\frac{3}{4}\right)^{k^2},$$

which is less then $\frac{1}{2n^2}$ if $k \ge 4(\log(4/3))^{-1/2}\sqrt{\log n \log\log n}$.

Furthermore

$$\sum_{\substack{1 \le a \le n \\ 1 \le x_1, \dots, x_k \le n \\ |FS(x_1 < \dots < x_k)| \ge k^3}} \left(\frac{1}{16}\right)^{|FS(x_1 < \dots < x_k)|} \le n \cdot \binom{n}{k} \left(\frac{1}{16}\right)^{k^3} <$$

$$< \frac{1}{2n^2}$$

holds if $k > 3\sqrt{\log_n}$.

By (2.5) we have

$$\sum_{n=1}^{\infty} Pr(H_n) < \infty,$$

so by the Borel-Cantelli lemma with probability 1, at most a finite number of events H_n occur.

$\square$

Note that we split the sum in (2.6) into two parts according the value $|FS(x_1, \ldots, x_k)|$. We mention here that for the sets $A_d = \{d, 3d, \ldots, kd\}$ $d = 1, 2, \ldots, \lfloor \frac{n}{k} \rfloor$, we have

$$|FS(A_d)| = \binom{k+1}{2} \sim k^2.$$

So we have to count these sets in the first sum which yields that our method works only if $k \gg \sqrt{\log n}$

In the next Proposition we will show that for a *random* sequence our bound, apart from the factor $\sqrt{\log \log n}$ is the best possible.

Proposition 2.8. *Let A be a random sequence of positive integers with $Pr(a \in A) = p > 0$. Then with probability 1, we have*

$$H_A(n) \gg_p \sqrt{\log n}.$$

Proof. Let $0 < p < 1$ be a fixed real number and let A be a random sequence of integers with $Pr(a \in A) = p > 0$ and let $k_n = \max_{a,k} = \{k : a+1, a+2, \ldots, a+k \in A\}$. By a theorem of Erdős and Rényi [ERe], with probability 1, $k_n = c_p \log n$. But let us observe that if $a, a+1, \ldots, a+k \in A$ then $H(a, 1, 2, \ldots \lfloor \sqrt{2k} - 1 \rfloor) \subset A$. Indeed, $H(a, 1, 2, \ldots \lfloor \sqrt{2k} - 1 \rfloor) \supset \{a, a+1, a+k\}$. It yields that with probability 1, we have

$$H_A(n) \gg_p \sqrt{\log n}.$$

□

Remark 2.9. 1. Recently Conlon, Fox and Sudakov [CFS] could move the $\sqrt{\log \log n}$ factor from the upper bound, so apart from a constant factor our result is strict. Their method is also probabilistic.

2. Cs.Sándor in [CSS] obtained a bound for the dimension to non-degenerate random Hilbert cube. His proof is also non-deterministic.

2.1.2 Hilbert cubes in thin sets

The density version of Szemerédi was rediscovered by many authors and proved in a same way (see e.g. [GR]). In fact one can state it in a stronger form:

Theorem 2.10. Let $A \subseteq [1, N]$ with $|A| > N^{4/5}$. Then there exists a Hilbert cube contained in A with dimension

$$\gg \log \frac{\log 3N}{\log(3N/|A|)}$$

In the present section we are going to investigate a similar question in thin sets as in the previous section.

Let $r_3(n)$ be the maximal number of integers that can be selected from the interval $[1, n]$ without including a three-term arithmetic progression.

Theorem 2.11 (Hegyvári [He04])**.** There exists a subset A of $[1, n]$ for which $|A| \geq \frac{1}{3} r_3(n)$ and

$$\max_{H \subset A \cap [1,n]} \dim H \leq \frac{1}{\log 2} \log \log n. \tag{2.7}$$

Corollary 2.12. For every c, $1/2 < c < 1$ there exists a sequence $A \subset [1, n]$ with

$$|A| = n \cdot e^{-(\log n)^c} \tag{2.8}$$

and

$$\frac{11}{10}(1 - c)(1 + o(1)) \log \log n \leq \max_{H \subset A \cap [1,n]} \dim H \leq \frac{1}{\log 2} \log \log n. \tag{2.9}$$

Proof. Let A be a maximal subset of $[1, n]$ which contains no three-term arithmetic progression. Hence $|A| = r_3(n)$. It is proved by Behrend that there is a set $A_{-1} \subset [1, n]$ which contains no three-term arithmetic progression and $|A_{-1}| > ne^{\alpha\sqrt{\log n}}$. So let $A_{-1} \supseteq A_0$, $|A_0| = ne^{\alpha\sqrt{\log n}}$. Now take a random 2-coloring of the elements of A_0 obtained by coloring each element independently either blue or red, where each color is equally likely. Fix a set $\{a, x_1, \ldots, x_k\}$ for which $H = H(a, x_1, \ldots, x_k) \subseteq A_0$ and H is non-degenerate (i.e. the vertices of the cube are distinct). Let X_H be the event that H is monochromatic.

The cube H is non-degenerate thus we have $Pr(X_H) = 2^{1-2^k}$. Furthermore there are $\binom{|A_0|}{k+1}$ possible choice for $a, x_1, \dots, x_k$ thus we conclude

$$Pr(S) \le \binom{|A_0|}{k+1} 2^{1-2^k} < |A_0|^{k+1} 2^{1-2^k}, \tag{2.10}$$

where $S = \{\text{For any } H \subseteq A_0,\ H \text{ is non-degenerate and monochromatic}\}$. An easy calculation shows that $Pr(S) < \frac{1}{2}$, provided

$$k \ge \frac{(1+o(1))}{\log 2} \log\log |A_0| = \frac{(1+o(1))}{\log 2} \log\log n$$

if n is large enough. It implies that with probability at least $\frac{1}{2}$ a random subset of A_0 does not contain a non-degenerate cube H with

$$\dim H > \frac{(1+o(1))}{\log 2} \log\log n. \tag{2.11}$$

Furthermore the number of occurrences of a given color has binomial distribution with expectation $|A_0|/2$ and standard deviation $\sqrt{|A_0|}/2$ thus by Chebyshevs inequality, for a random subset A of A_0 we have

$$Pr(|A| > |A_0|/3) > 1/2, \tag{2.12}$$

if n is large enough. Now by (2.11) and (2.13) we obtain that there is a subset A of A_0 for which

$$|A| > \frac{|A_0|}{3} \tag{2.13}$$

and if H is a non-degenerate cube of A, then

$$\dim H \le \frac{(1+o(1))}{\log 2} \log\log n. \tag{2.14}$$

Now we shall prove (2.7). Assume now to the contrary our assumption there exists a cube H in A for which

$$\dim H > \frac{(1+\varepsilon)}{\log 2} \log\log n.$$

for some $\varepsilon > 0$. By (2.14) H cannot be non-degenerate. Thus there exists an $x \in H$, for which

$$x = a + \epsilon_1 x_1 + \epsilon_2 x_2 \cdots + \epsilon_k x_k \tag{2.15}$$

and

$$x = a + \epsilon'_1 x_1 + \epsilon'_2 x_2 \cdots + \epsilon'_k x_k \qquad (2.16)$$

where $\epsilon_i, \epsilon'_i \in \{0,1\}$ and $(\epsilon_1, \dots, \epsilon_k) \neq (\epsilon'_1, \dots, \epsilon'_k)$. If there are common vertices (i.e. $\epsilon_i = \epsilon'_i = 1$) in the representation (2.15) and (2.16) then delete them, so we get an $x^* \in H$ which has at least two disjoint representations

$$x^* = a + \delta_1 x_1 + \delta_2 x_2 \cdots + \delta_k x_k \qquad (2.17)$$

and

$$x^* = a + \delta'_1 x_1 + \delta'_2 x_2 \cdots + \delta'_k x_k \qquad (2.18)$$

where $\delta_i, \delta'_i \in \{0,1\}$ and

$$\delta_i \cdot \delta'_i = 0 \qquad (2.19)$$

for $i = 1, 2, \dots, k$. By (2.17), (2.18) and (2.19) we obtain $a \in H$, $x^* \in H$,and $x^* + \delta'_1 x_1 + \delta'_2 x_2 \cdots + \delta'_k x_k \in H$. Here $x^* + \delta'_1 x_1 + \delta'_2 x_2 \cdots + \delta'_k x_k = x^* + x^* - a$, i.e. $\{a, x^*, 2x^* - a\} \subset H \subset A$. But $\{a, x^*, 2x^* - a\}$ forms a three-term arithmetic progression contained in A. This contradiction proves the theorem.

□

2.2 On Bergelson's theorem

The study some properties of $D(A) := A - A$ of dense sets in $\mathbb{Z}$ was a center problem in combinatorial number theory.

Erdős and Sárközy's unpublished result from the 60's states: if the upper density of an $A \subseteq \mathbb{N}$ is positive then $D(A) := A - A$ contains an arbitrarily long arithmetic progression.

On the iterated difference set $D(D(A))$ Bogolyubov obtained the following classical result:

Theorem 2.13 (Bogolyubov)**.** Let $A \subseteq \mathbb{N}$ with $\overline{d}(A) > 0$. Then there is a *Bohr set*

$$B(S, \varepsilon) = \{m \in \mathbb{Z} : \max_{s \in S} \|sm\| < \varepsilon\}$$

($\|x\| = \min_{n \in \mathbb{Z}} |x - n|$, the absolute fractional part) for which

$$D(D(A)) = A - A + A - A \supseteq B(S, \varepsilon).$$

On the other hand Kříž proved

Theorem 2.14 (Kříž)**.** There is a set A with positive upper density whose difference set $D(A)$ contains no Bohr set

So it is very reasonable to ask: What can we say about the structure of $D(A)$ when $\overline{d}(A) > 0$? In 1985 Bergelson proved [Be85] that in this case $D(A)$ is well-structured. Firstly he proved

Theorem 2.15 (Bergelson)**.** Let $A \subseteq \mathbb{N}$ with $\overline{d}(A) > 0$. For every k there exists an infinite set B of integers for which $A - A \supseteq B + B + \cdots + B$, ($k$ times)

His proof of this theorem is based on an ergodic theorem, namely Fürstenberg correspondence theorem (see also [Be85]).

Later Bergelson et al [Be97] gave a more general form of Theorem 2.15 which will be discussed in subsection 2.2.2

2.2.1 A combinatorial proof for Theorem 2.15 under restricted sum

The original proof of Theorem 2.15 is based on a deep ergodic theorem of Fürstenberg which was worked out just for the set of integers. We prove a related theorem in a more general structure, namely in $\mathbb{Z}^n$ (strictly speaking the proof below works not only in $\mathbb{Z}^n$; one can imitate it in more general structure, for instance in $\sigma-$finite (abelian) groups as well); neverheless we can guarantee a k-fold *resticted* sum $B\dot{+}B\dot{+}\dots\dot{+}B$, instead of the k-fold sum $B + B + \cdots + B$ in the difference set $A - A$. Before formally stating our theorem recall some definition.

Define the *discrete* rectangle of $\mathbb{Z}^n$ by

$$R = [a_1, b_1] \times [a_2, b_2] \times \dots [a_n, b_n] \cap \mathbb{Z}^n.$$

The volume of R is $|R| = \prod_i (b_i - a_i + 1)$.

Recall the notion of upper Banach density of A is

$$d^*(A) := \sup\{L : \forall m,\ \exists R_m,\ \min_i |b_i - a_i| \geq m, s.t.\ \frac{|A \cap R_m|}{|R_m|} \geq L\}.$$

We prove the following theorem in a more general set;

Theorem 2.16 (Hegyvári [He08])**.** Let $A \subseteq \mathbb{Z}^n$, with $d^*(A) = \gamma > 0$. For every integer M there is an infinite set $B \subseteq \mathbb{Z}^n$ such that

$$D(A) \supseteq= B \times M := B\dot{+}B\dot{+}\ldots\dot{+}B(M\ times).$$

Proof. Consider the integer lattice points $\{\mathbf{x}_i\}_{i=1}^{M^n}$; $\mathbf{x}_i = (x_{i_1}, x_{i_2}, \ldots, x_{i_n})$; $0 \leq x_{i_j} \leq M-1$. For $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, write

$$\mathbf{u} \equiv \mathbf{v}(modM)$$

if and only if

$$u_i \equiv v_i(modM)$$

for all $1 \leq i \leq n$. Let

$$A_i = \{\mathbf{a} \in A : \mathbf{a} \equiv \mathbf{x}_i(modM)\}.$$

Since $d^*(A) = \gamma > 0$ we have that $d^*(A_i) = \rho > 0$ for some i.

Let

$$A' = A_i - \mathbf{x}_i \subseteq L := \{\mathbf{u} \equiv \mathbf{0}(modM)\}.$$

Obviously

$$A' - A' = A_i - A_i \subseteq A - A.$$

Lemma 2.17. There exists a $finite$ set $U \subset L$ such that

$$A' - A' + U = L.$$

Proof of the Lemma:

Let $U = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r, \ldots\}$ be the maximal subset of $\mathbb{Z}^n$, such that the sets

$$\mathbf{u}_1 + A', \mathbf{u}_2 + A', \ldots, \mathbf{u}_r + A', \ldots$$

are pairwise disjoint.

We claim that $r \leq 4/\rho$.

Indeed since $d^*(A') = d^*(A_i) = \rho > 0$, there is a rectangle R such that $|R \cap A'| \geq \frac{\rho|R|}{2}$. Assume that the minimal length of edge of R is large enough, then we get

$$|R| \geq |R \cap \{(\mathbf{u}_1 + A') \cup \cdots \cup (\mathbf{u}_r + A')\}| =$$
$$|R \cap (\mathbf{u}_1 + A')| + \cdots + |R \cap (\mathbf{u}_r + A')| \geq r\frac{|R \cap A'|}{2} \geq r\frac{\rho|R|}{4}$$

which gives $r \leq 4/\rho$.

Now we prove that $A' - A' + U = L$. Assume to the contrary that there is an $x \in L$ for which

$$x \notin A' - A' + U$$

It means that for all $i = 1, 2, \dots r$

$$x + A' \cap \mathbf{u}_i + A' = \emptyset.$$

But it contradicts to the maximality of U. $\square$

We introduce an $r-$coloring

$$\chi(\mathbf{x}_1, \dots, \mathbf{x}_M) \mapsto \{1, 2, \dots, r\}$$

of all M element subsets of L as follows: for an $M-$tuple $\mathbf{x}_1, \dots, \mathbf{x}_M$ let

$$\chi(\mathbf{x}_1, \dots, \mathbf{x}_M) = \min\{i : \mathbf{x}_1 + \cdots + \mathbf{x}_M \in A' - A' + \mathbf{u}_i\}.$$

(Note the coloring is not necessary unique).

Lemma 2.18 (Ramsey)**.** Let X be a countable set and color all $M-$tuples of X by r colors. Then there exists an $infinite$ set B' which is monochromatic.

Now by this lemma we have that there is an infinite set $B' \subseteq L$ and an s, $1 \leq s \leq r$ for which every $M-$tuple $(\mathbf{x}_1, \dots, \mathbf{x}_M)$ of B'

$$\mathbf{x}_1 + \cdots + \mathbf{x}_M \in A' - A' + \mathbf{u}_s$$

holds.

Finally let

$$B := B' - \frac{\mathbf{u}_s}{M}.$$

Since $\mathbf{u}_s \in L$ we have that $\frac{\mathbf{u}_s}{M} \in \mathbb{Z}^n$. Thus we have

$$A' - A' + \mathbf{u}_s \supseteq B' \dot{+} B' \dot{+} \dots \dot{+} B' (M\ times) = (B + \frac{\mathbf{u}_s}{M}) \dot{+} \dots \dot{+} \dot{+} (B + \frac{\mathbf{u}_s}{M})(M\ times) =$$

$$= B\dot{+}B\dot{+}\ldots\dot{+}B + M\frac{\mathbf{u}_s}{M},$$

which implies

$$A - A \supseteq A' - A' \supseteq B\dot{+}B\dot{+}\ldots\dot{+}B(M\ times).$$

□

2.2.2 A stronger version of Theorem 2.15

In [Be97] the authors showed that whenever $\overline{d}(A) > 0$, $D(A)$ has a rich additive and multiplicative structure. For instance in Theorem 3 p.135. the following result proved

Theorem 2.19 (Bergelson et al)**.** Let B with $\overline{d}(B) > 0$. Then there is some sequence $\{x_n\}_{n=1}^{\infty}$ such that

$$\Big\{\sum_{n\in F} a_n x_n :\ F \text{ is a finite subset of } \mathbb{N} \text{ and for each } n \in F,\ a_n \in \{1,2\}\Big\} \cup$$

$$\cup \Big\{\prod_{n\in F} x_n^{a_n} :\ F \text{ is a finite subset of } \mathbb{N} \text{ and for each } n \in F,\ a_n \in \{1,2\}\Big\} \subseteq D(B).$$

In Theorem 5 they proved that $D(B)$ contains sums and products from a sequence where terms are allowed repeat a restricted number of times.

At the proof they used also a (deep) ergodic theorem. In the theorem below we can avoid this tool; instead of it we will utilize that $D(A)$ conatins almost a Bohr set.

Let $f : \mathbb{N}_+ \to \mathbb{N}_+$ be any function and $C \subseteq \mathbb{N};\ C \neq \emptyset$. We will use the following notations:

$$FS_f(C) := \Big\{\sum_{c_i\in X} w_i c_i : X \subseteq C,\ |X| < \infty;\ w_i \in [1, f(i)] \cap \mathbb{N}\Big\}.$$

Let the sum be zero, when X is the empty set.

Furthermore write

$$FP(C) := \Big\{ \prod_{c_i \in X} c_i : X \subseteq C;\ X \neq \emptyset,\ |X| < \infty \Big\}.$$

Clearly we have

$$FS_f(\{c_1, c_2, \dots c_n\}) = FS_f(\{c_1, c_2, \dots c_{n-1}\}) + \{0, c_n, \dots, f(n)c_n\}, \quad (2.20)$$

and

$$FP(\{c_1, c_2, \dots c_n\}) = FP(\{c_1, c_2, \dots c_{n-1}\}) \cdot \{1, c_n\}, \quad (2.21)$$

for every $\{c_1, c_2, \dots c_n\} \subseteq \mathbb{N};\ n \geq 2$, or equivalently,

$$FP(\{c_1, c_2, \dots c_n\}) = FP(\{c_1, c_2, \dots c_{n-1}\}) \cup c_n \cdot FP(\{c_1, c_2, \dots c_{n-1}\}).$$

Theorem 2.20 (Hegyvári-Ruzsa [HR16])**.** Let A be a set of integers $\overline{d}(A) > 0$. Let $f : \mathbb{N}_+ \to \mathbb{N}_+$ be any function. There exists an infinite set C of integers, such that

$$A - A \supseteq FS_f(C) \cup FP(C).$$

So we can conclude that $A - A$ contains both an additive and a multiplicative structure.

Proof. We start our proof by quoting Følner's theorem. We state it as a lemma:

Lemma 2.21 (Følner)**.** Let A be a set of integers with $\overline{d}(A) > 0$. There exists a Bohr-set $B(S, \varepsilon)$ such that

$$E := B(S, \varepsilon) \setminus (A - A)$$

has density 0.

See the proof in [Fo].

We have a Bohr set for which the exceptional set has density zero, i.e. for some $B = B(S, \varepsilon)$, $E := B(S, \varepsilon) \setminus (A - A)$, $d(E) = 0$.

Recall that every Bohr set has positive density, and for every pair of sets S, S' and for every k, $0 < k \cdot \varepsilon' \le \varepsilon$, we have

$$k \cdot B(S, \varepsilon') \subseteq B(S, \varepsilon), \tag{2.22}$$

and

$$B(S \cup S', \varepsilon) = B(S, \varepsilon) \cap B(S', \varepsilon) \tag{2.23}$$

(see e.g. [TV] p. 165).

We will proof the existence of the infinite set C inductively.

Let $K_1 := f(1)$. Since any Bohr set has positive density and the exceptional set has zero density, furthermore by (2.22) one can find an element c_1 from $B(S, \varepsilon/K_1)$ such that $ic_1 \notin E$, for $i = 1, 2, \dots K_1$. So we have

$$FS_f(\{c_1\}) \cup FP(\{c_1\}) = \{0, c_1, \dots, K_1 c_1\} \subseteq B \setminus E \subseteq A - A.$$

Assume now that the elements $c_1 < c_2 < \dots < c_n$ have been defined with the property

$$\mathcal{F}_n := FS_f(\{c_1, c_2, \dots, c_n\}) \cup FP(\{c_1, c_2, \dots, c_n\}) \subseteq B \setminus E \subseteq A - A.$$

Write $FP(\{c_1, c_2, \dots, c_n\}) = \{p_1 < p_2 < \dots < p_m\}$, and let $K := \max\{f(n+1), p_m\}$. Define

$$\varepsilon_1 = \frac{1}{K} \min\{\varepsilon - \|xs\| : x \in FS_f(\{c_1, c_2, \dots, c_n\}); s \in S\}, \tag{2.24}$$

and let $B_1 := B(S, \varepsilon_1)$. Note that $B(S, \varepsilon_1) \subseteq B = B(S, \varepsilon)$.

By (2.24) we have that for every non-negative integer $i \le K$, for every $u \in FS_f(\{c_1, c_2, \dots, c_n\})$, for every $c \in B_1$ and $s \in S$

$$\|s(u + ic)\| < \varepsilon$$

holds, hence

$$FS_f(\{c_1, c_2, \dots c_n\}) + \{0, c, 2c, \dots K \cdot c\} \subseteq B.$$

Now we claim that there exists an element $c \in B_1$, with $c > c_1$ for which,

$$FS_f(\{c_1, c_2, \dots c_n\}) + \{0, c, 2c, \dots K \cdot c\} \subseteq B \setminus E \subseteq A - A$$

also holds.

Assume to the contrary that for every $c \in B_1$ with $c > c_1$ there would be at least one element $x \in FS_f(\{c_1, c_2, \dots c_n\})$ and one integer $j \in [1, \dots, K]$ for which $x + jc \in E$. Since $d(B_1 \setminus [1, c_n]) > 0$, by the pigeonhole principle there would be an $x_0 \in FS_f(\{c_1, c_2, \dots c_n\})$, $j_0 \in [1, \dots, K]$ and a $B_1' \subseteq B_1$, such that $\underline{d}(B_1) > 0$ and $x_0 + j_0 B_1' \subseteq E$ contradicting the fact that $d(E) = 0$ and $\underline{d}(x_0 + j_0 B_1') > 0$.

Let c_{n+1} be any such c. Since $K \geq p_m$ and $0 \in FS_f(\{c_1, c_2, \dots, c_n\})$ we have

$$c_{n+1} \cdot FP(\{c_1, c_2, \dots, c_n\}) \subseteq \{0, c_{n+1}, 2c_{n+1}, \dots, K \cdot c_{n+1}\} \subseteq B \setminus E.$$

Then by (2.21) and by the inductive hypothesis $FP(\{c_1, c_2, \dots, c_n, c_{n+1}\}) \subseteq B \setminus E$. Moreover $K > f(n+1)$,

$$FS_f(\{c_1, c_2, \dots c_n, c_{n+1}\}) \subseteq$$

$$\subseteq FS_f(\{c_1, c_2, \dots c_n\}) + \{0, c_{n+1}, 2c_{n+1}, \dots, K \cdot c_{n+1}\} \subseteq B \setminus E.$$

Thus we have that

$$\mathcal{F}_{n+1} \subseteq B \setminus E \subseteq A - A,$$

as we wanted.

So our desired set is

$$C := \{c_1 < c_2 < \dots c_n < c_{n+1} < \dots\}.$$

$\square$

2.3 Character sums on Hilbert cubes

A frequently asked question of the theory of character sums is to bound the values of $|\widetilde{f(u)}|$ and $|\widehat{g}(x)|$.

Recently many authors investigate character sums on certain structured sets. For instance let us mention a result of Bourgain and Garaev or a recent work of Petridis and Shparlinski in which they investigated trilinear character sums. Further works are due to Garaev, Konyagin and Shkredov.

To understand better a Hilbert cube, in the present section we are going to investigate (additive and multiplicative) character sums on (multiplicative and additive) cubes. For this treatment we will estimate energies of cubes.

Let us start with the following observation of Montgomery (see e.g., [Ga]): Let $U \subseteq \mathbb{F}_p$ be an arbitrary subset and $A \subseteq U$ for which $|A| \ll \log p$. Let $A(x)$ be its characteristic function,

$$A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases},$$

then

$$\max_{r \neq 0} |\widehat{A}(r)| \gg |A|.$$

As a contrast we quote a paper of Ajtai et al. ([ASz]) where the authors construct a set $T \subseteq \mathbb{Z}_m$ for which

$$|T| = O(\log m (\log^* m)^{c' \log^* m}) \ c' > 0,$$

and

$$\max_{r \neq 0} |\widehat{T}(r)| \leq O(|T| / \log^* m)$$

(where $\log^* m$ is the multi-iterated logarithm) hold. For structured set note a theorem of Bourgain; if H is a multiplicative subgroup of $\mathbb{F}_p^*$ of order $|H| > e^{c \log p / \log \log p}$, then $|\sum_{h \in H} e_p(rh)| = o(|H|); \quad p \to \infty, (r \neq 0, c > 0)$ (see e.g., [Ga]).

Our aim of this section is to show that the L_1-norm of a character sum on a Hilbert cube is big in some respect.

We will prove:

Theorem 2.22. [Hegyvári [HE16]] Let $\Delta \in (0,1]$, $r > 1$, $r \in \mathbb{N}$, and let $H_r(x_0, a_1 < a_2 < \cdots < a_d)$ be an arbitrary Δ-degenerate Hilbert cube. We have

$$\sum_{\chi} \Big| \sum_{h \in H} \chi(h) \Big| \gg \begin{cases} \sqrt{p} |H|^{3/2 - \gamma_r/2} & |H| < p^{2/3} \\ p^{3/2} |H|^{-\gamma_r/2} & |H| \geq p^{2/3} \end{cases}$$

where $\gamma_r = \frac{\log_{r+1}(2r+1)}{\Delta}$.

Furthermore we investigate additive characters on Hilbert cubes of order 1. As we noted if $A \subseteq U \subseteq \mathbb{F}_p$ and $|A| \gg \log p$ then $\max_{r \neq 0} |\widehat{A}(r)| \geq c|A|$.

We are going to show that from a non-degenerate Hilbert cube we can select more elements having this property:

Theorem 2.23. [Hegyvári [HE16]] Let $H(x_0, a_1 < a_2 < \cdots < a_d)$ be an arbitrary non-degenerate Hilbert cube. For every $\xi \in \mathbb{F}_p^*$ there is a subset $H' \subseteq H$ with $|H'| \gg e^{c\sqrt{\log|H|}}$, such that

$$|\widehat{H'}(\xi)| \gg |H'|.$$

2.3.1 Energies of Hilbert cubes

Energies inform us about the arithmetical structure of the given set. It is easy to see that for both the additive and multiplicative energies we have

$$|A|^2 \ll E_+(A), E_\times(A) \ll |A|^3.$$

First of all we investigate multiplicative energy. We have

Proposition 2.24. *[Hegyvári [HE16]] Let $\Delta \in (0,1]$; $r > 1$, $r \in \mathbb{N}$ and let $H = H_r(x_0, a_1 < a_2 < \cdots < a_d)$ be an arbitrary Δ-degenerate Hilbert cube. We have*

$$E_\times(H) \ll \begin{cases} |H|^{\gamma_r} p & |H| < p^{2/3} \\ \frac{|H|^{3+\gamma_r}}{p} & |H| \geq p^{2/3} \end{cases} \tag{2.25}$$

where $\gamma_r = \frac{\log_{r+1}(2r+1)}{\Delta}$.

Remark 2.25. Note that the estimations above are nontrivial if H is not "too degenerate" (Δ is close to 1). For example find an $|H| \asymp p^{2/3}$. Since $1 < \log_{r+1}(2r+1) < \log_{r+1}(2r+2) = 1 + \frac{\log 2}{\log(r+1)}$, thus when r is "big", then $|H|^{\gamma_r} p$ is close to $|H|^{5/2}$, which is better than the trivial bound $|H|^3$.

Proof. Pick elements $h, h' \in H$. Then h and h' can be written as

$$h = x_0 + \sum_{i=1}^{d} \varepsilon_i a_i; \quad \varepsilon_i \in \{0, 1, \ldots, r\} \quad \text{and} \quad h' = x_0 + \sum_{i=1}^{d} \varepsilon_i' a_i; \quad \varepsilon_i' \in \{0, 1, \ldots, r\}.$$

Hence

$$h + h' = 2x_0 + \sum_{i=1}^{d} \eta_i a_i; \quad \eta_i \in \{0, 1, 2, \ldots, 2r\}.$$

So we have

$$|H+H| \le (2r+1)^d = |H|^{\frac{\log_{r+1}(2r+1)}{\Delta}}. \tag{2.26}$$

Now we need the following lemma:

Lemma 2.26. Let $A \subseteq \mathbb{F}_p$. Then

$$E_\times(A) \ll \max\Big\{|A+A|p, \frac{|A+A||A|^3}{p}\Big\}. \tag{2.27}$$

For (2.27) see e.g., [Ga, Lemma 3.4],

If $|A| \ge p^{\frac{2}{3}}$, we get that in (2.27) the second term dominates the first one, otherwise the first dominates the second one. Now one can estimate the energy of a Hilbert cube by (2.26).

□

Secondly for the additive energy we argue as follows: the set

$$H = \{x_0 = 0, 1 < 2 < \cdots < d\} \subseteq \mathbb{F}_p$$

shows that the additive energy of a Hilbert cube could be large (larger than the trivial lower bound $|H|^2$) for arbitrary dimension.

Let H be an arbitrary Hilbert cube. Denote by $R(x)$ the number of representations of x as a sum of two elements of H, i.e. let $R(x) = |\{(h_1, h_2) \in H : x = h_1 + h_2\}|$. It is easy to see that

$$|H|^2 = \sum_{x \in \mathbb{F}_p} R(x) \quad \text{and} \quad E_+(H) = \sum_{x \in \mathbb{F}_p} R^2(x).$$

Furthermore by the Cauchy inequality

$$|H|^4 = \Big(\sum_{x \in \mathbb{F}_p} R(x)\Big)^2 \le |H+H| \cdot \sum_{x \in \mathbb{F}_p} R^2(x) = |H+H|E_+(H).$$

Thus by (2.26) for Hilbert cubes of order one, we get

$$E_+(H) \ge |H|^{4-\log_2 3/\Delta}.$$

So when the Hilbert cube is non-degenerate, we have $E_+(H) \gtrsim |H|^{2.415}$.

In the rest of this section we are going to investigate the additive energy of *multiplicative* Hilbert cubes.

By a multiplicative Hilbert cube we mean the set

$$H^{\times}(x_0, a_1, a_2, \dots, a_d) = \Big\{x_0 \cdot \prod_{1\le i\le d} a_i^{\varepsilon_i}\Big\} \quad \varepsilon_i \in \{0,1\}$$

where $x_0, a_1, a_2, \dots, a_d \in \mathbb{Z}_p^*$ and define $H_r^{\times}$ $(r \in \mathbb{N})$ in the same way as in the additive case.

Let g be a primitive root modulo p, and write the elements of $\mathbb{Z}_p^*$ in the form $a = g^b$. For a subset X of $\mathbb{Z}_p^*$ write $indX := \{y : g^y \in X\}$.

Observe that $H^{\times}(x_0, a_1, a_2, \dots, a_d)$ is a multiplicative Hilbert cube if and only if,

$indH^{\times}(x_0, a_1, a_2, \dots, a_d)$ is an additive Hilbert cube.

This easily implies

Fact 2.27. Assume that $S \subseteq \mathbb{F}_p^*$, $|S| \gg p^{1-1/2^d}$ then S contains a multiplicative $d-$dimensional Hilbert cube.

(e.g. see in [GR])

We prove the following

Proposition 2.28. *[Hegyvári [HE16]] Let* $H^{\times} := H^{\times}(x_0, a_1, a_2, \dots, a_d) \subseteq \mathbb{F}_p^*$; $|H^{\times}| = p^{\alpha}$; $\alpha > \frac{13}{18}$ *be a multiplicative Hilbert cube and write* $H_2^{\times} = H_2^{\times}(x_0, a_1, a_2, \dots, a_d)$. *We have*

$$E_+(H^{\times}) \ll |H^{\times}|^3 \Big(\frac{|H_2^{\times}|}{p}\Big)^{1/5}.$$

It concludes

Corollary 2.29. Let $H^{\times} := H^{\times} := H^{\times}(x_0, a_1, a_2, \dots, a_d) \subseteq \mathbb{F}_p^*$; $|H^{\times}| = p^{\alpha}$; $\alpha > \frac{13}{18}$ be a multiplicative Hilbert cube and assume that $|H_2^{\times}| \ll |H^{\times}|^{1+\varepsilon}$ for some $\varepsilon > 0$. Then

$$E_+(H^{\times}) \ll |H^{\times}|^{3-\delta}$$

where $\delta = \frac{1-\alpha(1+\varepsilon)}{5\alpha}$.

i.e., we obtain a non-trivial bound for the additive energy.

Proof of Proposition 2.28

Write the additive energy of $H^\times$ in the form $E_+(H^\times) = \frac{|H^\times|^3}{K}$. Then by the Gowers, Balog-Szemerédi theorem (see [TV]) there is an $\overline{H}^\times \subseteq H^\times$, for which $|\overline{H}^\times| > \frac{|H^\times|}{K}$ and $|\overline{H}^\times + \overline{H}^\times| < K^5|\overline{H}^\times|$. We need the following

Lemma 2.30. Let $A \subseteq \mathbb{F}_p^*$. Then

$$|A + A||AA| \gg \min\left\{p|A|, \frac{|A|^4}{p}\right\}.$$

In particularly when $|A| \geq p^{2/3}$ then $|A + A||AA| \gg p|A|$.

This is Proposition 1.1 in [Ga].

To use the lower bound $p|\overline{H}^\times|$ we need $\frac{|H^\times|}{K} \geq p^{2/3}$ or equivalently

$$K < \frac{|H^\times|}{p^{2/3}} = p^{\alpha - 2/3}. \tag{2.28}$$

By this lemma and by $\overline{H}^\times \cdot \overline{H}^\times \subseteq H_2^\times$, we obtain

$$K^5|\overline{H}^\times| > |\overline{H}^\times + \overline{H}^\times| \gg \frac{|\overline{H}^\times|p}{|H_2^\times|}$$

thus

$$K \gg \left(\frac{p}{|H_2^\times|}\right)^{1/5}. \tag{2.29}$$

Note that $|H_2^\times| \geq |H^\times|$, thus for (2.28) and (2.29) we need

$$(p^{1-\alpha})^{1/5} < p^{\alpha - 2/3}$$

which holds, since $\alpha > \frac{13}{18}$.

Hence $E_+(H^\times) = \frac{|H^\times|^3}{K}$ can be bounded by $|H^\times|^3 \left(\frac{H_2^\times}{p}\right)^{1/5}$ as we claimed.

2.3.2 Proof of Theorem 2.22 and 2.23

Proof of Theorem 2.22. First we are going to detect a connection between the L_1 norm of a character sum and the multiplicative energy of an arbitrary subset. We need the following

Lemma 2.31. Let $A \subseteq \mathbb{F}_p^*$. Then

$$\sum_\chi \Big| \sum_{a\in A} \chi(a) \Big| \gg p \frac{|A|^{3/2}}{E_\times(A)^{1/2}}. \tag{2.30}$$

This lemma is a multiplicative analogous of an additive one (see [Ka]). Since this form is not stated explicitly in the literature, we include a simple proof here.

Proof. Write

$$\widetilde{A_\chi} := \sum_{a\in A} \chi(a).$$

Using the identity $\sum_{u\in\mathbb{F}_p^*} |\widetilde{f(u)}|^2 = (p-1)\sum_{x\in\mathbb{F}_p^*} |f(x)|^2$ we have

$$\sum_\chi \left|\widetilde{A_\chi}\right|^2 = (p-1)|A|.$$

By the Hölder inequality we get

$$(p-1)|A| = \sum_\chi \left|\widetilde{A_\chi}\right|^2 = \sum_\chi \left|\widetilde{A_\chi}\right|^{2/3} \left|\widetilde{A_\chi}\right|^{4/3} \le$$

$$\le \Big(\sum_\chi |\widetilde{A_\chi}| \Big)^{2/3} \Big(\sum_\chi |\widetilde{A_\chi}|^4 \Big)^{1/3}.$$

By the orthogonality $\Big(\sum_\chi |\widetilde{A_\chi}|^4 \Big)$ is just $(p-1)\cdot E_\times(A)$, so we get

$$\sqrt{\frac{(p-1)^3|A|^3}{(p-1)\cdot E_\times(A)}} \le \sum_\chi |\widetilde{A_\chi}|$$

from which we get the statement.

□

dc_1260_16

Now we can combine Proposition 2.24 and Lemma 2.31.

By Lemma 2.31 with H in place A we obtain

$$\sum_{\chi}\Big|\sum_{h\in H}\chi(h)\Big| \gg \begin{cases}\sqrt{p}|H|^{3/2-\gamma_r/2} & |H|<p^{2/3}\\ p^{3/2}|H|^{-\gamma_r/2} & |H|\geq p^{2/3}\end{cases}$$

□

Proof of Theorem 2.23. Let $k = 6\cdot\lfloor\sqrt{d}\rfloor$. Split the set $A = \{a_1 < a_2 < \cdots < a_d\}$ into blocks $A_{i+1} = \{a_{ik+1} < a_{ik+2} < \cdots < a_{(i+1)k}\}$; $i = 0,1,2,\ldots,\lfloor\frac{d}{k}\rfloor$, (leave the remaining rightmost elements of A if it is necessary) and let $B_{i+1} := \{e_p(\xi\cdot\sum_{j=ik+1}^{t} a_j) : \ ik+1 \leq t \leq (i+1)k\}$ be the corresponding sets. Since H is a non-degenerate Hilbert cube, we get that all sets B_i have k many elements. Hence there are $t_1 < t_2$ such that the difference of the arguments of $e_p(\xi\cdot\sum_{j=ik+1}^{t_1} a_j)$ and $e_p(\xi\cdot\sum_{j=ik+1}^{t_2} a_j)$ is at most $\frac{2\pi}{k}$, and thus

$$\arg\left(e_p(\xi\cdot\sum_{j=t_1+1}^{t_2} a_j)\right) \leq \frac{2\pi}{k} \tag{2.31}$$

Write $a'_{i+1} := \sum_{j=t_1}^{t_2} a_j$, let $m := \lfloor\frac{d}{k}\rfloor$, and write $\omega_{i+1} = e_p(\xi\cdot a'_{i+1})$; $\quad i = 0,1,\ldots m$. Here $m = c'\sqrt{d} = c\sqrt{\log|H|}$. We have

$$H' := \Big\{x_0 + \sum_{1\leq i\leq m}\varepsilon_i a'_i; \quad \varepsilon_i\in\{0,1\}\Big\} \subseteq H(x_0,a_1,a_2,\ldots,a_d)$$

Furthermore by (2.31) we argue that for all $\underline{\varepsilon} = \{\varepsilon_1,\ldots,\varepsilon_m\}\in\{0,1\}^m$

$$\arg\Big(\sum_{1\leq i\leq m}\varepsilon_i\omega_i\Big) \leq \frac{2\pi}{6} \tag{2.32}$$

By (2.32) we obtain that

$$|\widehat{H'}(\xi)| \gg 2^m = e^{c\sqrt{\log|H|}}$$

for some $c>0$.

□

2.4 On a problem of Brown, Erdős and Freedman

An old question in number theory is to find *structures* in certain sets, for example in the set of primes, in the set of squares e.t.c. . Brown, Erdős and Freedman [BEF] asked whether $\mathcal{Q}$ contains arbitrarily large Hilbert cubes.

Recall

$$F_{\mathcal{A}}(n) = \max\{k : \text{there is a k-cube in } \mathcal{A} \cap \{1, 2, \ldots, n\}\}.$$

So the question of Brown et al can be formulated in the following form: is it true that

$$F_{\mathcal{A}}(n) \to \infty$$

as $n \to \infty$?

A related old question is due to Erdős and Moser. They asked whether there are arbitrarily large sets $\mathcal{A} \subset \mathbb{N}$ such that for all $a, a' \in \mathcal{A}$; $a \neq a'$ we have $a + a' \in \mathcal{Q}$.

The question of Brown, Erdős and Freedman remains open; our goal in this section is to show that the dimensions $F_{\mathcal{Q}}(n)$ and $F_{\mathcal{P}}(n)$ are not too big.

Before results below a theorem of Rivat, Sárközy and Stewart [RSS] was known; they proved that $F_{\mathcal{Q}}(n) \ll \log n$. First we improve this result.

2.4.1 The case of squares and primes

Theorem 2.32 (Hegyvári-Sárközy [HS99])**.** For $n > n_0$ we have

$$F_{\mathcal{Q}}(n) < 48\sqrt[3]{\log n}.$$

To prove Theorem 1, first we shall have to study the modular analogue of the problem. Let $f(p)$ denote the cardinality of the largest subset $\mathcal{A} \subset \mathbb{Z}_p$ with the property that for some $d \in \mathbb{Z}_p$ every element of $d + FS(\mathcal{A})$ is a quadratic residue in $\mathbb{Z}_p$.

We will prove

Theorem 2.33 (Hegyvári-Sárközy [HS99])**.** For $\varepsilon > 0$, $p > p_0(\varepsilon)$ we have

$$f(p) < 12\sqrt[4]{p}.$$

Proof of Theorem 2.32 and 2.33. :

First we shall need the following result of Olson and its consequence:

Lemma 2.34. If p is a prime number and $a_1, a_2, \dots, a_s$ are non-zero residues modulo p such that $a_i \neq \pm a_j$ for $i \neq j$, then

$$|FS(a_1, a_2, \dots, a_s)| \geq \frac{1}{2}\min\{p+3, s(s+1)\}.$$

Thus we conclude the following

Corollary 2.35. If p is a prime number and $\mathcal{R} \subseteq \mathbb{Z}_p$ then we have

$$|FS(\mathcal{R})| \geq \frac{1}{2}\min\{p+3, (|\mathcal{R}|^2-1)/4\}.$$

Write

$$G(h,p) = \sum_{x=0}^{p-1} e_p(hx^2)$$

and shortly $G_0 = G(1,p)$. It is well-known that $|G_0| = \sqrt{p}$ and $|G(h,p)| = |G_0|$ for $h \neq 0$ and $G(0,p) = p$..

Assume that $d \in \mathbb{Z}_p$, $\mathcal{A} = \{a_1, a_2, \dots, a_k\} \subseteq \mathbb{Z}_p$. Split the cube into two parts;

$$B := d + FS(a_1, a_2, \dots, a_{[k/2]}) \quad C := FS(a_{[k/2]+1}, \dots, a_k),$$

so

$$B + C \subseteq H(d, a_1, a_2, \dots, a_k) \tag{2.33}$$

and so each elements of $B+C$ is quadratic residue in $\mathbb{Z}_p$.

Then by Corollyary 2.35 we have

$$\min\{|B|, |C|\} \geq \frac{1}{2}\min\{p+3, (\lfloor k/2 \rfloor^2-1)/4\}. \tag{2.34}$$

Let

$$T = \sum_{x=0}^{p-1}\Big(\sum_{b\in B} e_p(bx^2)\Big)\Big(\sum_{c\in C} e_p(cx^2)\Big).$$

Then by (2.33) we have

$$|T| = \Big|\sum_{x=0}^{p-1}\sum_{b\in B}\sum_{c\in C} e_p((b+c)x^2)\Big)\Big| = \Big|\sum_{b\in B}\sum_{c\in C} G(b+c,p)\Big| \geq$$

$$\geq \Big|\sum_{b\in B}\sum_{c\in C} G_0\Big| - \sum_{b\in B}\sum_{c\in C}|G_0 - G(b+c,p)| = |B||C||G_0| - \sum_{b\in B; c\in C; p|b+c}|G_0 - G(0,p)| \geq$$

$$\geq |B||C|\sqrt{p} - 2\sum_{b\in B; c\in C; p|b+c} 1 \geq |B||C|\sqrt{p} - 2\min\{|B|,|C|\}. \qquad (2.35)$$

We turn to the upper bound; by the Cauchy inequality

$$|T| = \sum_{x=0}^{p-1}\Big|\sum_{b\in B} e_p(bx^2)\Big|\Big|\sum_{c\in C} e_p(cx^2)\Big| \leq$$

$$\leq \Big(\sum_{x=0}^{p-1}\Big|\sum_{b\in B} e_p(bx^2)\Big|^2\Big)^{1/2}\Big(\sum_{x=0}^{p-1}\Big|\sum_{c\in C} e_p(cx^2)\Big|^2\Big)^{1/2}.$$

If x runs over $0, 1\dots, p-1$ then x^2 meets every residue class at most twice. Thus it follows that

$$|T| \leq \Big(2\sum_{y=0}^{p-1}\Big|\sum_{b\in B} e_p(by)\Big|^2\Big)^{1/2}\Big(2\sum_{y=0}^{p-1}\Big|\sum_{c\in C} e_p(cy)\Big|^2\Big)^{1/2} =$$

$$= 2\Big(\sum_{y=0}^{p-1}\Big|\sum_{b,b'\in B} e_p((b-b')y)\Big|^2\Big)^{1/2}\Big(\sum_{y=0}^{p-1}\Big|\sum_{c,c'\in B} e_p((c-c')y)\Big|^2\Big)^{1/2} =$$

$$= 2\sqrt{|B|p}\sqrt{|C|p} = 2p\sqrt{|B||C|}. \qquad (2.36)$$

It follows from (2.35) and (2.36) that

$$|B||C|\sqrt{p} \leq 2p(\sqrt{|B||C|} + \min\{|B|,|C|\} \leq 4p\sqrt{|B||C|}$$

whence

$$\min\{|B|,|C|\} \leq \sqrt{|B||C|} \leq 4\sqrt{p}. \qquad (2.37)$$

by (2.34), and (2.37) we have

$$\min\Big\{\frac{p+3}{2}, \frac{\lfloor k/2\rfloor^2 - 1}{8}\Big\} \leq 4\sqrt{p}. \qquad (2.38)$$

If $p > 57$ then $\frac{p+3}{2} > 4\sqrt{p}$ and thus it follows from (2.38)

$$\lfloor k/2\rfloor^2 - 1 \leq 32\sqrt{p}$$

For large p this implies
$$k = |A| < 12\sqrt[4]{p}$$
and this completes the proof of Theorem 2.33

$\square$

Now we are in the position to prove Theorem 2.32

Proof of Theorem 2.32. We start by a very important but simple result which called "Gallagher Larger Sieve":

Lemma 2.36. Let $A \subseteq [1, N]$ be a set of integers. Let $\mathcal{P}$ be any finite set of prime numbers and for each prime let $\nu(p)$ denote the number of residue classes modulo p that contain an element of A. We have
$$|A| \leq \frac{\sum_{p \in \mathcal{P}} \log p - \log n}{\sum_{p \in \mathcal{P}} \frac{\log p}{\nu(p)} - \log n}. \tag{2.39}$$

Using now this sieve we prove the following technical lemma:

Lemma 2.37. Let $K > 0$, $0 < \eta < 1$, $p_0 > 0$ and $\varepsilon > 0$, and write $C = (2K(1-\eta))^{1/(1-\eta)}$. Then there exists a number $n_0 = n_0(K, \eta, p_0, \varepsilon)$ such that if $n \in \mathbb{N}$, $n > n_0$, $\mathcal{A} \subset \{1, 2, \dots, n\}$ and, writing $U = C(\log n)^{1/(1-\eta)}$, we have
$$\nu(p) < Kp^\eta \tag{2.40}$$
for every prime p with $p_0 < p \leq U$ then
$$|\mathcal{A}| < (C + \varepsilon)(\log n)^{\eta/(1-\eta)}. \tag{2.41}$$

Proof. We use Lemma 2.36 with $\mathcal{P} = \{p : p \text{ prime};\ p_0 < p \leq U\}$. Then by (2.40) and the prime number theorem, for $n \to \infty$ the denominator in (2.39) is
$$\sum_{p \in \mathcal{P}} \frac{\log p}{\nu(p)} - \log n > \sum_{p_0 < p \leq U} \frac{\log p}{Kp^\eta} - \log n =$$
$$= \left(\frac{1}{K} + o(1)\right) \sum_{n \leq U/\log U} \frac{\log n}{(n \log n)^\eta} - \log n =$$
$$= \left(\frac{1}{K} + o(1)\right) \int_2^{U/\log U} \frac{(\log x)^{1-\eta}}{x^\eta} - \log n =$$

$$= \Big(\frac{1}{K} + o(1)\Big)\frac{1}{1-\eta}U^{1-\eta} - \log n = \Big(\frac{1}{K(1-\eta)} + o(1)\Big)U^{1-\eta} \qquad (2.42)$$

which is positive so that, indeed Lemma 2.36 can be applied.

Again by the prime number theorem, the numerator in (2.39) is

$$\sum_{p\in\mathcal{P}} \log p - \log n = \sum_{p_0<p\leq U} \log p - \log n = (1+o(1))U - \log n = (1+o(1))U. \qquad (2.43)$$

It follows from (2.39), (2.42) and (2.43) that

$$|\mathcal{A} \leq ("K(1-\eta) + o(1))U^{\eta} = (C+o(1))(\log n)^{\eta/(1-\eta)}$$

which proves 2.41 and this completes the proof the Lemma.

□

Now assume that there is a Hilbert $k-$cube $H(d, a_1, a_2, \dots, a_k)$ in $\mathcal{Q} \cap \{1, 2, \dots, n\}$.

This implies that for every prime p, every element of $H(d, a_1, a_2, \dots, a_k)$ is a quadratic residue in $\mathbb{Z}_p$. Thus by Theorem 2.33, the number of distinct residue classes amongst them is $\nu(p) < 12\sqrt[4]{p}$.

By using Lemma 2.37 with $K = 12;\ \eta = 1/4;\ \varepsilon = \frac{1}{100}$ it follows that for large n we have

$$k < (C+\varepsilon)(\log n)^{\eta/(1-\eta)} = \Big(18^{4/3} + \frac{1}{100}\Big)(\log n)^{1/3} < 48\sqrt[3]{\log n}.$$

□

We investigated the related problem in primes as well. We obtain the following:

Theorem 2.38 (Hegyvári-Sárközy [HS99])**.** For every $\varepsilon > 0$ and $n > n_0(\varepsilon)$ we have

$$H_{\mathcal{P}}(n) < (16+\varepsilon)\log n$$

Proof. We shall need the following result of Olson (which is derived from Lemma 2.34):

Lemma 2.39. If p is a prime, and $\mathcal{A}$ is set of distinct non-zero residue classesmodulo p, and
$$|\mathcal{A}| > \sqrt{4p-1}$$
then for every residue classes $x \in \mathbb{Z}_p$ we have $x \in FS(\mathcal{A})$.

Now we will prove that if $d \in \mathbb{N}$, $\mathcal{A}$, $|\mathcal{A}| = s$ is a finite subset of $\mathbb{N}$ and
$$d + FS(\mathcal{A}) \tag{2.44}$$
then defining $\nu(p)$ as in Lemma 2.36, we have
$$\nu(p) < 4\sqrt{p} + 3. \tag{2.45}$$
We will prove this by contradiction: assume that $\nu(p) \geq 4\sqrt{p} + 3$. Then there are integers
$$b_1, b_2, \dots, b_k \in \mathcal{A} \tag{2.46}$$
such that
$$k \geq 4\sqrt{p} + 2. \tag{2.47}$$
$$b_i \not\equiv b_j (mod\ p) \text{ for } 1 \leq i < j \leq k \tag{2.48}$$
$$b_i \not\equiv 0 (mod\ p) \text{ for } 1 \leq i \leq k. \tag{2.49}$$
Write $s = [k/2]$ so that by (2.47) we have
$$s > \frac{k}{2} - 1 \geq 2\sqrt{p} > \sqrt{4p-3}. \tag{2.50}$$
By Lemma 2.39 (and since $d, b_1, \dots, b_k$ are positive), it follows from (2.48),(2.49) and (2.50) that there are u, v with
$$u \in d + FS(\{b_1, \dots, b_s\}), \tag{2.51}$$
$$p|u;\ u > 0, \tag{2.52}$$
$$v \in FS(\{b_{s+1}, \dots, b_{2s}\}), \tag{2.53}$$
$$p|v;\ v > 0, \tag{2.54}$$
Then by (2.52) and (2.54) we have $p|u+v$, and $u+v \geq 2p$ so that $u+v$ is a composite number. Moreover, it follows from (2.46), (2.51) and (2.53) that
$$u + v \in d + FS(\{b_1, \dots, b_k\}) \subset d + FS(\mathcal{A})$$

which contradicts (2.44), and this completes the proof of (2.45).

By Lemma 2.36, it follows from (2.45) that if $n > n_1(\varepsilon)$ and

$$d + FS(A) \subset \mathcal{P} \cap [1, n]$$

then we have

$$|\mathcal{A}| = (16 + \varepsilon) \log n$$

which completes the proof.

Remark 2.40. Our results introduce many other results; e.g related to character sum estimation (Balasuriya and Shparlinski ([BaSh]), treatment and versions of Gallagher sieve (Croot and Elsholtz [CE]), and many improvements (Dietman-Elsholtz [DE1], [DE2], [W])

Let me mention that Wood observed – based on a work of Paturi, Saks and Zane – that the dimension of Hibert cube which contained in $\mathcal{P}$ is connected to the following problem: if C_n denotes the circuits Σ_2^3 (AND gates used as inputs, OR gate as output) tests whether the number $m = X_1 X_2 \dots X_n$ is a prime, then one can conclude the number of gates from the dimension $\dim(\mathcal{P})$.

2.4.2 On infinite Hilbert cubes

It is an interesting question that which well-know sequence contains an *infinite* Hilbert cube. Almost trivial that the set of squares $\mathcal{Q}$ and the set of all primes $\mathcal{P}$ do not contain an infinite cube.

Let $\mathcal{P}_k = \{n_1 < n_2 < \dots\}$ be the set of the positive integers composed of the primes not exceeding k. By a theorem of R. Tijdeman we know that

$$n_{k+1} - n_k \to \infty$$

as $k \to \infty$. Hence we conclude

Theorem 2.41 (Hegyvári-Sárközy)**.** The set $\mathcal{P}_k = \{n_1 < n_2 < \dots\}$ $(k \geq 2)$ does not contain an infinite cube.

Remark 2.42. Probably the set $\mathcal{W} := \{1; 4; 8; 9; \dots; n^k; \dots\}$ also possesses property above but this is not known, and presently this is probably beyond our reach.

Finally in this section we consider some result on special and general sets.

In [BR] Bergelson and Ruzsa proved the following interesting fact:

Theorem 2.43 (Bergelson-Ruzsa)**.** Let A be the sequence of squarefree numbers. For every $a_0 \in A$ contains an infinite Hilbert cube $H(a_0, x_1.x_2 \dots\}$ containing in A.

They derived this result from the following theorem:

Theorem 2.44 (Bergelson-Ruzsa)**.** Let $S \subset \mathbb{N}$ be a set such that $1 \notin S$ any two elements of S are coprime, and

$$\sum_{s \in S} \frac{1}{s} < \infty.$$

Then there is an infinite set X such that

$$FS(X) \subset B^c(S)$$

where $B^c(S)$ denotes the set of natural numbers that are not divisible by any element of S.

In [He08c] I obtained a related result:

Theorem 2.45 (Hegyvári)**.** Let $T := \{q_i : i \in \mathbb{N}\}$ be an increasing sequence of primes. Assume that there is an infinite Hilbert cube $H(a_0, x_1.x_2 \dots\} \subset B^c(T)$, where $B^c(T)$ denotes the set of natural numbers that are not divisible by any element of T. Then for each $n \in \mathbb{N}$,

$$H(n) < 8 \sum_{i=1}^{f(n)} q_i^{3/2}$$

where $f(n)$ is the smallest s for which $q_1 q_2 \cdots q_s \geq n$.

As a corollary we obtain

Corollary 2.46. Let $\alpha > 1$, and let $U := \{q_i : i \in \mathbb{N}\}$ be an increasing sequence of primes with

$$\lim_{k \to \infty} \frac{q_k}{k^\alpha} = 1,$$

and $H(a_0, x_1.x_2 \dots\} \subset B^c(U)$. Then we have

$$H(n) < c(\alpha) \Big(\frac{\log n}{\log \log n} \Big)^{\frac{3\alpha+2}{2}}.$$

dc_1260_16

We close this section a result on general set.

In [H79] E.G. Strauss proved that for every $\varepsilon > 0$ there exists a sequence with density $> 1-\varepsilon$ which does not contain an infinite Hilbert cube. On the other hand, it was proved in [Na] that every sequence of integers with density 1 contains an infinite Hilbert cube. Let us start with two remarks. Firstly note that for a given interval $I = [a, a+m]$, if a Hilbert cube $H(a_0, x_1 < \cdots < x_s)$ lies in I then clearly $s \ll \sqrt{m}$. Secondly if for some $A \subset [1, N]$, we would like to avoid A by an Hilbert cube, then statistically we have a gap with size $\frac{N}{|A|}$ and by the previous remark there is a cube with $|H| \sim c\sqrt{\frac{N}{|A|}}$. This argument works just in a finite case and completely false in the infinite case. However the next theorem shows that essentially apart from a $\log n$ factor a same conclusion remains true.

Theorem 2.47 (Hegyvári)**.** Let A be a sequence of integers and let $\omega : \mathbb{N} \to \mathbb{R}^+$ be any function and assume that $\omega(x) \to \infty$ as $x \to \infty$. Then there exists an infinite cube H which avoids A and for which

$$\limsup_{n\to\infty} \frac{H(n)}{\sqrt{n/A(n) \cdot \omega(n) \cdot \log^2 n}} > 0.$$

The proof of Theorem 2.47 can be found in [He08b].

□

Chapter 3

Additive Ramsey type problems

3.1 On a theorem of Raimi and Hindman

A branch of combinatorial analysis – called Ramsey theory – investigates partitions of certain structures. In [H79], p.180, Th 11.15] Hindman deals with the intersecting properties of a finite partition of the set $\mathbb{N}$ of positive integers. He gives an elementary proof for Raimi's theorem [RA68] which reads as follows:

Theorem 3.1. There exists $E \subseteq \mathbb{N}$ such that, whenever $r \in \mathbb{N}$ and $\mathbb{N} = \bigcup_{i=1}^{r} D_i$ there exist $i \in \{1, 2, \ldots, r\}$ and $k \in \mathbb{N}$ such that $(D_i + k) \cap E$ is infinite and $(D_i + k) \setminus E$ is infinite.

Hindman shows that the set E of natural numbers whose last non-zero entry in their ternary expansion is 1 satisfies this condition. Raimi's original proof used a topological result.

In the present section we are going to give a far-reaching generalization to this theorem.

Recall that a given a sequence $\{x_n\}_{n=1}^{\infty}$ in $\mathbb{N}$,

$$FS(\{x_n\}_{n=1}^{\infty}) = \{\textstyle\sum_{n \in F} x_n : F \text{ is a finite nonempty subset of } \mathbb{N}\}.$$

Now we state a generalization of Raimi's theorem.

Theorem 3.2 (Hegyvári [He05]). Let $A \subseteq \mathbb{N}$ be a sequence of integers such that there is a positive irrational γ for which $\{\langle \gamma x \rangle : x \in A\}$ is dense in $[0, 1)$. Let $r \in \mathbb{N}$ and let $\alpha_1, \alpha_2, \ldots, \alpha_r$ be positive real numbers such that $\sum_{i=1}^{r} \alpha_i = 1$. There exists a disjoint partition $\mathbb{N} = \bigcup_{i=1}^{r} E_i$ such that

(1) for every $i \in \{1, 2, \ldots, r\}$, $d(E_i) = \alpha_i$ and

(2) for each $t \in N$ and each partition $A = \bigcup_{j=1}^{t} F_j$, there exist $m \in \{1, 2, \ldots, t\}$ and a sequence $\{x_n\}_{n=1}^{\infty}$ in $\mathbb{N}$ such that for every $h \in FS(\{x_n\}_{n=1}^{\infty})$ and every $i \in \{1, 2, \ldots, r\}$, $(F_m + h) \cap E_i$ is infinite.

Notice that Raimi's theorem follows from the case $r = 2$, instead of an infinite set $\{x_n\}_{n=1}^{\infty}$ just a single integer k.

First we prove a technical lemma.

Lemma 3.3. Let $\{I_n\}_{n=1}^{\infty}$ be a sequence of pairwise disjoint intervals in $[0, 1)$ and assume that for every $\varepsilon > 0$ there exist $a \in [0, 1)$ and $m \in \mathbb{N}$ such that $\bigcup_{n=m+1}^{\infty} I_n \subseteq (a, a + \varepsilon)$. Let γ be a positive irrational number, and let $E = \{x \in \mathbb{N} : \langle \gamma x \rangle \in \bigcup_{n=1}^{\infty} I_n\}$. Then $d(E) = \sum_{n=1}^{\infty} \mu(I_n)$.

Proof of Lemma . Recall that if γ is a nonzero irrational number, then $\{\langle \gamma x \rangle : x \in \mathbb{N}\}$ is uniformly distributed mod 1. That is, if $0 \leq a < b \leq 1$, then

$$d(\{x \in \mathbb{N} : \langle \gamma x \rangle \in (a, b)\}) = b - a.$$

Let $\alpha = \sum_{n=1}^{\infty} \mu(I_n)$. Let $\varepsilon > 0$ be given and let $k \in \mathbb{N}$ be an integer such that $\sum_{n=1}^{k} \mu(I_n) > \alpha - \varepsilon$. Choose an $a \in [0, 1)$ and $m \in \mathbb{N}$ such that $\bigcup_{n=m+1}^{\infty} I_n \subseteq (a, a + \varepsilon)$. We may presume that $m \geq k$.

Let

$$F = \{x \in \mathbb{N} : \langle \gamma x \rangle \in \bigcup_{n=1}^{m} I_n\}$$

and let

$$G = \{x \in \mathbb{N} : \langle \gamma x \rangle \in \bigcup_{n=1}^{m} I_n \cup (a, a + \varepsilon)\}.$$

Now $\bigcup_{n=1}^{m} I_n \cup (a, a+\varepsilon)$ is a finite union of pairwise disjoint intervals of total length $\delta \leq \sum_{n=1}^{m} \mu(I_n) + \varepsilon$.

Therefore we have by the uniform distribution of $\{\langle \gamma x \rangle : x \in \mathbb{N}\}$ that $d(F) = \sum_{n=1}^{m} \mu(I_n)$ and $d(G) = \delta$. Thus $\underline{d}(E) \geq d(F) \geq \sum_{n=1}^{k} \mu(I_n) > \alpha - \varepsilon$ and $\overline{d}(E) \leq d(G) \leq \sum_{n=1}^{m} \mu(I_n) + \varepsilon \leq \alpha + \varepsilon$. $\square$

Proof of Theorem 3.2. Take a positive irrational γ for which $\{\langle\gamma x\rangle : x \in A\}$ is dense in $[0,1)$. Let $s_0 = 0$ and inductively for $i \in \{1,2,\ldots,r\}$, let $s_i = s_{i-1} + \alpha_i$ (so $s_r = 1$). For $i \in \{1,2,\ldots,r\}$ and $j \in \mathbb{N}$, let

$$J_{i,j} = \left[1 - \frac{1}{2^j} + \frac{s_{i-1}}{2^{j+1}}, 1 - \frac{1}{2^j} + \frac{s_i}{2^{j+1}}\right).$$

For $i \in \{1,2,\ldots,r\}$ let $J_i = \bigcup_{j=0}^{\infty} J_{i,j}$ and let $E_i = \{x \in \mathbb{N} : \langle\gamma x\rangle \in J_i\}$. Then $\mu(J_i) = \sum_{j=0}^{\infty} \dfrac{s_i - s_{i-1}}{2^{j+1}} = \alpha_i$ so by the lemma, $d(E_i) = \alpha_i$.

Now let $t \in N$ and let $A = \bigcup_{j=1}^{t} F_j$. We claim

Fact: *For any c, d with $0 \le c < d \le 1$ there exists $m \in \{1,2,\ldots,t\}$ and there exist a, b, with $c \le a < b \le d$ such that $\{\langle\gamma x\rangle : x \in F_m\}$ is dense in (a,b).*

To see this, suppose not. Let $a_0 = c$ and $b_0 = d$. Inductively let $j \in \{1,2,\ldots,t\}$. Then $\{\langle\gamma x\rangle : x \in F_j\}$ is not dense in (a_{j-1}, b_{j-1}) so pick a_j, b_j with $a_{j-1} \le a_j < b_j \le b_{j-1}$ such that $\{\langle\gamma x\rangle : x \in F_j\} \cap (a_j, b_j) = \emptyset$. When this process is complete one has that $(a_t, b_t) \cap \bigcup_{j=1}^{t} \{\langle\gamma x\rangle : x \in F_j\} = \emptyset$. That is, $(a_t, b_t) \cap \{\langle\gamma x\rangle : x \in A\} = \emptyset$, a contradiction.

Now for $n \in \mathbb{N}$, we inductively choose a_n, b_n, and $m(n)$ such that $m(n) \in \{1,2,\ldots,t\}$, $0 < a_n < b_n < 1$, $\{\langle\gamma x\rangle : x \in F_{m(n)}\}$ is dense in (a_n, b_n), $b_n \le a_{n+1}$, $a_{n+1} \ge 1 - \dfrac{b_n - a_n}{4}$, and $b_{n+1} - a_{n+1} \le \dfrac{b_n - a_n}{2}$.

Choose $m(1) \in \{1,2,\ldots,t\}$ and a_1, b_1 such that $0 < a_1 < b_1 < 1$ and $\{\langle\gamma x\rangle : x \in F_{m(1)}\}$ is dense in (a_1, b_1). Given $n \in \mathbb{N}$ and a_n and b_n, let

$$c = \max\left\{b_n, 1 - \frac{b_n - a_n}{4}\right\}$$

and

$$d = \min\left\{1, c + \frac{b_n - a_n}{2}\right\}.$$

Apply Fact to choose $m(n+1) \in \{1,2,\ldots,t\}$ and a_{n+1}, b_{n+1} with $c \le a_{n+1} < b_{n+1} \le d$ such that $\{\langle\gamma x\rangle : x \in F_{m(n+1)}\}$ is dense in (a_{n+1}, b_{n+1}). Then

$$b_n \le c \le a_{n+1},\ 1 - \frac{b_n - a_n}{4} \le c \le a_{n+1},$$

and

$$b_{n+1} \le d \le c + \frac{b_n - a_n}{2} \le a_{n+1} + \frac{b_n - a_n}{2}.$$

Now take $m \in \{1, 2, \dots, t\}$ such that $D = \{n : m(n) = m\}$ is infinite and enumerate D in increasing over as $\{n(k)\}_{k=1}^{\infty}$. For each $k \in \mathbb{N}$, let $c_k = a_{n(k)}$ and $d_k = b_{n(k)}$. Then for each k, $\{\langle \gamma x \rangle : x \in F_m\}$ is dense in (c_k, d_k), $d_k \le c_{k+1}$,

$$c_{k+1} \ge 1 - \frac{d_k - c_k}{4},$$

and

$$d_{k+1} - c_{k+1} \le \frac{d_k - c_k}{2}.$$

For each $k \in \mathbb{N}$ pick $x_k \in \mathbb{N}$ such that

$$\langle \gamma x_k \rangle \in \left(1 - d_k, 1 - c_k - \frac{d_k - c_k}{2}\right).$$

Notice that for any $k \in \mathbb{N}$ and $v \in \omega$, $d_{k+v} - c_{k+v} \le \dfrac{d_k - c_k}{2^v}$.

We show now by induction on $v \in \mathbb{N}$ that

$$\begin{aligned} & H \subseteq \mathbb{N},\ |H| = v, \text{ and } k = \min H \Rightarrow \\ & \Rightarrow \langle \gamma \textstyle\sum_{l \in H} x_l \rangle \in \left(1 - d_k, 1 - c_k - \dfrac{d_k - c_k}{2^v}\right). \end{aligned} \tag{3.1}$$

When $v = 1$, (3.1) holds directly, so assume that $v > 1$ and (3.1) holds for $v - 1$. Let $H \subseteq \mathbb{N}$ with $|H| = v$, let $k = \min H$, let $u = \max H$, and let $G = H \setminus \{u\}$. Then

$$\langle \gamma \sum_{l \in G} x_l \rangle < 1 - c_k - \frac{d_k - c_k}{2^{v-1}}$$

and

$$\langle \gamma x_u \rangle < 1 - c_u \le 1 - c_{k+v-1} \le \frac{d_{k+v-2} - c_{k+v-2}}{4} \le \frac{d_k - c_k}{2^v}$$

so

$$\langle \gamma \sum_{l \in G} x_l \rangle + \langle \gamma x_u \rangle < 1 - c_k - \frac{d_k - c_k}{2^{v-1}} + \frac{d_k - c_k}{2^v} = 1 - c_k - \frac{d_k - c_k}{2^v}.$$

Since $\langle\gamma\sum_{l\in G}x_l\rangle+\langle\gamma x_u\rangle<1$, we have that

$$\langle\gamma\sum_{l\in G}x_l\rangle+\langle\gamma x_u\rangle=\langle\gamma\sum_{l\in H}x_l\rangle$$

and so (3.1) is established.

Now let H be a finite nonempty subset of $\mathbb{N}$ and let $i\in\{1,2,\ldots,r\}$. We show that $(F_m+\sum_{l\in H}x_l)\cap E_i$ is infinite. Let $k=\min H$. Then by (3.1),

$$\langle\gamma\sum_{l\in H}x_l\rangle\in(1-d_k,1-c_k)$$

so

$$c_k+\langle\gamma\sum_{l\in H}x_l\rangle<1<d_k+\langle\gamma\sum_{l\in H}x_l\rangle.$$

Pick $j\in\mathbb{N}$ such that $1-\dfrac{1}{2^j}>c_k+\langle\gamma\sum\limits_{l\in H}x_l\rangle$. Then

$$c_k<1-\frac{1}{2^j}-\langle\gamma\sum_{l\in H}x_l\rangle+\frac{s_{i-1}}{2^{j+1}}<1-\frac{1}{2^j}-\langle\gamma\sum_{l\in H}x_l\rangle+\frac{s_i}{2^{j+1}}<d_k$$

and $\{\langle\gamma y\rangle:y\in F_m\}$ is dense in (c_k,d_k) and so

$$K=\left\{y\in F_m:1-\frac{1}{2^j}-\langle\gamma\sum_{l\in H}x_l\rangle+\frac{s_{i-1}}{2^{j+1}}<\langle\gamma y\rangle<1-\frac{1}{2^j}-\langle\gamma\sum_{l\in H}x_l\rangle+\frac{s_i}{2^{j+1}}\right\}$$

is infinite.

To complete the proof it suffices to show that if $y\in K$, then

$$y+\sum_{l\in H}x_l\in E_i.$$

Indeed, given $y\in K$,

$$\langle\gamma y\rangle+\langle\gamma\sum_{l\in H}x_l\rangle\in J_{i,j}$$

and $\langle\gamma y\rangle+\langle\gamma\sum_{l\in H}x_l\rangle<1$ so

$$\langle\gamma y\rangle+\langle\gamma\sum_{l\in H}x_l\rangle=\langle\gamma(y+\sum_{l\in H}x_l)\rangle$$

so $y+\sum_{l\in H}x_l\in E_i$ as required.

$\square$

Remark 3.4. Theorem 3.2 implies that for every t partition of the set $\mathbb{N} = \bigcup_{j=1}^{t} F_j$ not just one translation h of some F_m meets E_j : $(j = 1, \ldots, r)$ in an infinite set, rather each translations do, given h from an additive "cube".

A natural question is to ask the following: Is any infinite set $\{x_n\}_{n=1}^{\infty}$, such that Theorem 3.2 remains true if we want that the elements h included in $FS(\{x_n\}_{n=1}^{\infty}) \cup FP(\{x_n\}_{n=1}^{\infty})$, where $FP(\{x_n\}_{n=1}^{\infty})$ is a multiplicative cube defined by

$$FP(\{x_n\}_{n=1}^{\infty}) = \{\textstyle\prod_{n\in F} x_n : F \text{ is a finite nonempty subset of } \mathbb{N}\}\,?$$

Our combinatorial approach is not enough to prove this extension. Maybe some tools from ergodic theory would work.

3.2 A Ramsey type question of Sárközy

A set $\mathcal{A}$ of positive integers is called an asymptotic basis of order h if any large enough integer is a sum of at most h elements of $\mathcal{A}$, the integer h being the least one such that this property holds. In [AS3], A. Sárközy considered the problem of estimating the maximal order $H(k)$, as asymptotic bases, of the subsequences of primes having a positive relative density $1/k$. He obtained the upper bound $H(k) \ll k^4$ and the lower bound $H(k) \gg k \log\log k$. Later Ramaré and Ruzsa improved almost definitively this result by showing $H(k) \asymp k \log\log k$ (cf. [RR]).

A Ramsey type version of this problem is also due to Sárközy who raised the following question (see in [AS2]): one can see that for all $k \in \mathbb{N}$, there is a number $t = t(k)$ with the property that for any k-colouring of the set of squares every integer large enough can be represented as the monochromatic sum of at most t squares. Then what is the smallest number $t = t(k)$ having this property, and also the similar problem for the primes.

To describe our result we define the concept of the order of K partition.

Definition 3.5. For any integer positive K and any K-partition

$$\mathcal{U} = (\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_K)$$

of $\mathcal{A}$ as a union of K sets

$$\mathcal{A} = \bigcup_{k=1}^{K} \mathcal{A}_k,$$

dc_1260_16

we denote by $\mathrm{ord}(\mathcal{U})$ the least number h having the following property: for any sufficiently large integer n there exists k such that n is a sum of at most h elements of $\mathcal{A}_k$. We finally denote

$$\mathrm{ord}_K(\mathcal{A}) = \sup\{\mathrm{ord}(\mathcal{U}) : \mathcal{U} \text{ is a } K\text{-partition of } \mathcal{A}\}.$$

First we quote an important "finite type Kneser theorem" which is due to Sárközy

Lemma 3.6. Let N and k be positive integers and $\mathcal{A} \subset \{1, 2, \ldots, N\}$ such that

$$|\mathcal{A}| > \frac{N}{k} + 1.$$

Then there exist integers d, h, m such that

$$1 \leq d \leq k - 1,$$
$$1 \leq h \leq 118k,$$

and

$$\{(m+1)d, (m+2)d, \ldots, (m+N)d\} \subset h\mathcal{A}.$$

3.2.1 The squares

First we give an upper bound.

Theorem 3.7. [Hegyvári-Hennecart [HH07]]
Let K be an integer. Then

$$\mathrm{ord}_K(\mathcal{Q}) \leq c_3 (K \log K)^5.$$

where c_3 can be taken equal to 10^9.

Proof. Let

$$\mathcal{Q} = \bigcup_{k=1}^{K} \mathcal{Q}_k, \tag{3.2}$$

be a partition of the squares. Let N_0 be an integer large enough such that for any $N \geq N_0$

$$\pi(\sqrt{N}) - \pi(\sqrt{N}/2) \geq K + 1.$$

Take any $N \geq N_0$ and put

$$\mathcal{S}_k = \mathcal{Q}_k \cap [N/4, N], \qquad k = 1, 2, \ldots, K.$$

For each prime p, let

$$I_p = \{1 \leq k \leq K \, : \, \mathcal{S}_k \subset \mathbb{N}p\}.$$

We then define recursively the following, possibly empty, increasing sequence of prime numbers:

$$\begin{aligned} q_1 &= \min\{p \, : \, I_p \neq \emptyset\}, \\ q_j &= \min\{p \, : \, I_p \setminus (I_{q_1} \cup \cdots \cup I_{q_{j-1}}) \neq \emptyset\}, \quad j \geq 2. \end{aligned}$$

This sequence is clearly finite: $q_1 < q_2 < \cdots < q_r$ with $|I_{q_1} \cup \cdots \cup I_{q_r}| \leq K - 1$. We denote $\mathcal{K}'$ the complementary set of $I_{q_1} \cup \cdots \cup I_{q_r}$ in $\{1, 2, \ldots, K\}$. We have

$$\begin{aligned} \left| \bigcup_{k \in \mathcal{K}'} \mathcal{S}_k \right| &\geq \prod_{j=1}^{r} \left(1 - \frac{1}{q_j}\right) \frac{\sqrt{N}}{2} \\ &\geq \prod_{j=1}^{K-1} \left(1 - \frac{1}{p_j}\right) \frac{\sqrt{N}}{2} \\ &\geq \frac{\sqrt{N}}{4 \log K}, \end{aligned}$$

by an explicit lower bound in Mertens' formula, where $p_1 < p_2 < \cdots$ is the increasing sequence of prime numbers.

Hence there exists some $k \in \mathcal{K}'$ such that

$$|\mathcal{S}_k| \geq \frac{\sqrt{N}}{4K \log K}.$$

Let us denote by $r_{\mathcal{Q}}^{(5)}(n)$ the number of representation of n by five squares. Now we need a lemma:

Lemma 3.8. For any $n \geq 1$, we have

$$r_{\mathcal{Q}}^{(5)}(n) \leq 30 n^{3/2}. \tag{3.3}$$

The lemma is a simple consequence of Theorem 4, p. 180 in [EG].

Put $\mathcal{S} = \mathcal{S}_k$. We have

$$\left(\sum_{s\in\mathcal{S}} 1\right)^5 = |\mathcal{S}|^5 \geq \left(\frac{\sqrt{N}}{4K\log K}\right)^5,$$

and on the other hand

$$\begin{aligned}\left(\sum_{s\in\mathcal{S}} 1\right)^5 &= \sum_{n\in 5\mathcal{S}} r_{\mathcal{S}}^{(5)}(n) \leq \sum_{n\in 5\mathcal{S}} r_{\mathcal{Q}}^{(5)}(n)\\ &\leq |5\mathcal{S}| \max_{1\leq n\leq 5N} r_{\mathcal{Q}}^{(5)}(n) \leq 340N^{3/2}|5\mathcal{S}|,\end{aligned}$$

by (3.3), where we write $5\mathcal{S}$ for denoting the set of the sums of 5 elements from $\mathcal{S}$.

It satisfies $5\mathcal{S} \subset [5N/4, 5N]$ and

$$|5\mathcal{S}| \geq \frac{N}{c_1(K\log K)^5},$$

for some absolute constant $c_1 > 0$. Assuming N large enough, we deduce from Lemma 3.6 that there exist d with $1 \leq d \leq c_1(K\log K)^5$ such that for some

$$h \leq h_0 = c_2(K\log K)^5,$$

we have

$$\{(m+1)d, (m+2)d, \ldots, (m+5N)d\} \subset 5h\mathcal{S},$$

for some m such that

$$\frac{5hN}{4} \leq md \quad \text{and} \quad (m+5N)d \leq 5hN.$$

Since k belongs to $\mathcal{K}'$, we see that $\mathcal{S} = \mathcal{S}_k$ contains some integer s coprime to d and satisfying

$$\frac{N}{4} \leq s \leq N.$$

Thus any integer in the interval

$$\mathcal{L} := \{(m+1)d + (d-1)N, (m+2)d + (d-1)N + 1, \ldots, (m+5N)d\}$$

can be written as a sum $x+js$ where $x \in 5h\mathcal{S}$ and $0 \le j \le d-1$. By shifting $\mathcal{L}$ by multiples of s and taking the union of the given intervals $\mathcal{L}+js$, $0 \le j \le l$, we get

$$[(m+N)d, (m+5N)d + lN/4] \subset \bigcup_{j=5h}^{5h+d-1+l} j\mathcal{S}.$$

Applying this argument to $N+1$ instead of N, we get for any $l' \ge 0$

$$[(m'+N+1)d', (m'+5N+5)d' + l'(N+1)/4] \subset \bigcup_{j=5h'}^{5h'+d'-1+l'} j\mathcal{S}',$$

where

$$\mathcal{S}' = \mathcal{Q}_{k'} \cap \left(\frac{N+1}{4}, N+1\right], \quad 1 \le k' \le K, \quad 1 \le d' \le c_1(K \log K)^5, \quad 1 \le h' \le h_0,$$

and

$$(m'+N+1)d' \le (5h'-4d')(N+1) \le (5h'-4)(N+1) \le (5h_0-4)(N+1).$$

Since $(m+5N)d + lN/4 \ge 5Nd + lN/4$, letting $l = l(N) = 20h_0 - 20d - 15$, it follows that the intervals $I(N) = [(m+N)d, (m+5N)d + lN/4]$, N sufficiently large, where m, d depends on N, overlap. Thus any large integer is a monochromatic sum in terms of partition (3.2) of at most $25h_0 = c_3(K \log K)^5$ squares and we are done. $\square$

We now turn to obtain a lower bound of $\mathrm{ord}_K(\mathcal{Q})$.

Theorem 3.9 (Hegyvári-Hennecart [HH07]). Let K be an integer. Then

$$\mathrm{ord}_K(\mathcal{Q}) \ge K \exp\left((\log 2 + o(1)) \frac{\log K}{\log \log K}\right).$$

Proof. For any $s \ge 2$, let $M_s = p_1 p_2 p_3 \dots p_s$ where $p_1 < p_2 < p_3 < \cdots$ is the increasing sequence of prime numbers. We denote by R the set of all non-zero quadratic residues modulo M_s. Then

$$|R| = \frac{p_1 - 1}{2} \cdot \frac{p_2 - 1}{2} \cdot \dots \cdot \frac{p_s - 1}{2}.$$

Let us consider the following partition of the squares:

$$\mathcal{Q} = \bigcup_{j=1}^{s} \{m^2 \,:\, (m, M_{j-1}) = 1 \text{ and } p_j \mid m\} \cup \bigcup_{m \in R} \mathcal{Q} \cap (m + \mathbb{N}M_s).$$

This a K_s-partition with

$$K_s = s + \frac{p_1 - 1}{2} \cdot \frac{p_2 - 1}{2} \cdot \cdots \cdot \frac{p_s - 1}{2}.$$

Let n be a large square free multiple of M_s. If h is such that $h(m+qM_s) = n$ for some $m \in R$, then $M_s \mid h$. This yields $h \geq M_s$. We obtain

$$\mathrm{ord}_{K_s}(\mathcal{Q}) \geq M_s.$$

Now let $K \geq 2$ be an integer. Then there is an $s \geq 1$ such that $K_s \leq K < K_{s+1}$. Since $(\mathrm{ord}_K(\mathcal{Q}))_{K \geq 1}$ is not decreasing, we get

$$\mathrm{ord}_K(\mathcal{Q}) \geq \mathrm{ord}_{K_s}(\mathcal{Q}) \geq M_s = \frac{M_{s+1}}{p_{s+1}} \geq \frac{2^{s+1}K_{s+1}}{p_{s+1}} > \frac{2^{s+1}K}{p_{s+1}}.$$

Classic asymptotic estimates on the primes give

$$s + 1 = (1 + o(1)) \frac{\log K}{\log \log K} \text{ and } p_{s+1} = e^{(1+o(1)) \log s} = e^{(1+o(1)) \log \log K}.$$

□

3.2.2 The primes

First we give un upper bound:

Theorem 3.10. [Hegyvári-Hennecart [HH07]] Let K be an integer. Then

$$\mathrm{ord}_K(\mathcal{P}) \leq 1500K^3.$$

Proof. We need two lemmas:

Lemma 3.11. Let N be a large integer. Then for any $n \le N$, we have

$$r_{\mathcal{P}}^{(3)}(n) \ll \frac{N^2}{(\log N)^3}.$$

where $r_{\mathcal{P}}^{(3)}(n)$the number of representations of an integer as a sum of 3 primes.

and

Lemma 3.12. Let N be a large integer. Then

$$E(\mathcal{P} \cap (N/2, N]) \le \frac{N^3}{5(\log N)^4} \tag{3.4}$$

where $E(\cdot)$ as usual denotes the energy.

The proofs of the lemmas can be found in [MBN] (using different terminology).

Let

$$\mathcal{P} = \bigcup_{k=1}^{K} \mathcal{P}_k, \tag{3.5}$$

be a partition of the primes. By the prime number theorem, since $20^{1/4} > 2$, we can find an integer N_0 such that for any $N \ge N_0$, both (3.4) and

$$\pi(N) - \pi(N/2) \ge \frac{N}{20^{1/4} \log N} \tag{3.6}$$

are satisfied. Let $N \ge N_0$ and put

$$\mathcal{S}_k = \mathcal{P}_k \cap (N/2, N], \qquad k = 1, 2, \ldots, K.$$

For any $k = 1, \ldots, K$, we have by Cauchy-Schwarz inequality,

$$|\mathcal{S}_k|^4 \le |2\mathcal{S}_k| E(\mathcal{S}_k),$$

thus there exists k such that

$$\begin{aligned} |2\mathcal{S}_k| &\ge \frac{|\mathcal{S}_1|^4 + \cdots + |\mathcal{S}_K|^4}{E(\mathcal{S}_1) + \cdots + E(\mathcal{S}_K)} \\ &\ge \frac{|\mathcal{S}_1|^4 + \cdots + |\mathcal{S}_K|^4}{E(\mathcal{P} \cap (N/2, N])}. \end{aligned}$$

dc_1260_16

By Hölder inequality we get

$$|2\mathcal{S}_k| \geq \frac{(|\mathcal{S}_1| + \cdots + |\mathcal{S}_K|)^4}{K^3 E(\mathcal{P} \cap (N/2, N])} = \frac{(\pi(N) - \pi(N/2))^4}{K^3 E(\mathcal{P} \cap (N/2, N])}$$

giving by Lemma 3.12 and (3.6)

$$|2\mathcal{S}_k| \geq \frac{N}{4K^3}.$$

We put $\mathcal{S} = \mathcal{S}_k$. Since $2\mathcal{S} \subset (N, 2N]$, applying Lemma 3.6 to $2\mathcal{S} - N$ shows for N large enough that there exists an integer d with $1 \leq d \leq 4K^3$ such that for some

$$h \leq h_0 = 500K^3, \tag{3.7}$$

we have

$$hN + \{(m+1)d, (m+2)d, \ldots, (m+2N)d\} \subset 2h\mathcal{S},$$

for some m such that $(m+2N)d \leq hN$. Since $\mathcal{S}$ contains at least two primes, we can find a prime p in $\mathcal{S}$ which is coprime to d. Thus the following interval of consecutive integers

$$hN + \{(m+1)d + (d-1)N, (m+2)d + (d-1)N + 1, \ldots, (m+2N)d\}$$

is contained in $\bigcup_{j=2h}^{2h+d-1} j\mathcal{S}$. Now shifting this interval by successive multiples of some arbitrary element $p \in \mathcal{S}$, we get

$$hN + [(m+N)d, (m+2N)d + lN] \subset \bigcup_{j=2h}^{2h+d-1+2l} j\mathcal{S}.$$

Applying this with $N+1$ instead of N, we get for any $l' \geq 0$,

$$h'(N+1) + [(m'+N+1)d', (m'+2N+2)d' + l'(N+1)] \subset \bigcup_{j=2h'}^{2h'+d-1+2l'} j\mathcal{S}',$$

where

$$\mathcal{S}' = \mathcal{P}_{k'} \cap ((N+1)/2, N+1], \quad 1 \leq k' \leq K, \quad 1 \leq d' \leq 4K^3, \quad 1 \leq h' \leq h_0,$$

and

$$h'(N+1)+(m'+N+1)d' \le (2h'-d')(N+1) \le (2h_0-1)(N+1).$$

Since $hN+(m+2N)d+lN \ge (h+l+2d)N$, we get for $l = 2h_0-2d-h$

$$h'(N+1)+(m'+N+1)d' \le hN+(m+2N)d+lN.$$

It follows that we can cover all sufficiently large integers by sums of at most $3h_0$ monochromatic sums of primes, according to the given partition (3.5). and by (3.7) we proved the theorem.

Now we turn to the lower bound.

Theorem 3.13 (Hegyvári-Hennecart [HH07])**.** Let K be an integer. Then

$$\mathrm{ord}_K(\mathcal{P}) \ge (e^{\gamma}+o(1))K\log\log K.$$

Proof. Let us consider the partition

$$\mathcal{P} = \{p \in \mathcal{P} \,:\, p \mid M\} \cup \bigcup_{\substack{m=1\\(m,M)=1}}^{M} \mathcal{P}\cap(m+\mathbb{N}M)$$

$(M \ge 1)$ and the colouring classes induced by it. This is a K-partition with

$$K = 1+\varphi(M),$$

where φ is the Euler's totient function. Let us count the minimal number of monochromatic summands needed to represent a large positive integer n congruent to 0 modulo M: it is clearly sufficient to consider the chromatic classes $\mathcal{P}\cap(m+\mathbb{N}M)$, where $(m,M)=1$. Obviously any integer h such that $h(m+qM)=n$ for some m coprime to M and some $q \ge 0$ must satisfy $M \mid h$. Thus

$$\mathrm{ord}_{1+\varphi(M)}(\mathcal{P}) \ge M. \tag{3.8}$$

Now let $K \ge 2$ be any integer. Let the sequence $(M_s)_{s\ge1}$ be defined as in the previous section. There exists an $s \ge 1$ such that $1+\varphi(M_s) \le K < 1+\varphi(M_{s+1})$, or equivalently

$$p_s - 1 \le \frac{K-1}{\varphi(M_{s-1})} < (p_s-1)(p_{s+1}-1).$$

Let λ be the integral part of $\frac{K-1}{\varphi(M_{s-1})}$. Observe that $\lambda \geq p_s - 1$. We thus have

$$(\lambda+1)\varphi(M_{s-1}) > K-1 \geq \lambda\varphi(M_{s-1}) \geq \varphi(\lambda M_{s-1}).$$

Since the sequence $(\mathrm{ord}_K(\mathcal{P}))_{K\geq 1}$ is non decreasing, we deduce from (3.8)

$$\mathrm{ord}_K(\mathcal{P}) \geq \mathrm{ord}_{1+\varphi(\lambda M_{s-1})}(\mathcal{P}) \geq \lambda M_{s-1} = \left(\frac{\lambda}{\lambda+1}\right)\frac{(\lambda+1)\varphi(M_{s-1})}{\prod\limits_{p|M_{s-1}}\left(1-\frac{1}{p}\right)}$$
$$> \left(\frac{p_s-1}{p_s}\right)\frac{K-1}{\prod\limits_{p|M_{s-1}}\left(1-\frac{1}{p}\right)}.$$

From Mertens' formula, we obtain

$$\prod_{p|M_{s-1}}\left(1-\frac{1}{p}\right) = \frac{e^{-\gamma}+o(1)}{\log s} = \frac{e^{-\gamma}+o(1)}{\log\log K},$$

by using the estimate

$$\log K = (1+o(1))\log\varphi(M_s) = (1+o(1))\log M_s = (1+o(1))p_s = (1+o(1))s\log s,$$

deduced from the prime number theorem. □

Remark 3.14. 1. At the proof of Theorem 3.13 we us a similar approach what used Sárközy having a lower bound for the order as additive basis of a dense set of primes.

2. Akhilesh, Ramana and Ramaré and Guohua Chen improved the bounds both in the prime as well as the square case. see [AR14], [RR12] and [Ch16].

□

Chapter 4

Restricted addition and related results

Recall some notation which will be necessary in this chapter:

For $h \geq 1$, we will use the following notation for addition and restricted addition: $h\mathcal{A}$ will denote the set of sums of h not necessarily distinct elements of $\mathcal{A}$, and $h \times \mathcal{A}$, the set of sums of h pairwise distinct elements of $\mathcal{A}$.

In this sense for an infinite set of integers $A \subseteq \mathbb{N}$, the set of subset sums can be perform as $FS(A) = \cup_{h \geq q}(h \times \mathcal{A}) \cup \{0\}$ (zero comes form the empty set).

4.1 On a problem of Burr and Erdős

In [E], Erdős writes:

> *Here is a really recent problem of Burr and myself : An infinite sequence of integers $a_1 < a_2 < \cdots$ is called an asymptotic basis of order k, if every large integer is the sum of k or fewer of the a's. Let now $b_1 < b_2 < \cdots$ be the sequence of integers which is the sum of k or fewer distinct a's. Is it true that*
>
> $$\limsup(b_{i+1} - b_i) < \infty.$$

In other words the gaps between the b's are bounded. The bound may of course depend on k and on the sequence $a_1 < a_2 < \cdots$.

If $\mathcal{A}$ is an increasing sequence of integers $a_1 < a_2 < \cdots$, the largest asymptotic gap in $\mathcal{A}$, that is

$$\limsup_{i\to+\infty}(a_{i+1} - a_i),$$

is denoted by $\Delta(\mathcal{A})$.

The question of Burr and Erdős takes the shorter form: is it true that if $h(\{0\} \cup \mathcal{A}) \sim \mathbb{N}$, then

$$\Delta(\mathcal{A} \cup 2 \times \mathcal{A} \cup \cdots \cup h \times \mathcal{A}) < +\infty\,?$$

In the following theorem we disprove this conjecture (except if $h = 2$):

Theorem 4.1 (Hegyvári-Hennecart-Plagne [HHP])**.** (i) *If* $(\mathcal{A} \cup 2\mathcal{A}) \sim \mathbb{N}$ *then*

$$\Delta(\mathcal{A} \cup 2 \times \mathcal{A}) \leq 2.$$

If $2\mathcal{A} \sim \mathbb{N}$ *then* $\Delta(2 \times \mathcal{A}) \leq 2$*.*

(ii) *Let* $h \geq 3$*. There exists a set* $\mathcal{A}$ *such that* $h(\{0\} \cup \mathcal{A}) \sim \mathbb{N}$ *and*

$$\Delta(\mathcal{A} \cup 2 \times \mathcal{A} \cup \cdots \cup h \times \mathcal{A}) = +\infty.$$

There exists a set $\mathcal{A}$ *such that* $h\mathcal{A} \sim \mathbb{N}$ *and* $\Delta(h \times \mathcal{A}) = +\infty$*.*

Proof. Let us first consider the case $h = 2$. Clearly the odd elements in $2\mathcal{A}$ do belong to $2 \times \mathcal{A}$. This implies that if $2\mathcal{A} \sim \mathbb{N}$, then $\Delta(2 \times \mathcal{A}) \leq 2$. This also implies that the odd elements in $\mathcal{A} \cup 2\mathcal{A}$ are in $\mathcal{A} \cup (2 \times \mathcal{A})$. It follows that $\mathcal{A} \cup 2\mathcal{A} \sim \mathbb{N}$ implies $\Delta(\mathcal{A} \cup (2 \times \mathcal{A})) \leq 2$.

In the case $h \geq 3$, it is enough to construct an explicit example. We first introduce the sequence defined by $x_0 = h$ and $x_{n+1} = (3 \cdot 2^{h-2} - 1)x_n^2 + hx_n$ for $n \geq 0$, and let

$$\mathcal{A}_n = [0, x_n^2) \cup \left\{2^j x_n^2 \,:\, j = 0, 1, 2, \ldots, h-2\right\}.$$

Finally we define

$$\mathcal{A} = \{0\} \cup \bigcup_{n\geq 0} (x_n + \mathcal{A}_n)\,.$$

Since any positive integer less than or equal to $2^{h-1}-2$ can be written as a sum of at most $h-2$ (distinct) powers of 2 taken from $\{2^j : j = 0,1,\dots,h-2\}$, any integer in $[0,(2^{h-1}-1)x_n^2)$ can be written as a sum of $h-1$ elements of $\mathcal{A}_n$. Thus it follows

$$[0,(3\cdot 2^{h-2}-1)x_n^2) \subset \{0,2^{h-2}x_n^2\}+[0,(2^{h-1}-1)x_n^2) \subset \{0,2^{h-2}x_n^2\}+(h-1)\mathcal{A}_n \subset h\mathcal{A}_n.$$

We therefore infer that $[hx_n,x_{n+1}) \subset h(x_n+\mathcal{A}_n)$. Moreover, since $hx_n \le x_n^2$, we have $[x_n,hx_n] \subset [x_n,x_n^2] \subset x_n+\mathcal{A}_n$. It follows that, for any $n \ge 0$, we have

$$[x_n,x_{n+1}) \subset h\big((x_n+\mathcal{A}_n)\cup\{0\}\big) \subset h\mathcal{A}.$$

Consequently $h\mathcal{A} \sim \mathbb{N}$.

On the other hand, $(h-1)\mathcal{A} \not\sim \mathbb{N}$. Indeed, this assertion follows from the more precise fact that, for any $n \ge 0$, no integer in the range $[(2^{h-1}-1)x_n^2 + (h-1)x_n+1, 2^{h-1}x_n^2-1]$ (an interval of integers with a length tending to infinity with n) can be written as a sum of $h-1$ elements of $\mathcal{A}$. Let us prove this fact by contradiction and assume the existence of an integer

$$u \in [(2^{h-1}-1)x_n^2+(h-1)x_n+1, 2^{h-1}x_n^2-1] \cap (h-1)\mathcal{A}.$$

Since we have (using $h \ge 3$)

$$u \le 2^{h-1}x_n^2 - 1 < x_{n+1},$$

we deduce that

$$\begin{aligned} u &\in (h-1)\left(\{0\}\cup\bigcup_{i=0}^{n}(x_i+\mathcal{A}_i)\right) \\ &\subset (h-1)\big([0,x_n+x_n^2]\cup\big\{2^j x_n^2+x_n : j=1,2,\dots,h-2\big\}\big). \end{aligned}$$

In other words, we can express u as a sum of the form

$$\begin{aligned} u &= \alpha_{h-2}\left(2^{h-2}x_n^2+x_n\right)+\cdots+\alpha_1\left(2x_n^2+x_n\right)+\rho\left(x_n+x_n^2\right) \\ &= \left(2^{h-2}\alpha_{h-2}+\cdots+2\alpha_1+\rho\right)x_n^2+\left(\alpha_{h-2}+\cdots+\alpha_1+\rho\right)x_n, \end{aligned}$$

with $\alpha_1,\dots,\alpha_{h-2} \in \mathbb{N}$, ρ a positive real number and

$$\alpha_{h-2}+\cdots+\alpha_1+\rho \le h-1.$$

If we denote by $[\rho]$ the integral part of ρ, this implies that

$$\left(2^{h-2}\alpha_{h-2}+\cdots+2\alpha_1+[\rho]\right)x_n^2 \leq u \leq \left(2^{h-2}\alpha_{h-2}+\cdots+2\alpha_1+\rho\right)x_n^2+(h-1)x_n$$

and in view of $u \in [(2^{h-1}-1)x_n^2+(h-1)x_n+1, 2^{h-1}x_n^2-1]$, we deduce that

$$2^{h-2}\alpha_{h-2}+\cdots+2\alpha_1+[\rho] \leq 2^{h-1}-1$$

and

$$2^{h-2}\alpha_{h-2}+\cdots+2\alpha_1+\rho \geq 2^{h-1}-1.$$

We therefore obtain $2^{h-2}\alpha_{h-2}+\cdots+2\alpha_1+[\rho]=2^{h-1}-1$. We conclude by the facts that $\alpha_{h-2}+\cdots+\alpha_1+[\rho] \leq h-1$ and that the only decomposition of $2^{h-1}-1$ as a sum of at most $h-1$ powers of 2 is $2^{h-1}-1=1+2+2^2+\cdots+2^{h-2}$ that $\alpha_1=\cdots=\alpha_{h-2}=[\rho]=1$. From this, we deduce that $\rho \leq h-1-\alpha_1-\cdots-\alpha_{h-2}=1$ and finally $\rho=1$ which gives $u=(2^{h-1}-1)x_n^2+(h-1)x_n$, a contradiction. Since $h\mathcal{A} \sim \mathbb{N}$, we deduce that $\mathcal{A}$ is an asymptotic basis of order h.

Concerning restricted addition, we see that for $l \geq h-2$, we have

$$\max(l \times \mathcal{A}_n) \leq (2^{h-1}-2)x_n^2+(l-h+2)x_n^2=(2^{h-1}+l-h)x_n^2.$$

Hence

$$x_{n+1}-\max\left(l \times (x_n+\mathcal{A}_n)\right) \geq (2^{h-2}-l+h-1)x_n^2+(h-l)x_n.$$

If $l \leq 2^{h-2}+h-2$, then $x_{n+1}-\max\left(l \times (x_n+\mathcal{A}_n)\right) \geq x_n^2-(2^{h-2}-2)x_n$ which tends to infinity as n tends to infinity. □

Having Theorem 4.1 at hand, the next natural question is then: assume that $h\mathcal{A} \sim \mathbb{N}$, that is $h\mathcal{A}$ contains all but finitely many positive integers, is it true that there exists an integer k such that $\Delta(k \times \mathcal{A}) < +\infty$? If so, k could depend on $\mathcal{A}$. But, suppose that such a k exists for all $\mathcal{A}$ satisfying $h\mathcal{A} \sim \mathbb{N}$: is this value of k uniformly (with respect to $\mathcal{A}$) bounded from above (in term of h)? If so, write $k(h)$ for the maximal possible value:

$$k(h)=\max_{h\mathcal{A}\sim\mathbb{N}} \min\{k \in \mathbb{N} \text{ such that } \Delta(k \times \mathcal{A}) \text{ is finite}\}.$$

Theorem 4.1 implies that $k(2)$ does exist and is equal to 2. No other value of $k(h)$ is known but we believe that the following conjecture is true.

Conjecture 4.2. The function $k(h)$ is well-defined in the sense that for any integer $h \geq 1$, $k(h)$ is finite.

One can read from the proof of Theorem 4.1 that if for every h, $k(h)$ exists, then

Theorem 4.3. *Let $h \geq 2$. We have*

$$k(h) \geq 2^{h-2} + h - 1.$$

According to what obviously happens in the case of usual addition, it would be of some interest to establish, for any given set of integers $\mathcal{A}$, the monotonicity of the sequence $\big(\Delta(h \times \mathcal{A})\big)_{h \geq 1}$:

Conjecture 4.4. Let $\mathcal{A}$ be a set of positive integers, then the sequence $\big(\Delta(h \times \mathcal{A})\big)_{h \geq 1}$ is non-increasing.

More interestingly, we will show the following partial result in the direction of Conjecture 4.4:

Theorem 4.5 (Hegyvári-Hennecart-Plagne)**.** *Let $\mathcal{A}$ be a set of positive integers and h be the smallest positive integer such that $\Delta(h \times \mathcal{A})$ is finite. Then there exists an increasing sequence of integers $(h_j)_{j \geq 0}$ with $h_0 = h$ such that for any $j \geq 1$, one has $h_j + 2 \leq h_{j+1} \leq h_j + h + 1$ and $\Delta(h_{j+1} \times \mathcal{A}) \leq \Delta(h_j \times \mathcal{A})$.*

This shows that for a given set of positive integers $\mathcal{A}$, the inequality $\Delta\big((h+1) \times \mathcal{A}\big) \leq \Delta(h \times \mathcal{A})$ holds for any h belonging to some set of positive integers having a positive lower asymptotic density bounded from below by $1/(h+1)$.

Remark 4.6. At the proof of Theorem 4.5 we will use a combinatorial lemma, called "sunflower lemma". Recently this lemma is frequently used at additive problems. In my knowledge before us only Erdős and Sárközy used it to solve (rather a different) problem.

Proof of Theorem 4.5. Let $\mathcal{A}$ be such that $d = \Delta(h \times \mathcal{A}) < +\infty$. This implies that for any sufficiently large x,

$$A(x) = |\mathcal{A} \cap [1, x]| \geq Cx^{1/h},$$

for some positive constant C depending only on d. Now, the number of subsets of $\mathcal{A} \cap [1, x]$ with cardinality $h+1$ is equal to the binomial coefficient $\binom{A(x)}{h+1} \gg x^{1+1/h}$ where the implied constant depends on both $\mathcal{A}$ and h. Choose an x such that $\binom{A(x)}{h+1} \geq (h+2)!\, h^{h+2} x$. It thus exists an integer n less than $(h+1)x$ such that

$$n = a_1^{(i)} + \cdots + a_{h+1}^{(i)}, \qquad \text{for } i = 1, \ldots, (h+1)!\, h^{h+2},$$

where the $(h+1)!\, h^{h+2}$ sets $E_i = \{a_1^{(i)}, \ldots, a_{h+1}^{(i)}\}$ of $h+1$ pairwise distinct elements of $\mathcal{A}$ are distinct. We now make use of the following intersection theorem for systems of sets due to Erdős and Rado (cf. Theorem III of [**?**]):

Lemma 4.7 (Erdős-Rado)**.** Let m, q, r be positive integers and E_i, $1 \leq i \leq m$, be sets of cardinality at most r. If $m \geq r!\, q^{r+1}$, then there exist an increasing sequence $i_1 < i_2 < \cdots < i_{q+1}$ and a set F such that $E_{i_j} \cap E_{i_k} = F$ as soon as $1 \leq j < k \leq q+1$.

By applying this result with $q = h$ and $r = h+1$, we obtain that there are $h+1$ sets E_{i_j}, $j = 1, \ldots, h+1$, and a set F, with $0 \leq |F| \leq h-1$, such that $E_{i_j} \cap E_{i_k} = F$ if $1 \leq j \neq k \leq h+1$. Observe that we must have $0 \leq |F| \leq h-1$ since the E_i's are distinct and the sum of all elements of E_i is equal to n for any i. We obtain that the integer

$$n' = n - \sum_{a \in F} a$$

can be written as a sum of $h+1-|F|$ pairwise distinct elements of $\mathcal{A}$ in at least $h+1$ ways, such that all summands occurring in any of these representations of n' in $(h+1-|F|) \times \mathcal{A}$ are pairwise distinct (equivalently, this means that the set $\cup_{j=1}^{h+1} E_{i_j} \smallsetminus F$ has exactly $(h+1)(h+1-|F|)$ distinct elements). This shows that

$$n' + (h \times \mathcal{A}) \subset (2h+1-|F|) \times \mathcal{A},$$

and finally $\Delta(h \times \mathcal{A}) = \Delta\big(n' + (h \times \mathcal{A})\big) \geq \Delta(h_1 \times \mathcal{A})$, where $h_1 = 2h+1-|F|$.

Iterating this process, we get an increasing sequence $(h_j)_{j \geq 0}$, with $h_0 = h$, such that

$$\Delta(h_j \times \mathcal{A}) = \Delta\big(n' + (h_j \times \mathcal{A})\big) \geq \Delta(h_{j+1} \times \mathcal{A}),$$

where h_{j+1} is of the form $h_j + h + 1 - |F_j|$ for some set F_j satisfying $0 \leq |F_j| \leq h-1$. We conclude that $h_j + 2 \leq h_{j+1} \leq h_j + h + 1$, as stated. $\square$

To finish this section we mention two related problems. We quote from the book [ERG] written by Erdős and Graham where two of these conjectures are explicitly stated (page 52): Is it true that if $\mathrm{ord}(A) = r$, then $r \times A$ has positive (lower) density? If $\overline{d}(sA) > 0$ then must $s \times A$ also have positive upper density?

In [HHP2] we gave an affirmative answer to these questions. We prove a more general theorem (which is also valid, but with no interest, if $\underline{d}(hA) = 0$)

Theorem 4.8 (Hegyvári-Hennecart-Plagne)**.** *Let $\mathcal{A}$ be a set of positive integers such that for some integer h $\underline{d}(hA) > 0$ then*

$$\underline{d}(h \times A) \geq \frac{1}{h^h \exp(\pi\sqrt{2h/3})}\underline{d}(hA).$$

See the proof in [HHP2].

Finally we will show that a relative small part of $2\mathcal{A}$ lies outside of $2 \times \mathcal{A}$:

Theorem 4.9 (Hegyvári-Hennecart-Plagne)**.** *For any finite set $\mathcal{A}$ of non negative integers with $|\mathcal{A}| \geq 2$, one has*

$$|(2\mathcal{A}) \setminus (2 \times \mathcal{A})| \ll \frac{|2\mathcal{A}|(\log\log|2\mathcal{A}|)^{5/4}}{(\log|2\mathcal{A}|)^{1/4}}.$$

Proof of Theorem 4.9. Let $\mathcal{B}$ be the subset of $\mathcal{A}$ defined as

$$\mathcal{B} = \{a \in \mathcal{A} \text{ such that } 2a \notin 2 \times \mathcal{A}\}.$$

We let

$$B = |\mathcal{B}| = |(2\mathcal{A}) \setminus (2 \times \mathcal{A})| = |2\mathcal{A}| - |2 \times \mathcal{A}|.$$

Clearly $\mathcal{B}$ does not contain any non trivial triple in arithmetic progression, because if $a + b = 2c$ with $a, b, c \in \mathcal{B}$ and $a \neq b$ then $2c = a + b \in 2 \times \mathcal{A}$ contrary to the fact that c is in $\mathcal{B}$. Thus we may apply the following lemma:

Lemma 4.10. Let $\mathcal{A}$ be a finite set of non negative integers of cardinality n. If $\mathcal{A}$ does not contain any 3-term arithmetic progression, then

$$|2\mathcal{A}| \geq \frac{n^{5/4}}{2r_3(n)^{1/4}}.$$

See the proof in [Ru].

So by the lemma we obtain

$$|2\mathcal{B}| \geq \frac{B^{5/4}}{2r_3(B)^{1/4}}.$$

Now, since $\mathcal{A}$ contains $\mathcal{B}$, we have

$$|2\mathcal{A}| \geq \frac{B^{5/4}}{2r_3(B)^{1/4}}$$

and finally, by a result of Sanders [San],

$$|2\mathcal{A}| \geq \kappa_1 B \frac{\log|\mathcal{B}|^{1/4}}{(\log\log B)^{5/4}}$$

for some positive constant κ_1. Clearly this lower bound implies

$$|(2\mathcal{A}) \setminus (2 \times \mathcal{A})| = B \leq \kappa \frac{|2\mathcal{A}|(\log\log|2\mathcal{A}|)^{1/4}}{\log|2\mathcal{A}|^{1/4}}.$$

for some constant κ.

$\square$

Remark 4.11. 1. The theorem above was independently proved by T. Schoen [So]

2. Originally in [HHP2] we gave the bound

$$|(2\mathcal{A}) \setminus (2 \times \mathcal{A})| \leq \kappa \frac{|2\mathcal{A}|}{(\log\log|2\mathcal{A}|)^{1/4}}.$$

using Roth's result [Roth]. (Sanders' result is later than our theorem).

4.2 On complete sequences

A set $A \subseteq \mathbb{N}$ is said to be *complete* if there exists a threshold number n_0 such that every natural number greater than n_0 is the sum of distinct terms taken from A. This concept was introduced by Erdős in the 60's. The simplest example for a complete set is the powers of two: $\{2^n : n = 0, 1, \dots\}$ where clearly the threshold is $n_0 = 0$. An infinite subset $A \subseteq \mathbb{N}$ is called *subcomplete* if there is an infinite arithmetic progression in $FS(A)$.

In the literature there are many interesting results: K.F. Roth and Gy. Szekeres proved that if $f(x) \in \mathbb{R}[x]$ and f maps integers to integers then the set $F := \{f(n) : n \in \mathbb{N}\}$ is complete if and only if for any prime p there is an integer k such that p does not divide $f(k)$, and the leading coefficient of $f(x)$ is positive. (They used ingenious analytic techniques).

There are many generalization of it; e.g. S. Burr investigated some perturbation of values of F. He proved that the set $F' := \{s_n = f(n) + O(n^\alpha) : n \in \mathbb{N}\}$ is subcomplete, provided the values of F' are positive integers, f is non-constant, and $0 < \alpha < 1$.

For thinner sequence I quote here a theorem of Zeckendorf who proved that every positive integer N has a unique representation as the sum of non-consecutive Fibonacci numbers.

Many other problems and results can be found in [ERG], section "Complete Sequences".

In the next two sections I discuss two earlier results of mine.

4.2.1 Completeness of thin sequences

A natural question of Erdős asked how dense a sequence A which is subcomplete has to be. He conjectured that $a_{n+1}/a_n \to 1$ (as $n \to \infty$) implies the subcompleteness. However in 1960 J. W. S. Cassels (cf. [Ca]) showed that for every $\varepsilon > 0$ there exists a sequence A for which

$$a_{n+1} - a_n = o(a_n^{1/2+\varepsilon})$$

and A is not complete.

In 1962 Erdős proved the following theorem:

Theorem 4.12 (Erdős)**.** Let $A \subseteq \mathbb{N}$ be an infinite increasing sequence of integers, for which

$$A(n) > Cn^{\frac{\sqrt{5}-1}{2}}$$

$(C > 0)$. Then A is subcomplete

A couple of years later J. Folkman improved it to $A(n) > n^{1/2+\varepsilon}$ ($\varepsilon > 0$; $n > n_0(\varepsilon)$).

In 2000 I arrived very close to the best. I proved

Theorem 4.13 (Hegyvári)**.** Let $A = \{0 < a_1 < a_2 < \dots\}$ be an infinite sequence of integers. Assume that

$$A(n) > 300\sqrt{n \log n}$$

for $n > n_0$. Then A is subcomplete.

We mention here that $300\sqrt{n \log n}$ cannot be replaced by $\sqrt{2n}$; it is easy to construct a sequence A for which $A(n) > \sqrt{2n}$ and A is not subcomplete.

In the same year a little bit later and independently Łuczak and Schoen [LS00] also proved essentially in the same way this theorem.

The proofs based on a theorem of Sárközy (which theorem was also proved independently by Freiman using Fourier analysis).

The proof of Theorem 4.13 can be found in [He00] and in ([TV] p.480; p.482).

Finally in 2006 Szemerédi and Vu [SzV] could complete the problem of Erdős; apart from the constant factor they proved the conjecture:

Theorem 4.14 (Szemerédi-Vu)**.** Let $A = \{0 < a_1 < a_2 < \dots\}$ be an infinite sequence of integers. Assume that

$$A(n) \gg \sqrt{n}$$

Then A is subcomplete.

As they wrote there: *"The proof presented here combines arguments from Hegyvári's paper [11] and new ideas..."*

([11]=[HE00]).

4.2.2 Completeness of exponential type sequences

As we mentioned the simplest example for a complete set is the powers of two: $\{2^n : n = 0, 1, \dots\}$ where clearly the threshold is $n_0 = 0$ and furthermore the set $S_p = \{p^n : n = 0, 1, \dots\}; p \in \mathbb{N}$ is complete if and only if $p = 2$. An easy counting argument shows that if a set A is complete, than

$$A(n) := \sum_{a \in A; a \leq n} 1 > \log_2 n - t_A, \tag{4.1}$$

for some t_A. Thus it is reasonable to ask on a slightly denser sequence. A longstanding question of Erdős was strengthen by J. Birch in 1959 who proved that the sequence $S_{p,q} = \{p^n q^m : n, m = 0, 1, \dots\}$, $(1 < p, q \in \mathbb{N}\ (p, q) = 1)$ is complete (see in [Bi]).

A few years later Cassels proved in [Ca] a more general theorem which includes the Birch's theorem.

Theorem 4.15 (Cassels,1960)**.** Let $A \subseteq \mathbb{N}$ and assume that

$$\lim_{n \to \infty} \frac{A(2n) - A(n)}{\log \log n} = \infty,$$

and for every real number θ, $(0 < \theta < 1)$ $\sum_{i=1}^{\infty} \|a_i \theta\| = \infty$.

Then the sequence A is complete.

Later H. Davenport remarked that there is a stronger version of Erdős' conjecture which is not covered by Cassels' theorem: considering (4.1) there should be a threshold $K = K(p, q)$ for which the set $S_{p,q}(K) = \{p^n q^m : n = 0, 1, \dots\ 0 \leq m \leq K\}$ will be complete.

As Erdős wrote

"Of course the exact value of $K(p, q)$ is not known and no doubt will be very difficult to determine."

In [He00b] I gave a quantitative upper bound for the function $K(p, q)$

Theorem 4.16 (Hegyvári)**.** For every integers $p, q > 1$ and $(p, q) = 1$ there exists $K = K(p, q)$ such that the sequence $Y_K = \{p^n q^m : n = 0, 1, \dots\ 0 \leq m \leq K\}$ is complete. Moreover

$$K(p, q) \leq 2p^{2c^{2^{2q^{4p+3}}}}$$

where $c = 1152 \log_2 p \log_2 q$.

I should mention that my theorem has many improvements.

Firstly J. Fang – based on my idea and a result of V. Vu – could reduce one step of my proof obtaining

Theorem 4.17 (J. Fang [FG])**.**

$$K(p,q) \le p^{c^{2^{q^{2p+3}}}}$$

where $c = 1152 \log_2 p \log_2 q$.

Further improvements by Y-G. Chen and J. Fang

Theorem 4.18 (Y-G. Chen J. Fang [CFG])**.**

$$K(p,q) \le c^{2^{q^{2p+3}}}$$

where $c = 1152 \log_2 p \log_2 q$,

Very recently Bergelson and Simmons [BS] obtained the best bound, using a very deep theorem of Fürstenberg.

I close this section with two further results. The first is related to the recent result of Bergelson and Simmons and a question of Erdős and Graham [p. 53 in [ERG]). In this booklet they asked the following: Let $S(t,\alpha) = \{s_1, s_2, \dots, \}$ with $s_n = \lfloor t\alpha^n \rfloor$. For what values of t and α is $S(t,\alpha)$ complete? (As they wrote: There seems to be little hope of proving this at present since it is not even known what is the distribution of $\lfloor (3/2)^n \rfloor$.) In [HR] with G. Rauzy we prove

Theorem 4.19 (Hegyvári-Rauzy [HR])**.** Let $B = \{b_1 < b_2 < \dots\} \subset \mathbb{N}$, $\alpha > 0$. Then the set

$$\{b_m[2^n\alpha] : b_m \in B \ n \in \mathbb{N}\}$$

is subcomplete.

The second is related to the completeness of exponential type sequences. Chen-Fang and myself proved the following result: Let

$$S_p = \{p^s : s \ge 0; s \in \mathbb{N}\}$$

be the sequence of p powers, and let $F_0 = 0, F_1 = 1; \quad n > 1; \ F_n = F_{n-1} + F_{n-2}$ be the sequence of Fibonacci sequence. Finally denote by

$$\mathcal{F}_k(n) := \{F_k < F_{k+1} < \dots F_n\},$$

the k, n-*truncated* sequence of $\{F_i\}$, where $n > k$.

Theorem 4.20 (Chen-Fang-Hegyvári [CFH])**.** For any integer $p > 1$ and any integer $k \geq 1$, there exists an integer $n \leq p^2 F_{k+2p-2}^2 + pF_{k+2p-2}$ such that $S_p F_k(n)$ is complete.

The Theorem 4.13 and Theorem 4.16 are relatively old result of mine. These results can be found in papers which will be included as a supplements at the end of this work.

Chapter 5

Expanding and covering polynomials

5.1 Expanding polynomials

The well-known Cauchy-Davenport theorem states that for any pair of sets A, B in $\mathbb{Z}_p$ such that $A + B \neq \mathbb{Z}_p$, we have $|A + B| \geq |A| + |B| - 1$ and this estimation is sharp; for arithmetic progressions A, B with common difference yield $|A+B| = |A|+|B|-1$. Now a natural question arises; what can we say on the image of a two-variable (or more generally multivariable) polynomial. One can ask which polynomial f blows up its domain, i.e. if for any $A, B \subseteq \mathbb{Z}_p$, $|A| \asymp |B|$ then $f(A, B) := \{f(a, b) : a \in A; b \in B\}$ is ampler (in some uniform meaning) than $|A|$. As we remarked earlier, the polynomial $f(x, y) = x + y$ and similarly $g(x, y) = xy$ are not admissible.

Let us say that a polynomial $f(x, y)$ is an *expander* if $|f(A, B)|/|A|$ tends to infinity as p tends to infinity (a more precise definition will be given above).

Nevertheless by the well known sum-product estimation we know that one of them blows up its domain.

Theorem 5.1 (Bourgain-Katz-Tao)**.** Let $A \subset \mathbb{F}_p$ for which

$$p^{\delta} < |A| < p^{1-\delta}.$$

Then one has a bound of the form

$$\max\{|A + A|, |AA|\} \geq c(\delta)|A|^{1+\varepsilon}$$

for some $\varepsilon = \varepsilon(\delta) > 0$

see the proof in [BKT] (later results omitted the δ-restriction)

Remark 5.2. This theorem gives immediately a three-variable expanding polynomial. Indeed, we have two cases

When $|AA| \gg |A|^{1+\varepsilon}$, then for every element $a \in A$ we have

$$|AA + A| \geq |AA + a| \gg |A|^{1+\varepsilon}.$$

When $|A + A| \gg |A|^{1+\varepsilon}$, then again for every $b \in A$ ($b \neq 0$ by the well-known Plnnecke-Ruzsa's inequality we get

$$|bA + bA| \gg |A|^{1+\varepsilon}$$

hence

$$|A|^{1+\varepsilon} \ll |bA + bA| \leq \frac{|bA + A|^2}{|A|},$$

so we get $|AA + A| \gg |A|^{1+\varepsilon/2}$,

Thus by this remark the challenging question is to find two-variable expanding polynomials.

Definition 5.3. For any prime number p, let $F : \mathbb{F}_p^k \to \mathbb{F}_p$ be an arbitrary function in 2-variable in $\mathbb{F}_p$. This function is said to be *expander*, if for any α, $0 < \alpha < 1$, there exist $\epsilon = \epsilon(\alpha) > 0$ such that for any pair $A, B \in \mathbb{F}_p$ with

$$|A|, |B| \asymp p^{\alpha}$$

one has

$$|F(A, B)| \gg p^{\alpha+\epsilon},$$

where

$$F(A, B) = \{F(a, b) : a \in A;\ b \in B\}$$

It is reasonable to try the polynomials:

$$F_1(x, y) = f(x, y) + g(x, y),\ F_2(x, y) = f(x, y)/g(x, y),$$

$$F_3(x, y) = f(x, y) \cdot g(x, y),\ F_4(x, y) = f(g(x, y), y),\ F_5(x, y) = g(x, f(x, y)).$$

It is easy to see that $F_1(x,y)$ and $F_2(x,y)$ are not expander.

Indeed $F_1(x,y)$ can be written in the form $(x+1)(y+1)-1$. Thus if A and B are geometric sequences (with common quotient) -1, then F_1 does not blow up its domain. This observation leads us the following; in order to exhibit expanders of the type $f(x)+h(x)g(y)$ we thus have to assume that f and g are affinely independent,

Definition 5.4. Let $f(x) \in \mathbb{Z}[x]$ and $g(y) \in \mathbb{Z}[y]$. We say that f and g are affinely independent, if there is no $(u,v) \in \mathbb{Z}^2$ such that $f(x) = uh(x)+v$ or $h(x) = uf(x)+v$.

Indeed, if $F(x,y)$ is a polynomial in the form $F(x,y) = f(x) + (uf(x) + v)g(y)$ where $u,v \in \mathbb{F}_p$ and f, g are integral polynomials, then it is not expander.

It is clear if $u = 0$, since in this case

$$F(x,y) = f(x) + vg(y)$$

and – say – $A_d = \{a : dk = f(a) \text{ and } 1 \le k \le p/3\}$ and $B_d = \{b : dk = vg(b)\text{an } 1 \le k \le p/3\}$ $(d \neq 0)$, then they (and they sum) are arithmetic progressions.

If $u \neq 0$, then $F(x,y) = (f(x)+vu^{-1})(1+ug(y))-vu^{-1}$; and $(f(A)+vu^{-1})$ and $(1+ug(B))-vu^{-1}$ are geometric sequences (with common quotient), (i.e. A and B are "inverse image" of them) then $F(x,y)$ is not an expander.. In order to exhibit expanders of the type $f(x)+h(x)g(y)$, we thus have to assume that f and g are affinely independent, namely there is no $(u,v) \in \mathbb{Z}^2$ such that $f(x) = uh(x)+v$ or $h(x) = uf(x)+v$.

According to the literature the first known explicit construction is due to J. Bourgain (see [B]) who proved that the polynomial $F_5(x,y) = x^2 + xy$ is an expander. More precisely he proved that if $p^\varepsilon < |A| \asymp |B| < p^{1-\varepsilon}$ then $|f(A,B)|/|A| > p^\gamma$, where $\gamma = \gamma(\varepsilon)$ is a positive but inexplicit real number.

In my best knowledge in [HH09] we gave first explicitly an infinite class of expanding polynomials.

5.1.1 Infinite class of expanding polynomials in prime fields

The main tools what we will use, two Szemerédi-Trotter type inequalities:

Proposition 5.5 (Bourgain-Katz-Tao Theorem [BKT]). *Let $\mathcal{P}$ and $\mathcal{L}$ be respectively a set of points and a set of lines in $\mathbb{F}_p^2$ such that*

$$|\mathcal{P}|, |\mathcal{L}| < p^\beta$$

for some β, $0 < \beta < 2$. Then

$$|\{(P, L) \in \mathcal{P} \times \mathcal{L} \ : \ P \in L\}| \ll p^{(3/2-\gamma)\beta} \quad \textit{(as } p \textit{ tends to infinity)},$$

for some $\gamma > 0$ depending only on β.

and another inequality which gives explicit bound to the expanding measure:

Proposition 5.6 (L.A. Vinh [LAV]). *Let $d \geq 2$. Let $\mathcal{P}$ be a set of points in $\mathbb{F}_p^d$ and $\mathcal{H}$ be a set of hyperplanes in $\mathbb{F}_p^d$. Then*

$$|\{(P, H) \in \mathcal{P} \times \mathcal{H} \ : \ P \in H\}| \leq \frac{|\mathcal{P}||\mathcal{H}|}{p} + (1 + o(1))p^{(d-1)/2}(|\mathcal{P}||\mathcal{H}|)^{1/2}.$$

Now we can proof the following:

Theorem 5.7 (Hegyvári-Hennecart [HH09]). Let $k \geq 1$ be an integer and f, g be polynomials with integer coefficients, and define for any prime number p, the map F from $\mathbb{Z}^2$ onto $\mathbb{Z}$ by

$$F(x, y) = f(x) + x^k g(y)$$

Assume moreover that $f(x)$ is affinely independent to x^k. Then F induces an expander.

Proof. For p sufficiently large, the image $g(B)$ of any subset B of $\mathbb{F}_p$ has cardinality at least $|B|/\deg(g)$. It follows that we can restrict our attention to maps of the type $F(x, y) = f(x) + x^k y$. We let $d := \deg(f)$.

Let A and B be subsets of $\mathbb{F}_p$ with cardinality $|A| \asymp |B| \asymp p^\alpha$. For any $z \in \mathbb{F}_p$, we denote by $r(z)$ the number of couples $(x, y) \in A \times B$ such that $z = F(x, y)$, and by C the set of those z for which $r(z) > 0$. By Cauchy-Schwarz inequality, we get

$$|A|^2|B|^2 = \Big(\sum_{z \in \mathbb{F}_p} r(z)\Big)^2 \leq |C| \times \Big(\sum_{z \in \mathbb{F}_p} r(z)^2\Big).$$

One now deal with the sum $\sum_{z\in\mathbb{F}_p} r(z)^2$ which can be rewritten as the number of quadruples $(x_1, x_2, y_1, y_2) \in A^2 \times B^2$ such that

$$f(x_1) + x_1^k y_1 = f(x_2) + x_2^k y_2. \tag{5.1}$$

For fixed $(x_1, x_2) \in A^2$ with $x_1 \neq 0$ or $x_2 \neq 0$, (5.1) can be viewed as the equation of a line ℓ_{x_1,x_2} whose points (y_1, y_2) are in $\mathbb{F}_p^2$. For (x_1, x_2) and (a, b) in A^2, the lines ℓ_{x_1,x_2} and $\ell_{a,b}$ coincide if and only if

$$\begin{cases} (x_1 b)^k = (a x_2)^k \\ b^k(f(x_2) - f(x_1)) = x_2^k(f(b) - f(a)), \end{cases}$$

or equivalently

$$\begin{cases} (x_1 b)^k = (a x_2)^k \\ (b^k - a^k)(f(x_2) - f(x_1)) = (x_2^k - x_1^k)(f(b) - f(a)). \end{cases} \tag{5.2}$$

At this point observe that by our assumption, there are only finitely many prime numbers p such that $f(x) = ux^k + v$ for some $(u, v) \in \mathbb{F}_p^2$, in which case the second equation in (5.2) holds trivially for any x_1 and x_2. We assume in the sequel that p is not such a prime number.

Let $(a, b) \in A^2$ such that $a \neq 0$ or $b \neq 0$. Assume for instance that $b \neq 0$. By (5.2) we get $x_1 = \frac{\zeta a x_2}{b}$ for some k-th root modulo p of unity ζ. Moreover, we obtain

$$b^k\Big(f(x_2) - f(\zeta\frac{a x_2}{b})\Big) - x_2^k(f(b) - f(a)) = 0, \tag{5.3}$$

which is a polynomial equation in x_2. If we write $f(x) = \sum_{0\le j\le d} f_j x^j$ then

$$b^k(f(x) - f(\zeta\frac{ax}{b})) = \sum_{1\le j\le d} b^k(1 - \frac{\zeta^j a^j}{b^j}) f_j x^j$$

is a polynomial which could be identically equal to $x^k(f(b) - f(a))$ only if the following two conditions are satisfied:

$$f(b) - f(a) = (b^k - a^k) f_k,$$
$$f_j \neq 0 \Rightarrow b^j = \zeta^j a^j.$$

Since $f(x)$ is assumed to be affinely independent to x^k, we necessarily have $f_j \neq 0$ for some $0 < j \neq k$. If $b^j = \zeta^j a^j$ for ζ being a k-th root of unity in $\mathbb{F}_p$, then $b = \eta a$ where η is some $(kd!)$-root of unity in $\mathbb{F}_p$. Let

$$X := \{(a, b) \in A^2 \ : \ b^{kd!} \neq a^{kd!}\}.$$

Since there are $kd!$ many $(kd!)$-roots of unity in $\mathbb{F}_p$, We have $|A^2 \setminus X| \leq kd!|A|$, hence $|X| \geq \frac{|A|^2}{2}$ for p large enough.

If $(a, b) \in X$, then (5.3) has at most $\max(k, d)$ many solutions x_2, thus (5.2) has at most $k \max(k, d)$ many solutions (x_1, x_2). We conclude that the number of distinct lines $\ell_{a,b}$ when (a, b) runs in A^2 is $c(k, f)|A|^2$ where $c(k, f)$ can be chosen equal to $(2k \max(k, d))^{-1}$, for p large enough. The set of all these pairwise distinct lines $\ell_{a,b}$ is denoted by $\mathcal{L}$, its cardinality satisfies $|A|^2 \ll |\mathcal{L}| \leq |A|^2$, as observed before. Let $\mathcal{P} = B^2$. Then putting $N := |A|^2 \asymp |B|^2$, we have by Proposition 5.5

$$\big\{(p, \ell) \in \mathcal{P} \times \mathcal{L} \ : \ p \in \ell\big\} \ll N^{3/2-\delta}$$

for some $\delta > 0$. Hence the number of solutions of the system (5.2) is $O(N^{3/2-\delta}) = O(|A|^2|B|^{1-2\delta})$. Finally $|C| \gg |B|^{1+2\delta}$, which is the desired conclusion. $\square$

As a corollary of Theorem 5.6 Vinh derived the following:

Corollary 5.8 (Vinh)**.** Let $\mathcal{P}$ be a collection of points and $\mathcal{L}$ be a collection of lines in ${\mathbb{F}_p}^2$. Suppose that $|\mathcal{P}|, |\mathcal{L}| \leq N = p^\alpha$ with $1 + \beta \leq \alpha \leq 2 - \beta$ for some $0 < \beta < 1$. Then we have

$$|\{(p, l) \in \mathcal{P} \times \mathcal{L} \ : \ p \in l\}| \leq 2N^{\frac{3}{2}-\frac{\beta}{4}}.$$

Using this statement we can state a quantitative form of the above theorem in a certain range of the domains.

Theorem 5.9 (Hegyvári-Hennecart [HH09])**.** Let F as in Theorem 5.7 and $\alpha > 1/2$. For any pair (A, B) of subsets of $\mathbb{F}_p$ such that $|A| \asymp |B| \asymp p^\alpha$, we have

$$|F(A, B)| \gg |A|^{1+\frac{\min\{2\alpha-1;2-2\alpha\}}{2}}.$$

5.1.2 Complete expanders

We start this section to introduce the notion of complete expander.

Definition 5.10. Let $I \subset (0,1)$ be a non empty interval. A family $\{F\}$ of two variables functions is called *complete expander according to* I if for any $\alpha \in I$, for any prime number p and any pair (A, B) of subsets of $\mathbb{F}_p$ satisfying $|A|, |B| \asymp p^\alpha$, we have

$$|F(A,B)| \geq cp^{\min\{1;2\alpha\}}.$$

It is known that a random $f(x, y)$ is complete expanders with a large probability. Nevertheless, we can show that some explicit expanders are not complete, in particular Bourgain's function $F(x, y) = x^2 + xy = x(x + y)$.

Now we claim two negative answers:

Proposition 5.11. *Let $k \geq 2$ be an integer, $u \in \mathbb{Z}$ and $F(x,y) = x^{2k} + ux^k + x^k y = x^k(x^k + y + u)$. Then for any α, $0 < \alpha \leq 1/2$, F is not a complete expander according to $\{\alpha\}$.*

and

Proposition 5.12. *Let $f(x)$ and $g(y)$ be non constant integral polynomials and $F(x, y) = f(x)(f(x) + g(y))$. Then F is not a complete expander according to $\{1/2\}$.*

For the proof of Theorem 5.11 and 5.12 we need the following lemma which is due to Erdős.

Lemma 5.13 (Erdős Lemma)**.** There exists a positive real number δ such that the number of different integers ab where $1 \leq a, b \leq n$ is $O(n^2/(\ln n)^\delta)$.

(A sharper result due to G. Tenenbaum [T] implies that δ can be taken equal to $1 - \frac{1+\ln\ln 2}{\ln 2}$ in this statement.)

Proof of Theorem 5.11. Let L be a positive integer such that $L < \sqrt{p}/2$. The set of k-th powers in $\mathbb{F}_p^*$ is a subgroup of $\mathbb{F}_p^*$ with index $l = \gcd(k, p-1) \leq k$. Thus there exists $a \in \mathbb{F}_p^*$ such that $[1, L]$ contains at least L/l residue classes of the form ax^k, $x \in \mathbb{F}_p^*$. We let $A = \{x \in \mathbb{F}_p^* : ax^k \in [1, L]\}$, which has cardinality at least L since each k-th power has l k-th roots modulo p. We let $B = \{y \in \mathbb{F}_p : a(y+u) \in [1, L]\}$. We clearly have $|B| = L$. Moreover the elements of $F(A, B)$ are of the form $x^k(x^k + y + u)$ with $x \in A$ and $y \in B$, thus are of the form $a'^2x'y'$ where $x', y' \in [1, 2L]$ and $aa' = 1$ in $\mathbb{F}_p$. By Erdős Lemma, we infer $|F(A, B)| = O(L^2/(\ln L)^\delta) = o(L^2)$. $\square$

dc_1260_16

Proof of Theorem 5.12. We shall need the following result:

Lemma 5.14. Let $u \in \mathbb{F}_p$, L be a positive integer less than $p/2$ and $f(x)$ be any integral polynomial of degree $k \geq 1$ (as element of $\mathbb{F}_p[x]$). Then the number $N(I)$ of residues $x \in \mathbb{F}_p$ such that $f(x)$ lies in the interval $I = (u-L, u+L)$ of $\mathbb{F}_p$ is at least $L-(k-1)\sqrt{p}$.

Proof. Let J be the indicator function of the interval $[0, L)$ of $\mathbb{F}_p$ and let

$$T := \sum_{h \in \mathbb{F}_p} \widehat{J * J}(h) S_f(-h, p) e_p(hu),$$

where the exponential sum

$$S_f(h, p) := \sum_{x \in \mathbb{F}_p} e_p(hf(x))$$

is known to satisfy the bound $|S_f(h,p)| \leq (k-1)\sqrt{p}$ whenever $h \neq 0$ in $\mathbb{F}_p$ and p is an odd prime number. On the one hand, we have

$$\begin{aligned} T &= p\widehat{J * J}(0) + \sum_{h \in \mathbb{F}_p \setminus \{0\}} \widehat{J * J}(h) S_f(-h, p) e_p(hu) \\ &\geq pL^2 - k\sqrt{p} \sum_{h \in \mathbb{F}_p \setminus \{0\}} |\widehat{J * J}(h)| \\ &\geq pL^2 - kLp^{3/2}, \end{aligned}$$

by the bound for Gaussian sums and Parseval Identity. Hence

$$T \geq pL(L - k\sqrt{p}) \tag{5.4}$$

On the other hand,

$$\begin{aligned} T &= \sum_{h \in \mathbb{F}_p} \sum_{y \in \mathbb{F}_p} \sum_{z \in \mathbb{F}_p} J(z)J(y+z)e_p(h(y+u)) \sum_{x \in \mathbb{F}_p} e_p(hf(x)) \\ &= \sum_{x \in \mathbb{F}_p} \sum_{y \in \mathbb{F}_p} \sum_{z \in \mathbb{F}_p} J(z)J(y+z) \sum_{h \in \mathbb{F}_p} e_p(h(y+u-f(x))) \\ &= p \sum_{x \in \mathbb{F}_p} d_L(f(x)-u), \end{aligned}$$

where $d_L(z)$ denotes the number of representations in $\mathbb{F}_p$ of z under the form $j - j'$, $0 \leq j, j' < L$. Since obviously $d_L(z) \leq L$ for each $z \in \mathbb{F}_p$, we get

$$T \leq pLN(I).$$

Combining this bound and (6.3), we deduce the lemma. $\square$

Now we complete our proof.

We choose p large enough so that both $f(x)$ and $g(y)$ are not constant polynomials modulo p. Let $L = k\sqrt{p}$, and define A (resp. B) to be the set of the residue classes x (resp. y) such that $f(x)$ (resp. $g(y)$) lies in the interval $(0, 2L)$. By the previous lemma, one has $|A|, |B| \geq \sqrt{p}$. Moreover for any $(x, y) \in A \times B$, we have $f(x)$ and $f(x) + g(y)$ in the interval $(0, 4L)$. By Erdős Lemma, the number of residues modulo p which can be written as $F(x, y)$ with $(x, y) \in A \times B$, is at most $O(L^2/(\ln L)^\delta) = o(p)$, as p tends to infinity.

$\square$

Remark 5.15. We did not discuss the polynomials $F_3(x, y) = f(x, y) \cdot g(x, y)$ and $F_4(x, y) = f(g(x, y), y)$ yet. Here $F_4(x, y) = f(g(x, y), y) = (x + 1)y$ which covered by our Theorem 5.7 – and recently many authors improve the expanding measure of it (in the form $|A(A + 1)|$).

In 2015 T. Tao discovered a very deep theorem which describes expanding polynomials with two variables under a restriction of the domain. His theorem also covers F_3. (see [TaoEx])

Before this theorem we could prove just a conditional version.

5.2 Covering polynomials and sets

Bounds for exponential sums are related to additive questions in $\mathbb{F}_p$. In [S] Sárközy investigated the following problem: let $A, B, C, D \subseteq \mathbb{F}_p$ be non-empty sets. Then the equation

$$a + b = cd$$

is solvable in $a \in A, b \in B, c \in C, d \in D$ provided $|A||B||C||D| > p^3$. This simple equation has many interesting consequences. We merely mention here just an improvements of the modular Fermat theorem which was firstly investigated by Schur. One can ask the more general question of investigating the solvability of

$$a + b = F(c, d) \tag{5.5}$$

where $F(x, y)$ is a two variables polynomial with integer coefficients.

One can read easily from this result, that this problem is equivalent to the following problem: let $G(x, y, z, w) = x + y + F(x, y)$. Now what condition guaranties that for sets $A, B, C, D \subseteq \mathbb{F}_p$, $G(A, B, C, D)$ covers everything, i.e.

$$G(A, B, C, D) = \mathbb{F}_p?$$

In the present section we collect some result on this topic.

Let $A, B \subseteq \mathbb{F}_p$ and let $H < \mathbb{F}_p^*$. We ask the solvability of the equation

$$a + b = h;\ (a, b, h) \in A \times B \times H.$$

Restricting the cardinality of H to some region we improve the result of Sárközy:

Theorem 5.16 (Hegyvári [HE12])**.** Let $A, B \subseteq \mathbb{F}_p$, $H < \mathbb{F}_p^*$. Write $|A||B| = p^{2-2\alpha}$ and $|H| = p^\beta$. Then the equation

$$a + b = h;\ (a, b, h) \in A \times B \times H$$

is solvable, provided

$$\beta > \frac{8\alpha + 1}{3}.$$

Essentially in the same way we can prove a more general result. Assume that $C, D \subseteq \mathbb{F}_p^*$, and assume that the cardinality of the generating subgroups of C and D are close to $|C|$ and $|D|$ respectively. We have

Theorem 5.17 (Hegyvári [HE12])**.** Assume that $C, D \subseteq \mathbb{F}_p^*$, $A, B \subseteq \mathbb{F}_p$. Let $|A||B| = p^{2-2\alpha}; |C| = p^\beta,\ |D| = p^\gamma,\ \langle C\rangle = G_1, \langle D\rangle = G_2, |G_1| = p^\delta, |G_2| = p^\theta, \max\{\delta, \theta\} < 3/4$. Then the equation

$$a + b = cg \quad (a, b, c, g) \in A \times B \times G_1 \times G_2,$$

is solvable, provided

$$\frac{5}{16}(\beta + \gamma) > \alpha + \frac{1 + \delta + \theta}{8}.$$

Corollary 5.18 (Hegyvári [HE12])**.** Let $A, B \subseteq \mathbb{F}_p$, $H < \mathbb{F}_p^*$. Write $|H| = p^\beta$. Then the equation

$$a + b = h;\ (a, b, h) \in A \times B \times H$$

is solvable, provided

$$|A||B||H|^2 > p^{\frac{9+5\beta}{4}}.$$

Note when $0 < \beta < \frac{3}{5}$, then this bound is better then the Sárközy's p^3 (the reason is that we can utilize the arithmetic structure od the sets).

Proof of Theorem 5.16 and 5.17. Proving theorems above we need some lemmas. Firstly we quote a well-known condition to the solvability like (5.5).

Lemma 5.19. Let $F(x, y) \in \mathbb{Z}[x, y]$ and let $S(r) = \sum_{c\in C,g\in D} e(r(F(c, d)))$, $r \in \mathbb{F}_p^*$.

Assume that for some $M > 0$, $\max_{r\neq 0} |S(r)| \leq M$. If

$$\sqrt{|A||B|}|C||D| > pM,$$

then the equation $a + b = F(c, d)$ $(a, b, c, d) \in A \times B \times C \times D$, is solvable.

For the proof see e.g. [S],[Ga].

A well-known estimation for the double exponential sums is

$$\Big|\sum_{x\in X,y\in Y} e(xy)\Big| < \sqrt{p|X||Y|}$$

noted by Vinogradov. This bound is non-trivial in the range $|X||Y| \gg p$. For our purpose we need the opposite range.

Lemma 5.20. Let $A, B \subseteq \mathbb{Z}_N$ $r \neq 0$. Write $S(r) = \sum_{x\in A} \sum_{y\in B} e(rxy)$, we have

$$|S(r)|^8 \leq N \cdot |A|^4 \cdot |B|^4 E_+(A)E_+(B),$$

where $E_+(\cdot)$ is the additive energy.

It is a result of Bourgain and Garaev. For seek of completeness we show the short proof.

Proof. We will use Cauchy inequality three times: Firstly respect to the variables from A:

$$|S(r)|^2 \leq |A| \sum_{y,y' \in B} \left| \sum_{x \in A} e(rx(y-y')) \right|.$$

In the second step (replace now A with $B \times B$) again, and denote $d(z) = \{(y, y' \in B; z = y - y'\}$ the representation function. Then we have

$$|S(r)|^4 \leq |A|^2|B|^2 \sum_{z \in \mathbb{Z}_N} d(z) \left| \sum_{x \in A} e(rxz) \right|^2.$$

Finally again by the Cauchy inequality

$$|S(r)|^8 \leq |A|^4|B|^4 \sum_{z \in \mathbb{Z}_N} d^2(z) \sum_{z \in \mathbb{Z}_N} \left| \sum_{x \in A} e(rxz) \right|^4 = N \cdot |A|^4 \cdot |B|^4 E_+(A)E_+(B).$$

□

The third lemma which will be necessary for us is the following ([TV] Ch. 9):

Lemma 5.21. Let $G < \mathbb{F}_p^*$, $|G| \ll p^{3/4}$, $Y \subseteq G$, then

$$E_+(Y) \ll |G||Y|^{3/2}.$$

Now our task is to give a bound for M.

Firstly we will do it under the condition of Theorem 5.17 and after for the simplicity we end the proof under the condition of Theorem 5.16. Assume that $C, D \subseteq \mathbb{F}_p^*$ and let the generating subgroup of C and D, $\langle C \rangle = G_1$, $\langle D \rangle = G_2$ respectively.

By Lemma 5.20 and 5.21 we conclude that

$$\begin{aligned} |S(r)| &\leq |C|^{1/2}|D|^{1/2}(pE_4^+(C)E_4^+(D))^{1/8} \ll \\ &\ll p^{1/8}|C|^{11/16}|D|^{11/16}|G_1|^{1/8}|G_2|^{1/8}. \end{aligned} \tag{2.1}$$

By Lemma 5.19 we obtain that the equation $a + b = cd$ $(a, b, c, d) \in A \times B \times C \times D$, is solvable, provided

$$|A|^{1/2}|B|^{1/2}|C|^{5/16}|D|^{5/16} \gg p^{9/8}|G_1|^{1/8}|G_2|^{1/8}. \tag{2.2}$$

Writing $|A||B| = p^{2-2\alpha}; |C| = p^{\beta}, \ |D| = p^{\gamma}, |G_1| = p^{\delta}, |G_2| = p^{\theta}$ (2.2) is equivalent to

$$1 - \alpha + \frac{5}{16}(\beta + \gamma) > \frac{9 + \delta + \theta}{8},$$

which gives Theorem 5.17. When $|A||B| = p^{2-2\alpha}; |H| = p^{\beta}$, it gives the constraint

$$\beta > \frac{8\alpha + 1}{3}.$$

and we obtain Theorem 5.16.

□

We merely mention that functions $F_1(x,y) = xy + x^2 h_1(y)$ and $F_2(x,y) = x^2 y + x h_2(y)$, $(h_i(y) \in \mathbb{Z}[y]; \ i = 1, 2$ non-zero polynomials) are admissible for the equation (5.5). Namely Bourgain gave the bounds

$$\Big| \sum_{c \in C, d \in D} e_p(F_i(c,d)) \Big| = O(|C||D|p^{-\varepsilon}),$$

where ε is a positive constant (see Propositions 3.6 and 3.7 in [B]).

So we have

Fact 5.22 (Hegyvári-Hennecart [HH09])**.** Let F_i be one of the two families of functions defined above. There exist real numbers $0 < \delta, \delta' < 1$ such that for any p and for any sets $A, B, C, D \subseteq \mathbb{F}_p$ fulfilling the conditions

$$|C| > p^{1/2-\delta}, \quad |D| > p^{1/2-\delta} \quad |A||B| > p^{2-\delta'},$$

there exist $a \in A, b \in B, c \in C, d \in D$ solving the equation

$$a + b = F_i(c,d) \ i = 1, 2. \tag{5.6}$$

Observe that in this case we obtain a better assumption to the solvabilty than p^3.

We finish this section to show that some sum-product set covers a given prime field.

Theorem 5.23. [Hegyvári [He09]] Let $A \subseteq \mathbb{F}_p$, $|A| > 2$, and let $q(x) = 1+u_1x+\cdots+u_Dx^D$ be a non-constant polynomial, and let $Q = \langle q(r) : r \in \mathbb{F}_p\rangle$ be a multi-set of the values.

There exists a multi-subset B of Q, $c_1 > 0$ for which

$$|B| < c_1 \log \frac{\log p/D}{log|A|} + 2D + 3 \tag{5.7}$$

and

$$FP_{mult}(B) * A = \sum_{h \in FP_{mult}(B)} h \cdot A = \mathbb{F}_p.$$

Proof of 5.23. For the proof we need the following lemma:

Lemma 5.24. Let $A, B \subseteq \mathbb{F}_p$. Let $S(r) := |\{A + q(r) \cdot B\}|$. We have

$$\max_{r \in \mathbb{F}_p} S(r) \geq \frac{p|A||B|}{p + D|A||B|}, \tag{5.8}$$

where $D = \deg q(x)$.

The idea that we used in the proof of the lemma is similar to the one in [GK].

Proof of Lemma 5.24. Denote by $R(r,m)$ the number of representations of m in the form $m = a+q(r)\cdot b$. Fix an element $r \in \mathbb{F}_p$. One now deals with the sum $\sum_m R^2(r,m)$. It counts the number of quadruples $(a, a', b, b') \in A\times A\times B\times B$ such that $a + q(r) \cdot b = a' + q(r) \cdot b'$. Note that $a \neq a'$ if and only if $b \neq b'$. Hence at the diagonal case we obtain

$$\sum_r \sum_{m;a=a'} R^2(r,m) = \sum_r |A||B| = p \cdot |A||B|. \tag{5.9}$$

Assume $a \neq a'$ and write the equality $a + q(r) \cdot b = a' + q(r) \cdot b'$ in the form $q(r) = \frac{a'-a}{b-b'}$. In the variable r we get at most D many solutions, thus we argue that

$$\sum_r \sum_{m;a\neq a'} R^2(r,m) = \sum_{m;a\neq a'} \sum_r R^2(r,m) \leq D \cdot |A|^2|B|^2. \tag{5.10}$$

By (5.9) and (5.10)

$$\sum_r \sum_m R^2(r,m) \leq p \cdot |A||B| + D \cdot |A|^2|B|^2. \tag{5.11}$$

Let $R^2(r_0,m) := \min_r R^2(r,m)$.

By (5.11)

$$p \cdot \sum_m R^2(r_0,m) \leq \sum_r \sum_m R^2(r,m) \leq p \cdot |A||B| + D \cdot |A|^2|B|^2,$$

and hence

$$\sum_m R^2(r_0,m) \leq |A||B| + p^{-1} \cdot D \cdot |A|^2|B|^2.$$

By the Cauchy inequality

$$(\sum_m R(r_0,m))^2 \leq S(r_0)(\sum_m R^2(r_0,m)),$$

and by the simple observation

$$\sum_m R(r_0,m) = |A||B|,$$

we obtain

$$|A|^2|B|^2 \leq S(r_0)(|A||B| + p^{-1} \cdot D \cdot |A|^2|B|^2),$$

hence (5.8). □

Now we follow an iteration step. We define a sequence of sets $A_0, A_1, \ldots$ and sequence $b_0, b_1, \ldots$ of the values of the range of Q as follows: let $A_0 = A$ and $b_0 = q(0) = 1$. By Lemma 5.24 we obtain an r_1, such that $S(r_1) \geq \frac{p|A|^2}{p+D|A|^2}$; so let $A_1 = A_0 + q(r_1) \cdot A_0$ and thus

$$|A_1| \geq \frac{p|A_0|^2}{p + D|A_0|^2}.$$

Generally assume that the sets $A_0, A_1, \ldots A_k$ and the sequence $b_0, b_1, \ldots b_k$ have been defined. Then by Lemma 5.24 we have an r_{k+1}, such that for the set $A_{k+1} := A_k + q(r_{k+1}) \cdot A_k$ we obtain

$$|A_{k+1}| \geq \frac{p|A_k|^2}{p + D|A_k|^2}. \tag{5.12}$$

Repeat this process unless we have $\frac{p}{p+D|A_n|^2} < \frac{9}{10}$, or equivalently

$$|A_n| > \sqrt{\frac{p}{9D}}. \tag{5.13}$$

We prove that this process is terminated, i.e. there exists an n for which (5.13) holds. From (5.12) and from the definition of n we conclude that for $1 \le k < n$

$$|A_{k+1}| \ge \frac{p|A_k|^2}{p + D|A_k|^2} \ge \frac{9}{10}|A_k|^2,$$

and by induction it is not too hard to check that

$$|A_k| \ge \frac{10}{9} \cdot \left(9|A|/10\right)^{2^k}. \tag{5.14}$$

By (5.13) and (5.14) we have that

$$n \le c_1 \log \frac{\log p/D}{log|A|} \tag{5.15}$$

for some $c_1 > 0$.

Repeat this process once more, then an easy calculation shows that $|A_{n+1}| \ge \frac{p}{10D}$. Finally let $r_{n+2} = \cdots = r_{n+2+10D} = 0$, and then by the Cauchy-Davenport inequality we obtain that

$$A_{n+2+10D} = \mathbb{F}_p$$

provided p is large enough, compared D.

In the rest of the proof we check that for the set B (5.7) holds and $A_{n+2+10D} = FP_{mult}(B) \cdot A$. For $0 \le k \le n+2+10D$, $b_k = q(r_k)$, $B = \{b_k : 0 \le k \le n+2+10D\}$ hence by (5.15) we obtain (5.7).

Finally by induction we prove that

$$A_k = FP_{mult}(b_0, \dots, b_k)A. \tag{5.16}$$

For $k = 0$ $A_0 = q(0)A = A$. From (5.16)

$$A_{k+1} = A_k + b_{k+1}A_k = FP_{mult}(b_0, \dots, b_k)A + b_{k+1} \cdot FP_{mult}(b_0, \dots, b_k)A,$$

where in the first term there are those $h \in FP_{mult}(b_0, \dots, b_{k+1})$ which do not contain b_{k+1}, while in the second there are the ones which do. $\square$

Chapter 6

Structure result for cubes in Heisenberg groups

Let p be a prime number and $\mathbb{F}$ the field with p elements. We denote by H_n the $(2n+1)$-dimensional Heisenberg linear group over $\mathbb{F}$ formed with the upper triangular square matrices of size $n+2$ of the following kind

$$[\underline{x}, \underline{y}, z] = \begin{pmatrix} 1 & \underline{x} & z \\ 0 & I_n & {}^t\underline{y} \\ 0 & 0 & 1 \end{pmatrix},$$

where $\underline{x} = (x_1, x_2, \ldots, x_n)$, $\underline{y} = (y_1, y_2, \ldots, y_n)$, $x_i, y_i, z \in \mathbb{F}$, $i = 1, 2, \ldots, n$, and I_n is the $n \times n$ identity matrix.

We have $|H_n| = p^{2n+1}$. and we recall the product rule in H_n:

$$[\underline{x}, \underline{y}, z][\underline{x}', \underline{y}', z'] = [\underline{x} + \underline{x}', \underline{y} + \underline{y}', \langle \underline{x}, \underline{y}' \rangle + z + z'],$$

where $\langle \cdot, \cdot \rangle$ is the inner product, that is $\langle \underline{x}, \underline{y} \rangle = \sum_{i=1}^{n} x_i y_i$.

So this set of $(n+2) \times (n+2)$ matrices form a group whose unit is $e = [\underline{0}, \underline{0}, 0]$.

As group-theoretical properties of H_n, we recall that H_n is non abelian and two-step nilpotent, that is the double commutator satisfies

$$[[a, b], c] = aba^{-1}b^{-1}cbab^{-1}a^{-1}c^{-1} = e$$

for any $a, b, c \in H_n$, where the commutator of a and b is defined as $[a, b] := aba^{-1}b^{-1}$.

The Heisenberg group possesses an interesting structure in which we can prove that in general there is no *good model* for a subset A with a small *squaring constant* $|A \cdot A|/|A|$ unlike for subsets of abelian groups. To know what we mean on good model let us recall the notion of *Freiman isomorphism.*

Let $s \geq 2$ be an integer and $A \subset H$ and $B \subset G$ be subsets of arbitrary (multiplicative) groups. A map $\pi : A \to B$ is said to be a Freiman s-homomorphism if for any $2s$-tuple $(a_1, \ldots, a_s, b_1, \ldots, b_s)$ of elements of A and any signs $\epsilon_i = \pm 1$, $i = 1, \ldots, s$, we have

$$a_1^{\epsilon_1} \ldots a_s^{\epsilon_s} = b_1^{\epsilon_1} \ldots b_s^{\epsilon_s} \Longrightarrow \pi(a_1)^{\epsilon_1} \ldots \pi(a_s)^{\epsilon_s} = \pi(b_1)^{\epsilon_1} \ldots \pi(b_s)^{\epsilon_s}.$$

Observe that in the case of abelian groups, we may set, without loss of generality, all the signs to $+1$. If moreover π is bijective and π^{-1} is also a Freiman s-homomorphism, then π is called a Freiman s-isomorphism from A into G. In this case, A and B are said to be Freiman s-isomorphic.

Green and Ruzsa proved in that a structural result holds for small squaring of finite set A in an abelian group. Namely A has a good Freiman model, that is a relatively small finite group G and a Freiman s-isomorphism from A into G. It reads as follows:

Theorem 6.1 (Green-Ruzsa)**.** Suppose that G is abelian, and that $|A+A| \leq K|A|$. Let $s \geq 2$. Then there is an abelian group G' with $|G' \leq (10sK)^{10K^2}|A|$ such that A is Freiman s-isomorphic to a subset of G'.

In 2007 B. Green gave an example showing that there need not exist good models in the *non-abelian* setting. His counterexample worked in Heisenberg groups. In 2012 we (Hegyvári-Hennecart) improved a result of him (based on Green's approach) but also includes arguments coming from group theory and Fourier analysis with additional tools, e.g. a recent incidence theorem due to Vinh (discussed in Chapter 4).

So it was our starting in the world of Heisenberg groups.

6.1 Structure results

Lately many new results pop up on expansion of Lie-type simple groups. Helfgott proved that for $A \subset SL_n(\mathbb{Z}_p)$, $|A \cdot A \cdot A| > |A|^{1+\varepsilon}$ (where $\varepsilon > 0$

is an absolute constant) unless A is contained in a proper subgroup. Or a nice and deep result (called "Convolution bound") of Babai-Nikolov-Pyber, which ensures that if $A \subseteq SL_2(\mathbb{Z}_p)$, and $|A| \sim p^{5/2}$ then $|A^2|$ covers at least one third of the group.

Nevertheless it is very less known on the structure of (k-fold) product sets in this non-abelian groups.

Certainly the general question is very hard and cannot be handled easily.

We will restrict our attention to subsets that will be called *cubes*.

Let $B \subseteq H_n$, and write the projections of B onto each coordinates by $X_1, X_2, \dots, X_n,\ Y_1, Y_2, \dots, Y_n$ and Z, i.e. one has $[\underline{x}, \underline{y}, z] \in B$, $\underline{x} = (x_1, x_2, \dots, x_n), \underline{y} = (y_1, y_2, \dots, y_n)$, if and only if $x_i \in X_i$ or $y_i \in Y_i$ for some i, or $z \in Z$.

Definition 6.2. A subset B of H_n is said to be a *cube* if

$$B = [\underline{X}, \underline{Y}, Z] := \{[\underline{x}, \underline{y}, z] \text{ such that } \underline{x} \in \underline{X},\ \underline{y} \in \underline{Y},\ z \in Z\}$$

where $\underline{X} = X_1 \times \dots \times X_n$ and $\underline{Y} = Y_1 \times \dots \times Y_n$ with non empty-subsets $X_i, Y_i \subset \mathbb{F}^*$.

Theorem 6.3. [Hegyvári-Hennecart [HH13]] For every $\varepsilon > 0$, there exists a positive integer n_0 such that if $n \geq n_0$, $B \subseteq H_n$ is a cube and

$$|B| > |H_n|^{3/4+\varepsilon}$$

then there exists a non trivial subgroup G of H_n, namely its center $[\underline{0}, \underline{0}, \mathbb{F}]$, such that $B \cdot B$ contains a union of at least $|B|/p$ many cosets of G.

We stress the fact that n_0 depends only on ε and that this result is valid uniformly in p.

Remark 6.4. The statement in Theorem 6.3 can be plainly extended to any subset $B' \subset H_n$ which derives from a cube B by conjugation : $B' = P^{-1}BP$ where P is a given element of H_n.

Furthermore we will show that the exponent $3/4 + \varepsilon$ in Theorem 6.3 cannot be essentially reduced to less than $1/2$:

Proposition 6.5. *[Hegyvári-Hennecart [HH13]] For any n and p there exists a cube $B \subseteq H_n$ such that*

$$|B| \geq \frac{\sqrt{p}}{4(2n)^n} |H_n|^{1/2}$$

and the only cosets contained in $B \cdot B$ are cosets of the trivial subgroup of H_n.

Choosing p large relative to n in this result implies the desired effect.

6.1.1 Fourier analysis for a sum-product estimate

We will use the following sum-product estimate:

Proposition 6.6. *Let $n, m \in \mathbb{N}$, $X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots Y_n \subseteq \mathbb{F}^* = \mathbb{F}\setminus\{0\}$, $Z \subseteq \mathbb{F}$. We have*

$$mZ + \sum_{j=1}^{n} X_j \cdot Y_j := \Big\{ z_1 + \cdots + z_m + \sum_{j=1}^{n} x_j y_j,\ z_i \in Z,\ x_j \in X_j,\ y_j \in Y_j \Big\} = \mathbb{F},$$

provided

$$|Z|^2 \prod_{i=1}^{n} |X_i||Y_i| > p^{n+2}. \tag{6.1}$$

Proof. Let $X_i(t)$ (resp. $Y_i(t)$ and $Z(t)$) be the indicator of the set X_i (resp Y_i and Z). One defines

$$f_i(t) = \frac{1}{|X_i|} \sum_{a \in X_i} Y_i\Big(\frac{t}{a}\Big), \quad i = 1, 2, \ldots, n.$$

Notice that $0 \leq f_i(t) \leq 1$, and $f_i(t) > 0$ if and only if $t \in X_i \cdot Y_i$. The Fourier transform of f_i is

$$\widehat{f_i}(r) = \sum_x f_i(x) e(xr)$$

where $e(x) = \exp(2\pi i x/p)$ as usual.

An easy calculation shows that for every $i = 1, 2, \ldots, n$

$$\widehat{f_i}(r) = \frac{1}{|X_i|} \sum_{a \in X_i} \widehat{Y_i}(ra)$$

and

$$\widehat{f}_i(0) = \frac{1}{|X_i|} \sum_{a \in X_i} \widehat{Y}_i(0) = |Y_i|, \tag{6.2}$$

since $\widehat{Y}_i(0) = \sum_x Y_i(x) = |Y_i|$. Using the Cauchy inequality and the Parseval equality we get if $p \nmid r$

$$|\widehat{f}_i(r)| \leq \frac{1}{\sqrt{|X_i|}} \sqrt{\sum_x |\widehat{Y}_i(x)|^2} = \sqrt{\frac{p|Y_i|}{|X_i|}} \tag{6.3}$$

Let $u \in \mathbb{F}$. Let S be the number of solutions of the equation

$$u = z_1 + z_2 + \cdots + z_m + \sum_{j=1}^{n} x_j y_j, \quad z_i \in Z,\ x_j \in X_j,\ y_j \in Y_j.$$

We can express S by the mean of the Fourier transforms of Z and f_i as follows:

$$pS = \sum_{r \in \mathbb{F}_p} \widehat{Z}(r)^m \prod_{i=1}^{n} \widehat{f}_i(r) e(-ru).$$

Our task is to show that this exponential sum is positive if the desired bound for the cardinalities (6.1) holds. Separating $r = 0$ and using (6.2) we can bound S as

$$\begin{aligned} pS \geq & |Z|^m \prod_{i=1}^{n} |Y_i| - \sum_{r \neq 0} |\widehat{Z}(r)|^m \prod_{i=1}^{n} |\widehat{f}_i(r)| \\ \geq & |Z|^m \prod_{i=1}^{n} |Y_i| - |Z|^{m-2} \prod_{i=1}^{n} \sqrt{\frac{p|Y_i|}{|X_i|}} \sum_{r \neq 0} |\widehat{Z}(r)|^2 \\ \geq & |Z|^m \prod_{i=1}^{n} |Y_i| - p|Z|^{m-1} \prod_{i=1}^{n} \sqrt{\frac{p|Y_i|}{|X_i|}} \end{aligned}$$

by the Parseval equality and (6.3). Hence $S > 0$ whenever

$$|Z|^2 \prod_{i=1}^{n} |X_i||Y_i| > p^{n+2}.$$

This completes the proof. $\square$

Remark 6.7. The idea what we used at the proof of Proposition above essentially the same what is in [He09]

Proof of Theorem 6.3. By the remark preceding Theorem 6.3 we may plainly assume that $|Z| < p/2$.

By the assumption on the cube B we have

$$|B| = |Z| \left(\prod_{i=1}^{n} |X_i||Y_i| \right) > |H_n|^{3/4+\varepsilon} = p^{3n/2+3/4+\varepsilon(2n+1)}. \tag{6.4}$$

For each i, there exists an element $a_i \in \mathbb{F}$ such that the number of solutions to the equation $a_i = x_i + x_i'$, $x_i, x_i' \in X_i$, is at least $|X_i|^2/p$. We denote by $\tilde{X}_i = X_i \cap (a_i - X_i)$ the set of the elements $x_i \in X_i$ such that $a_i - x_i \in X_i$. We thus have $|\tilde{X}_i| \geq |X_i|^2/p$. We similarly define $\tilde{Y}_i = Y_i \cap (b_i - Y_i)$ for some appropriate b_i and also have $|\tilde{Y}_i| \geq |Y_i|^2/p$. It follows by (6.4) that

$$|Z|^2 \left(\prod_{i=1}^{n} |\tilde{X}_i||\tilde{Y}_i| \right) \geq \frac{(|Z| \prod_{i=1}^{n} |X_i||Y_i|)^2}{p^{2n}} > p^{n+3/2+\epsilon(4n+2)}.$$

Hence for $n > 1/8\epsilon$ we obtain from Proposition 6.6 that $2Z + \sum_{i=1}^{n} \tilde{X}_i \cdot \tilde{Y}_i = \mathbb{F}$ and consequently

$$B \cdot B \supseteq [(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n), \mathbb{F}],$$

that is $B \cdot B$ contains at least one coset of the non trivial subgroup $G = [\underline{0}, \underline{0}, \mathbb{F}]$ of H_n.

In fact we may derive from the preceding argument a little bit more: for any index i we have

$$\sum_{a_i \in \mathbb{F}} |X_i \cap (a_i - X_i)| = |X_i|^2, \quad \sum_{b_i \in \mathbb{F}} |Y_i \cap (b_i - Y_i)| = |Y_i|^2,$$

hence

$$\prod_{i=1}^{n} \left(\sum_{a_i \in \mathbb{F}} |X_i \cap (a_i - X_i)| \right) \left(\sum_{b_i \in \mathbb{F}} |Y_i \cap (b_i - Y_i)| \right) = \prod_{i=1}^{n} |X_i|^2 |Y_i|^2,$$

or equivalently by developing the product

$$\sum_{\underline{a},\underline{b}\in\mathbb{F}^n}\prod_{i=1}^{n}|X_i\cap(a_i-X_i)||Y_i\cap(b_i-Y_i)|=\prod_{i=1}^{n}|X_i|^2|Y_i|^2. \tag{6.5}$$

We denote by E the set of all pairs $(\underline{a},\underline{b})\in\mathbb{F}^n\times\mathbb{F}^n$ such that

$$|Z|^2\prod_{i=1}^{n}|X_i\cap(a_i-X_i)||Y_i\cap(b_i-Y_i)|>p^{n+2}.$$

For such a pair $(\underline{a},\underline{b})$, the coset $[\underline{a},\underline{b},\mathbb{F}]$ is contained in $B\cdot B$ by the above argument. Then by (6.5)

$$\left(\prod_{i=1}^{n}|X_i||Y_i|\right)|E|+p^{n+2}(p^{2n}-|E|)>\left(\prod_{i=1}^{n}|X_i||Y_i|\right)^2$$

hence

$$|E|>\frac{\prod_{i=1}^{n}|X_i|^2|Y_i|^2-p^{3n+2}}{\prod_{i=1}^{n}|X_i||Y_i|-p^{n+2}}.$$

For $n>1/\epsilon$, we have by (6.4) and the fact that $|Z|\le p$

$$\prod_{i=1}^{n}|X_i||Y_i|>p^{3n/2+7/4},$$

hence

$$|E|\ge(1-p^{-3/2})\prod_{i=1}^{n}|X_i||Y_i|=(1-p^{-3/2})\frac{|B|}{|Z|}.$$

Since $|Z|\le p/2$, we thus have shown that $B\cdot B$ contains at least $2(1-p^{-3/2})|B|/p\ge|B|/p$ cosets $[\underline{a},\underline{b},\mathbb{F}]=[\underline{a},\underline{b},0][\underline{0},\underline{0},\mathbb{F}]$, as we wanted. $\square$

Proof of Proposition 6.5. Since B is a cube, $B\cdot B$ is contained in a cube which takes the form $[\underline{U},\underline{V},W]$ where $\underline{U}$, $\underline{V}\subset\mathbb{F}^n$ are direct products of subsets of $\mathbb{F}$ and $W\subset\mathbb{F}$. Since any non trivial subgroup of H_n has at least one of his $(2n+1)$ coordinate projections equals to $\mathbb{F}$, it suffices to prove that neither W is equal to $\mathbb{F}$, nor U, nor V contains a subset of the type $\{x_1\}\times\cdots\times\mathbb{F}\times\cdots\times\{x_n\}$.

Let $B = [R, R, Z]$ where

$$R = \left\{ (r_1, r_2, \ldots, r_n) \in \mathbb{F}^n \mid 0 \le r_i < \sqrt{(p-1)/2n} \right\}$$

and

$$Z = \left\{ z \in \mathbb{F} \mid 0 \le z < p/4 \right\}.$$

We have $|B| \ge p^{n+1}/4(2n)^n$ and

$$B \cdot B \subseteq [R + R, R + R, Z + Z + \langle R, R \rangle].$$

Clearly $R+R \subseteq \left[0, \sqrt{2p/n}\right]^n$, $Z+Z \subseteq \left[0, (p-1)/2\right)$ and $\langle R, R \rangle \subseteq \left[0, (p-1)/2\right]$. Hence the statement. $\square$

We close this section some further results. For $U \subset \mathbb{F}^2$ and $Z \subset \mathbb{F}$ we define the so called *semi-cube* A in $H = H_3$ by

$$A = \{[x, y, z] \text{ such that } (x, y) \in U,\ z \in Z\}.$$

As a main result we prove [HH12]

Theorem 6.8. Let $A = U \rtimes Z$ be a semi-cube in H. If $|A| \ge 2^{-1/3} p^{8/3}$ then the four-fold product set $A \cdot A \cdot A \cdot A$ contains at least $|U| \left(1 - \frac{p^4}{\sqrt{2}|A|^{3/2}}\right)$ cosets of the type $[x, y, \mathbb{F}]$.

Bibliography

[AR14] P. Akhilesh, D. S. Ramana, A chromatic version of Lagrange's four squares theorem, Monatshefte fr Mathematik, May 2014

[ASz] M. Ajtai, H. Iwaniec, J. Komlós, J. Pintz, and E. Szemerédi: Construction of a thin set with small Fourier coefficients, *Bull. London Math. Soc.* **22** (1990) 583-590

[BaSh] Sanka Balasuriya and Igor E. Shparlinski, Character sums with subsequence sums, Periodica Mathematica Hungarica Volume 55, Number 2, 215-221

[Be85] V. Bergelson, Sets of reccurence of $\mathbb{Z}^m$-actions and properties of sets of differences, *J. London Math. Soc.* (2) 31 (1985), 295–304

[Be97] V. Bergelson, P. Erdős, N. Hindman, T. Łuczak, Dense difference sets and their combinatorial structure. The mathematics of Paul Erds, I, 165175, Algorithms Combin., 13, Springer, Berlin, 1997.

[BR] V. Bergelson and I. Z. Ruzsa, Squarefree numbers, IP sets and ergodic theory, Paul Erdős and his Mathematics I, Bolyai Society Mathematical Studies, 11, Budapest, 2002, 147-160.

[BS] Vitaly Bergelson and David Simmons: New examples of complete sets, with connections to a diophantine theorem of Frstenberg, arXiv: 1507.02208

[Bi] B.J. Birch: Note on a problem of Erdős, Proc. Camb. Philos. Soc. 55 (1959), p. 370-373

[B] Bourgain, J., More on the sum-product phenomenon in prime fields and its application, Int. J. of Number Theory **1** (2005), 1–32.

[BKT] Bourgain J., Katz N. and Tao T., A sum-product theorem in finite fields and application, Geom. Funct. Anal. **14** (2004), 27–57.

[BEF] T.C. Brown, P. Erdős, and A.R. Freedman, Quasi progressions and descending waves, J. Comb. Th. Ser. A 53 (1990), 8195.

[Ca] J. W. S. Cassels, On the representation of integers as sums of distinct summands taken from a fixed set, Acta Sci. Math. (Szeged) 21 (1960), 111124.

[CFG] Yong-Gao Chen and Jin-Hui Fang, Remark on the completeness of an exponential type sequence,Acta Mathematica Hungarica, Volume 136, Number 3 (2012), 189-195,

[Ch16] Guohua Chen, On monochromatic sums of squares of primes, Journal of Number Theory Volume 162 2016, Pages 180-189

[CFS] D. Conlon, J. Fox and B. Sudakov, Short proofs of some extremal results, Combinatorics, Probability and Computing, Volume 23, Issue 01 (2014), pp 8-28

[DE1] R. Dietmann, C. Elsholtz: Hilbert cubes in progression-free sets and in the set of squares, Israel J.of Math. (2012),

[DE2] R. Dietmann, C. Elsholtz: Hilbert cubes in arithmetic sets, Revista Matemtica Iberoamericana, Vol 31, Issue 4, 2015, pp. 1477-1498

[DES1] R. Dietman, C. Elsholtz, I. E. Shparlinski, On gaps Between Primitive Roots int he Hamming Metric The Quarterly Journal of Mathematics (2012), 1-13

[DES2] R. Dietman, C. Elsholtz, I. E. Shparlinski, Prescribing binary digits of squarefree numbers and quadratic residues, To appear in Transactions of the AMS

[CE] Croot, E. S., III; Elsholtz, C. On variants of the larger sieve. Acta Math. Hungar. 103 (2004), no. 3, 243–254.

[E] P. Erdős, Some of my new and almost new problems and results in combinatorial number theory, Number Theory (Eger, 1996), de Gruyter, Berlin, 1998.

[ERG] P. Erdős and R. L. Graham, Old and new problems and results in combinatorial number theory, Monogr. Enseign. Math. 28 (1980).

[ERe] P. Erdős and A. Rényi, On a new law of large numbers, J. Anal. Math. 22 (1970), 103-111.

[FG] J. H. Fang, A note on the completeness of an exponential type sequence, Chin. Ann. Math. Ser. B. 32 (2011) 527-532

[Fo] E. Følner, Note on a generalization of a theorem of Bogoliuboff, *Math. Scand.* 2 (1954), 224–226

[Ga] M. Garaev: Sums and products of sets and estimates of rational trigonometric sums in fields of prime order, *Russian Math. Surveys* **65:4** 599-658 (2010)

[GK] A.A. Glibichuk, S.V. Konyagin: Additive properties of product sets in fields of prime order, Centre de Recherches Mathmatiques CRM Proceedings and Lecture Notes AMS 2006 (1-8)

[GR] R.L. Graham, B. Rothschild, and J. Spencer: RamseyTheory, Wiley Interscience (1980)

[EG] E. Grosswald, *Representations of integers as sums of squares*, Springer-Verlag, New York, 1985.

[He96] N. Hegyvári, On representation problems in the additive number theory, Acta Math. Hung. 72 (1-2) (1996), 35-44.

[H97] N. Hegyvári, On the dimension of the Hilbert cubes. J. Number Theory 77 (1999), no. 2, 326–330.

[HS99] N. Hegyvári, A. Sárközy, On Hilbert cubes in certain sets. Ramanujan J. 3 (1999), no. 3, 303–314.

[HR] N. Hegyvári, G. Rauzy, On the completeness of certain sequences. Publ. Math. Debrecen 55 (1999), no. 3-4, 245–252.

[He00] N. Hegyvári, On the representation of integers as sums of distinct terms from a fixed set Acta Arith. 92.2 2000. 99-104

[He00b] N. Hegyvári, On the completeness of an exponential type sequence. Acta Math. Hungar. 86 (2000), no. 1-2, 127–135

[HN] N. Hegyvári, Note on difference sets in $\mathbb{Z}^n$ (Period. Math. Hungar. Vol 44 (2), 2002, pp. 183-185

[He04] N. Hegyvári, On Combinatorial Cubes, The Ramanujan Journal, 2004, Volume 8, Issue 3, pp 303307

[He05] N. Hegyvári, On intersecting properties of partitions of integers, Combin. Probab. Comput. (14) 03, (2005), 319-323

[HH07] N. Hegyvári, F. Hennecart,On Monochromatic sums of squares and primes, Journal of Number Theory, Volume 124, Issue 2, 2007, Pages 314-324

[He08] N. Hegyvári, Additive Structure of Difference Sets, seminar Advanced Courses in Mathematics CRM Barcelona, Thematic Seminars Chapter 4 p 253-265

[He08c] N. Hegyvári, IP sets, Hilbert cubes, Publ. Math. Debrecen 72/1-2 (2008), 45-53

[He08b] N. Hegyvári, On additive and multiplicative Hilbert cubes Journal of Combinatorial Theory, Series A 115 (2008) 354360

[He09] N. Hegyvári,On sum-product bases, Ramanujan J. (2009) 19:p 1-8

[HR16] N. Hegyvári, I.Z. Ruzsa, Additive Structure of Difference Sets and a Theorem of Følner, AUSTRALASIAN JOURNAL OF COMBINATORICS Volume 64(3) (2016), Pages 437-443

[CFH] Y.G. Chen s J-H Fang and N. Hegyvári, Ont the subset sums of exponential type sequences, Acta Arithmetca 173 no.2 p.141-150

[HH09] N. Hegyvári, F. Hennecart, Explicit Constructions of Extractors and Expanders Acta Arith. 140 (2009), 233-249.

[HE12] N. Hegyvári, Some Remarks on Multilinear Exponential Sums with an Application, Journal of Number Theory Volume 132, Issue 1, January 2012, Pages 94-102

[HH12] N. Hegyvári, F. Hennecart, A Note on Freiman models in Heisenberg groups Israel Journal, 2012, Volume 189, Issue 1, pp 397-411

[HH13] N. Hegyvári, F. Hennecart,N. Hegyvari and F. Hennecart, A structure result for bricks in Heisenberg groups, Journal of Number Theory 133(9) (2013): 29993006.

[HE16] N. Hegyvári, Note on character sums of Hilbert cubes, Journal of Number Theory Volume 160: pp. 526-535. (2016)

[HHP] N. Hegyvári, F. Hennecart and A. Plagne, Answer to the Burr-Erds question on restricted addition and further results, Combinatorics, Probability and Computing, Volume 16, Issue 05, Sep 2007, pp 747-756

[HHP2] N. Hegyvári, F. Hennecart and A. Plagne, A proof of two Erdos conjectures on restricted addition and further results, J. reine angew. Math. (Crelle) 560 (2003), 199220

[Hil] D. Hilbert, ber die irreducibilität ganzer rationaler Functionen mit ganzzahligen Koeflizienten, J. Reine Angew. Math. 110 (1982), 104-129.

[H79] N. Hindman: Ultrafilters and combinatorial number theory. Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979), pp. 119–184, Lecture Notes in Math., 751, Springer, Berlin, 1979.

[Ka] A. A. Karatsuba: An Estimate of the L_1-Norm of an Exponential Sum, *Mathematical Notes*, **64**, (3), 1998 401-404

[LS00] Łuczak, Tomasz; Schoen, Tomasz On the maximal density of sum-free sets. Acta Arith. 95 (2000), no. 3, 225-229

[MBN] M. B. Nathanson, *Additive number theory. The classical bases*, Graduate Texts in Mathematics, 164. Springer-Verlag, New York, 1996.

[Na] M.B. Nathanson Elementary Methods in Number Theory, Grad. Texts in Math., vol. 195, Springer, 2000.

[RA68] R. Raimi, Translation properties of finite partitions of the positive integers, Fund. Math. 61 (1968) 253-256.

[RR] O. Ramaré, I. Z. Ruzsa, Additive properties of dense subsets of sifted sequences, *J. Théor. Nombres Bordeaux* **13** (2001), 557–581.

[RR12] D. S. Ramana, O. Ramaré, Additive energy of dense sets of primes and monochromatic sums Israel Journal of Math, (2014), Volume 199, Issue 2, pp 955-974

[RSS] J. Rivat, A. Sárközy, and C.L. Stewart, Congruence properties of the -function on sumsets, Illinois J. Math., 43:(1) pp. 1-18. (1999)

[RS] K. R. Roth and G. Szekeres, Some asymptotic formulae in the theory of partitions, Quarterly J. Math. 5 (1954), pp. 241-259.

[Roth] K. F. Roth, On certain sets of integers, *J. London Math. Soc.* **28** (1953), 104–109.

[Ru] I. Z. Ruzsa, Arithmetical progressions and the number of sums, *Period. Math. Hungar.* **25** (1992), 105–111.

[San] T. Sanders, On Roths theorem on progressions, Ann. of Math. 174 (2011), 619636.

[CSS] Sándor, Csaba Non-degenerate Hilbert cubes in random sets. J. Thor. Nombres Bordeaux 19 (2007), no. 1, 249-261

[AS1] A. Sárközy, Finite addition theorems I, *J. Number Theory* **48** (1994), 197–218.

[AS2] A. Sárközy, Unsolved problems in number theory, Period. Math. Hung. 42 (2001), 1735.

[AS3] A. Sárközy, On finite addition theorems III, *Astérisque* **258** (1999), 109–127.

[S] Sárközy, A., On sums and products of residues modulo p, Acta Arith. **118** (2005), 403–409.

[So] T. Schoen, The cardinality of restricted sumsets, preprint (2002).

[SzV] Szemerédi, E.; Vu, V. H. Finite and infinite arithmetic progressions in sumsets. Ann. of Math. (2) 163 (2006), no. 1, 1–35.

[TV] T. Tao, V.H. Vu, Additive combinatorics, p.526, Cambridge University Press, 2006

[TaoEx] T. Tao: Expanding Polynomials over finite fileds of large characteristic, and a regularity lemma for definable sets, Contrib. Disc. Math. Volt 10. n. 1 p 22-98

[T] Tenenbaum G., Sur la probabilité qu'un entier possède un diviseur dans un intervalle donné, Compositio Math. **51** (1984), 243–263.

[LAV] Vinh L.A., Szemerédi–Trotter type theorem and sum-product estimate in finite fields, *European J. Combin.* **32** (2011), 1177-1181.

[W] Woods, Alan R. Subset sum "cubes" and the complexity of primality testing. Theoret. Comput. Sci. 322 (2004), no. 1, 203–219.

Appendix A

Supplement 1

ACTA ARITHMETICA
XCII.2 (2000)

On the representation of integers as sums of distinct terms from a fixed set

by

NORBERT HEGYVÁRI (Budapest)

Introduction. Let A be a strictly increasing sequence of positive integers. The set of all the subset sums of A will be denoted by $P(A)$, i.e. $P(A) = \{\sum \epsilon_i a_i : a_i \in A;\ \epsilon_i = 0 \text{ or } 1\}$. A is said to be *subcomplete* if $P(A)$ contains an infinite arithmetic progression. A natural question of P. Erdős asked how dense a sequence A which is subcomplete has to be. He conjectured that $a_{n+1}/a_n \to 1$ implies the subcompleteness. But in 1960 J. W. S. Cassels (cf. [1]) showed that for every $\varepsilon > 0$ there exists a sequence A for which $a_{n+1} - a_n = o(a_n^{1/2+\varepsilon})$ and A is not subcomplete. In 1962 Erdős [2] proved that if $A(n) > Cn^{(\sqrt{5}-1)/2}$ $(C > 0)$ then A is subcomplete, where $A(n)$ is the counting function of A, i.e. $A(n) = \sum_{a_i \le n} 1$. In 1966 J. Folkman [4] improved this result showing that $A(n) > n^{1/2+\varepsilon}$ $(\varepsilon > 0)$ implies the subcompleteness.

In this note we improve this result. In Section 3 we prove

THEOREM 1. *Let $A = \{0 < a_1 < a_2 < \ldots\}$ be an infinite sequence of integers. Assume that $A(n) > 300\sqrt{n \log n}$ for $n > n_0$. Then A is subcomplete.*

We mention here that $300\sqrt{n \log n}$ cannot be replaced by $\sqrt{2n}$; it is easy to construct a sequence A for which $A(n) > \sqrt{2n}$ and A is not subcomplete.

The main tool for the proof of Theorem 1 is a remarkable theorem of G. Freiman and A. Sárközy (they proved it independently, see [5] and [7]). We are going to use it as Lemma 3.

We use the following notations. The cardinality of the finite set S is denoted by $|S|$. The set of positive integers is denoted by $\mathbb{N}$. $A + B$ denotes

---

2000 *Mathematics Subject Classification*: 11B75, 11A67.
Key words and phrases: subcomplete sequence, additive representations.
Research partially supported by Hungarian National Foundation for Scientific Research, Grant No. T025617, Grant No. T029759 and by DIMACS (Center for Discrete Mathematics and Theoretical Computer Science) NSF-STC-91-19999.

the set of integers that can be represented in the form $a+b$ with $a \in A$, $b \in B$. We write $X_1 + \ldots + X_n = (X_1 + \ldots + X_{n-1}) + X_n$, $n = 3, 4, \ldots$

Acknowledgements. I would like to express my thanks to Prof. G. Freiman for his helpful comments and suggestions.

1. Preliminaries. First we prove

Proposition. *Let $A = \{0 < a_1 < a_2 < \ldots\}$ be an infinite sequence of integers. Assume that $A(n) > 2\sqrt{n \log n}$ for $n > n_0$. Then for every d there exists an $L > 0$ and an infinite sequence $\{y_1 < y_2 < \ldots\}$ in $P(A)$ for which $d \,|\, y_i$ and $y_{i+1} - y_i < L, i = 1, 2, \ldots$*

Proof. $A(n) > 2\sqrt{n \log n}$ implies

$$a_n < \frac{n^2}{\log n}. \tag{1.1}$$

Let $U_i = \{a_{(i-1)d+1} < \ldots < a_{id}\}$. We need some lemmas.

Lemma 1. *If $d \in \mathbb{N}$ and $u_1, \ldots, u_d$ are integers, then there is a sum of the form*

$$u_{i_1} + \ldots + u_{i_t} \qquad (1 \le i_1 < \ldots < i_t \le d)$$

such that $d \,|\, u_{i_1} + \ldots + u_{i_t}$.

Proof. Either there is a k, $1 \le k \le d$, such that $d \,|\, u_1 + \ldots + u_k$ or there are k, m with $k < m$ and $u_1 + \ldots + u_k \equiv u_1 + \ldots + u_m \pmod d$ so that $d \,|\, u_{k+1} + \ldots + u_m$.

By Lemma 1, for every i there exists y_i such that $d \,|\, y_i = a_{i_1} + \ldots + a_{i_t}$, $a_{i_1} < \ldots < a_{i_t}$ and $\{a_{i_1}, \ldots, a_{i_t}\} \subseteq U_i$. Furthermore by (1.1) we get

$$y_i < d a_{id} < d \frac{(id)^2}{\log i} = d^3 \frac{i^2}{\log i}$$

or equivalently

$$Y(n) > \frac{\sqrt{n \log n}}{d^3}, \quad \text{where} \quad Y = \{y_1, y_2, \ldots\}.$$

Now if $y_m = a_{i_1} + \ldots + a_{i_t} = a_{j_1} + \ldots + a_{j_u}$, $\{a_{i_1}, \ldots, a_{i_t}\} \subseteq U_r$, $\{a_{j_1}, \ldots, a_{j_u}\} \subseteq U_s$ for some m and $r < s$ then clearly $u < t \le d$. This implies that if we renumber the elements $y_1, y_2, \ldots$ so that $y_1 \le y_2 \le \ldots$ and $y_i = y_{i+v}$ for some i then $v \le d$. Thus we conclude that there is a sequence $Y^* = \{y_1 < y_2 < \ldots\}$ in $P(A)$ for which $d \,|\, y_i$ and $Y^*(n) \ge Y(n)/d \ge \sqrt{n \log n}/d^4$ or $y_i < d^9 i^2 / \log i$ $(i = 1, 2, \ldots)$.

Lemma 2. *Let $Y = \{y_1 < y_2 < \ldots\}$ be a sequence of positive integers and let $P(Y) = \{s_1 < s_2 < \ldots\}$. Assume that there exists n^* such that for*

dc_1260_16

$n > n^*$ we have

$$y_{n+1} \leq \sum_{i=1}^{n} y_i.$$

Then there is $L > 0$ such that $s_{i+1} - s_i < L$ for every i.

We omit the easy proof (see [6]).

By Lemma 2 the proof of the Proposition will be complete if we check that the sequence Y^* defined in Lemma 1 satisfies the condition $y_{n+1} \leq \sum_{i=1}^{n} y_i$ for large n.

Assume contrary to the assertion that there are infinitely many n for which $y_{n+1} > \sum_{i=1}^{n} y_i$. Then

$$d^9 \frac{(n+1)^2}{\log(n+1)} > y_{n+1} > \sum_{i=1}^{n} y_i \geq \sum_{i=1}^{n} i > \frac{n^2}{2},$$

which is impossible if n is large enough. This proves the Proposition.

2. Arithmetic progressions

Definition. Let $A(d,l) = \{a + kd : 0 \leq k \leq l\}$ be an arithmetic progression.

In this section we prove

Theorem 2. *Let A be an infinite sequence of positive integers. Assume that $A(n) > 200\sqrt{n \log n}$ for $n > n_0$. Then there exists a $\Delta > 0$ such that for every $l \in \mathbb{N}$ there is an arithmetic progression $A(d,l) = \{u + kd : 0 \leq k \leq l\} \subset P(A)$ and $d < \Delta$.*

To prove Theorem 2 we shall use the following important lemma:

Lemma 3. *Let $0 < a_1 < \ldots < a_k \leq n$ be an increasing sequence of integers. Assume that $n > 2500$ and $k > 100\sqrt{n \log n}$. Then there exist integers d, b, z such that $1 \leq d \leq 100\sqrt{n/\log n}$, $z > \frac{1}{7} n \log n$, $b < 7z/\log n$ and*

$$\{sd : b \leq s \leq z\} \subseteq P(\{a_1, \ldots, a_k\}).$$

Lemma 3 is a special case of Theorem 4 in [7].

Now we prove the following

Lemma 4. *Let $A_i := A(D_i, H_i) = \{a_i + tD_i : 0 \leq t \leq H_i\}$ $(i = 1, 2, \ldots)$ be an infinite sequence of arithmetic progressions. Assume that $\lim_{i\to\infty} H_i = \infty$ and*

$$H_i > D_1 + D_{i+1} \tag{2.1}$$

for every $i \geq 1$. Then for every T there is an n for which $A_1 + \ldots + A_n$ contains an arithmetic progression $A(d,h)$ with $d \leq D_1$ and $h > T$.

Thus we are led to construct a long arithmetic progression with bounded difference.

Proof. We shall prove that for every n, $A_1+\ldots+A_n$ contains an $A(d,h)$, where

$$d \le D_1, \qquad h \ge H_n - D_1. \tag{2.2}$$

By the condition $\lim_{i\to\infty} H_i = \infty$, (2.2) completes the proof.

We show (2.2) by induction on n. For $n=1$, (2.2) is trivial. Assume now that $n \ge 2$ and the assertion holds with $1,\ldots,n-1$ in place of n.

By the inductive hypothesis there exists $A(d',h') \subseteq A_1+\ldots+A_{n-1}$ with $d' \le D_1, h' \ge H_{n-1} - D_1$. Since

$$A_1+\ldots+A_n = (A_1+\ldots+A_{n-1}) + A_n \supseteq A(d',h') + A_n$$

it is enough to show that there exists $A(d,h)$ with

$$A(d,h) \subseteq A(d',h') + A_n \quad \text{and} \quad d \le D_1, h \ge H_n - D_1.$$

Let $d = (d', D_n)$ and $u = d'/d, w = D_n/d$. Now $(u,w)=1$. Then

$$\begin{aligned} A(d',h') + A_n &= \{a + td' : 0 \le t \le h'\} + \{a_n + sD_n : 0 \le s \le H_n\} \\ &= \{a + a_n + d(tu+sw) : 0 \le t \le h', 0 \le s \le H_n\}. \end{aligned}$$

It follows from a result of Frobenius (cf. [3]) that if $(u,w)=1$ and if $t \ge w$ then every integer in the interval $[(u-1)(w-1)+1, H_n w]$ can be represented in the form

$$tu + sw, \qquad 0 \le t \le w,\ 0 \le s \le H_n.$$

By (2.1) we infer $h' \ge H_{n-1} > D_n + D_1 \ge D_n/d = w$. Thus by Frobenius' result we get

$$A(d',h') + A_n \supset A(d,h) := \{(a + a_n + duw) + rd : 0 \le r \le H_n w - uw\},$$

where $h = H_n w - uw = (H_n - u)w \ge H_n - u \ge H_n - d'/d \ge H_n - D_1$ and $d \le d' \le D_1$.

This completes the proof of the lemma.

Now define the infinite sequence of integers $[e^{20}] + 1 = n_0 < n_1 < \ldots$ where

$$n_i = n_{i-1}^2, \qquad i = 1, 2, \ldots$$

Let $B_i := (n_{i-1}, n_i] \cap A$. Now $|B_i| = A(n_i) - A(n_{i-1}) > 200\sqrt{n_i \log n_i} - n_{i-1} > 200\sqrt{n_i \log n_i} - \sqrt{n_i} > 100\sqrt{n_i \log n_i}$ since $n_i \ge n_0 = [e^{20}] + 1$. By Lemma 2 there are arithmetic progressions

$$A(D_i, H_i) = \{a_i + kD_i : 0 \le k \le H_i\} \subseteq P(B_i),$$

where

$$D_i \,|\, a_i, \qquad D_i \le 100\sqrt{\frac{n_i}{\log n_i}}, \qquad \frac{1}{8} n_i \log n_i < H_i \tag{2.3}$$

if n_i is large enough. Since $B_i \cap B_j = \emptyset$, for $i \neq j$ we get $A(D_1, H_1) + \ldots + A(D_n, H_n) \subset P(A)$ for every $n \in \mathbb{N}$.

Proof of Theorem 2. In view of Lemma 4 taking the arithmetic progressions $A(D_1, H_1), A(D_2, H_2), \ldots$ given above we have to show that for $i = 1, 2, \ldots,$

$$H_i > D_1 + D_{i+1}.$$

By (2.3),

$$H_i > \frac{1}{8} n_i \log n_i \geq 20e^{10} + 100 \frac{n_i}{\sqrt{\log n_i}} \geq D_1 + D_{i+1}.$$

Thus for every l there is an arithmetic progression $A(D_n, H_n) \subset P(A)$ where $H_n > l$ and $D_n < D_1$.

Theorem 2 is proved.

3. Proof of Theorem 1. Let $B = \{a_{2n-1} : n = 1, 2, \ldots\} \subset A$, $C = A \setminus B$. Now if $n > n_0$ then

$$B(n) \geq 300 \sqrt{\frac{n}{2} \log \frac{n}{2}} \geq 200 \sqrt{n \log n} \quad \text{and} \quad C(n) \geq 200 \sqrt{n \log n}.$$

By Theorem 2 there is a Δ such that for every l there is an arithmetic progression $A(d, l) = \{u + kd : 0 \leq k \leq l\} \subseteq P(B)$ and $d \leq \Delta$. Let $D = \text{l.c.m.}[1, 2, \ldots, [\Delta]]$. By the Proposition there are an L and an infinite sequence $\{x_1 < x_2 < \ldots\}$ in $P(C)$ for which $D \mid x_i$ and $x_{i+1} - x_i < L$ $(i = 1, 2, \ldots)$. Now choose an arithmetic progression $A(d, l)$ contained in $P(B)$, $l > L$. Here $d < \Delta$, thus $d \mid D$ and $d \mid x_i$, $i \in \mathbb{N}$, as well.

We claim $\{kd : (x_1 + u)/d \leq k\} \subset P(A)$. Indeed, let $pd \in [x_j, x_{j+1})$, $x_j > x_1 + u$. This yields that there exists an $i \leq j$ for which $x_1 + u < pd - x_i < u + Ld$.

Now $d \mid x_i$ so $pd - x_i = u + td$, $t < L$. This means $pd = x_i + u + td \in P(A)$.

Theorem 1 is proved.

Addendum (December 8, 1999). I have learned that T. Łuczak and T. Schoen proved a theorem essentially equivalent to my Theorem 1. They obtained their result independently and later.

References

[1] J. W. S. Cassels, *On the representation of integers as sums of distinct summands taken from a fixed set*, Acta Sci. Math. (Szeged) 21 (1960), 111–124.

[2] P. Erdős, *On the representation of large integers as sums of distinct summands taken from a fixed set*, Acta Arith. 7 (1962), 345–354.

[3] P. Erdős and R. L. Graham, *On a linear diophantine problem of Frobenius*, ibid. 21 (1972), 399–408.

[4] J. Folkman, *On the representation of integers as sums of distinct terms from a fixed sequence*, Canad. J. Math. 18 (1966), 643–655.
[5] G. Freiman, *New analytical results in subset-sum problem*, Discrete Math. 114 (1993), 205–218.
[6] R. L. Graham, *Complete sequences of polynomial values*, Duke Math. J. 31 (1964), 275–286.
[7] A. Sárközy, *Finite addition theorems II*, J. Number Theory 48 (1994), 197–218.

ELTE TFK
Eötvös University
Markó u. 29
H-1055 Budapest, Hungary
E-mail: Norb@ludens.elte.hu

Received on 2.4.1996
and in revised form on 21.5.1999 (2956)

Appendix B

Supplement 2

dc_1260_16

Acta Math. Hungar.
86 (1–2) (2000), 127–135.

ON THE COMPLETENESS OF AN EXPONENTIAL TYPE SEQUENCE

N. HEGYVÁRI[1] (Budapest)

Abstract. We investigate the Birch's sequence $Y_K = \left\{ p^\alpha q^\beta \mid p, q > 1,\ \alpha, \beta \in \mathbf{N}_0,\ 0 \leqq \beta \leqq K \right\}$ giving a partial answer for a question of P. Erdös.

1. Introduction

A set A of positive integers is said to be complete if there is an N such that every natural number greater than N is the sum of distinct terms taken from A. Trivially the set $\left\{ p^\alpha \mid \alpha \in \mathbf{N}_0, p > 1 \right\}$ is complete if and only if $p = 2$. A slightly denser sequence than the previous one is $Y = \left\{ p^\alpha q^\beta \mid p, q > 1, \alpha, \beta \in \mathbf{N}_0 \right\}$ and it is a plausible conjecture that Y is complete if and only if $(p, q) = 1$. This was an old conjecture of P. Erdős which was proved by J. Birch [1] in 1959. A few years later J. W. Cassels [2] established a more general theorem.

THEOREM A. *Let A be a sequence of positive integers and let $A(n)$ be its counting function, i.e. let $A(n) = \sum_{\alpha_i \leqq n} 1$. Assume*

$$\lim_{n \to \infty} \frac{A(2n) - A(n)}{\log \log n} = \infty$$

and for every real θ, $0 < \theta < 1$, $\sum_{i=1}^{\infty} \|\alpha_i \theta\| = \infty$. Then A is complete.

It is not too hard to prove that Cassels' theorem is covered by Birch's result.

Nevertheless — as H. Davenport remarked — there is a stronger version of Erdős' conjecture which does not follow from the Cassels' result. He mentioned [1] that it is possible to improve the proof of Birch which gives that for every p, q, $(p, q) = 1$ there exists an integer $K = K(p, q)$ such that the sequence $Y_K = \left\{ p^\alpha q^\beta \mid p, q > 1,\ \alpha, \beta \in \mathbf{N}_0,\ 0 \leqq \beta \leqq K \right\}$ is complete. (For the sequence Y_K we have $|Y_K| < K \cdot \log_p n$ $p > 1$ and so the first condition of

[1] Research partially supported by Hungarian National Foundation for Scientific Research, Grant No. T025617 and No. T029759 and by DIMACS (Center for Discrete Mathematics and Theoretical Computer Science) NSF-STC-91-19999.

Theorem A is not valid.) Indeed, it is not too hard to derive this statement from the Birch's result.

As Erdős mentioned in [6], "of course the exact value of $K(p,q)$ is not known and no doubt will be very difficult to determine".

The aim of this paper is to give an upper bound for $K(p,q)$. We prove

THEOREM. *For every positive integers p,q there exists $K=K(p,q)$ such that the set*

$$Y_K=\left\{p^\alpha q^\beta \,|\, p,q>1,\ \alpha,\beta\in \mathbf{N}_0,\ 0\leqq \beta\leqq K\right\}$$

is complete. Furthermore we have

$$K(p,q)\leqq 2p^{2c^{2^{2q^{4p+3}}}},$$

where $c=1152\log_2 p\log_2 q$.

The basic idea of the proof of the theorem is similar to that developed in [1], although our method and terminology are completely different. Related questions are investigated in [4] and [5].

2. Definitions, notation

Denote by $\mathbf{N}$ and $\mathbf{N}_0$ the set of positive integers and non-negative integers, respectively. For $A,B\subset \mathbf{N}$ and $k\in\mathbf{N}$ denote by $A+B=\left\{a+b\,|\,a\in A;\, b\in B\right\}$ and $kA=\left\{ka\,|\,a\in A\right\}$. Let

$$P(A)=\left\{\sum \varepsilon_i a_i\,|\,\varepsilon_i=0 \text{ or } 1;\ \sum\varepsilon_i<\infty\right\}.$$

Let $A=\{a_1<a_2<\ldots\}\subseteqq \mathbf{N}$; $x,y\in P(A)$. We call (x,y) *disjoint* if there are $X,Y\subset\mathbf{N}$, $X\cap Y=\{\emptyset\}$ and $x=\sum_{i\in X}a_i$; $y=\sum_{j\in Y}a_j$. Call $Z\subset P(A)$ a d-set if the elements of Z are pairwise disjoint. The sets X, Y are disjoint if for every $x\in X$, $y\in Y$ x and y are disjoint.

We call $\sum\varepsilon_{k,s}p^kq^s$ ($\varepsilon_{k,s}=0$ or 1) a representation of n if $n=\sum\varepsilon_{k,s}p^kq^s$. Let us say that p^kq^s is a term of n if $\varepsilon_{k,s}=1$.

We shall use the notation $Y_K=\left\{p^\alpha q^\beta\,|\,p,q>1,\ (p,q)=1,\ \alpha,\beta\in\mathbf{N},\ 0\leqq\beta\leqq K\right\}$ and if the powers of q are even numbers in all terms p^kq^s then we write $Y_{2p,2}=\left\{p^kq^{2m}\,|\,0\leqq k,\ 1\leqq m\leqq 2p\right\}$.

3. Lemmas

LEMMA 1. *Let* $A = \{0 < a_1 < \ldots < a_n < \ldots\}$ *be a sequence of integers. Assume that there is an* n_0 *such that for every* $n > n_0$, $a_n < a_1 + a_2 + \ldots + a_{n-1}$. *Then* $P(A)$ *has bounded gaps, i.e. if* $P(A) = \{x_1 < x_2 < \ldots\}$, *then for every* k *we have* $x_{k+1} - x_k < \Delta$, *where* $\Delta \leqq a_1 + \ldots + a_{n_0}$.

The proof of Lemma 1 is straightforward or see [3]. □

LEMMA 2. *Let* p, q *be positive integers. Let* $Y_{2p,2} = \{p^k q^{2m} \mid 0 \leqq k, 1 \leqq m \leqq 2p\}$. *Then* $P(Y_{2p,2}) = \{x_1 < x_2 < \ldots\}$ *has bounded gaps; in fact, for every* $n, x_{n+1} - x_n \leqq \Delta$, *where* $\Delta \leqq 2q^{4p+2}$.

PROOF. Assume $p^k q^{2m} \in Y_{2p,2}$ for which

$$2q^{4p} < p^k q^{2m}. \tag{1}$$

For brevity let $x := p^k q^{2m}$. We prove that

$$x < \sum_{p^t q^{2s} < x;\ p^t q^{2s} \in Y_{2p,2}} p^t q^{2s}. \tag{2}$$

Thus by Lemma 1 we get that $P(Y_{2p,2})$ has bounded gaps.

Now

$$\sum_{p^t q^{2s} < x;\, p^t q^{2s} \in Y_{2p,2}} p^t q^{2s} = \sum_{s=1}^{2p} q^{2s} \sum_{p^t < x/q^{2s}} p^t = \sum_{s=1}^{2p} q^{2s} \cdot \frac{p^{T+1} - 1}{p-1} \tag{3}$$

where $p^T \leqq \frac{x}{q^{2s}} < p^{T+1}$. By (1) and (3) we have

$$\sum_{p^t q^{2s} < x;\, p^t q^{2s} \in Y_{2p,2}} p^t q^{2s} > \sum_{s=1}^{2p} q^{2s} \frac{x/q^{2s} - 1}{p-1}$$

$$> \sum_{s=1}^{2p} \frac{x - q^{2s}}{p-1} > 2p \cdot \frac{x - q^{2s}}{p-1} > 2p \cdot \frac{x}{2} \cdot \frac{1}{p-1} > x$$

since $x = p^k q^{2m} > 2q^{4p} > q^{2p+1}$. Now we give an upper bound for the biggest gap in $P(Y_{2p,2})$. Let $p^k q^{2m} \in Y_{2p,2}$ be the least element for which (1) holds. Clearly we have $2q^{4p+2} \geqq p^k q^{2m}$ which is an upper bound for the length of the biggest gap of $P(Y_{2p,2})$. □

LEMMA 3. *Let $c, d \geqq 2$ be integers and let $(c,d) = 1$. Let $Y_A = \{c^\alpha d^\beta \mid \alpha \in \mathbf{N},\ 1 \leqq \beta \leqq A := [5\log_2 c] + 1\}$. Let $x \geqq d^{4A}$. Then there is a number n, $1 \leqq n \leqq x$ which has at least two representations $n = \sum_{y\in Y_A} \varepsilon_y y = n = \sum_{y\in Y_A} \varepsilon'_y y$ where $\varepsilon_y, \varepsilon'_y \in \{0,1\}$ and $\sum_{y\in Y_A} \varepsilon_y \cdot \varepsilon'_y = 0$ (i.e. the representations are disjoint).*

PROOF. Since $\sum_{c^k \leqq \sqrt[4]{x}} 1 = \left[\frac{1}{4}\frac{\log_2 x}{\log_2 c}\right]$, $d^A \leqq \sqrt[4]{x}$ and $(c,d) = 1$ we have

$$u := \left|Y_A \cap \left[1, \sqrt{x}\right]\right| > \left[\frac{1}{4}\frac{\log_2 x}{\log_2 c}\right] \cdot A > \frac{1}{5}\frac{\log_2 x}{\log_2 c} \cdot 5\log_2 c = \log_2 x.$$

Furthermore

$$\sum_{y\in Y_A,\ y \leqq \sqrt{x}} y < \left(\sqrt{x}\right)^2 = x.$$

Thus we have $P\left(Y_A \cap \left[1, \sqrt{x}\right]\right) \subset [1, x]$. There are $2^u > x$ subset sums of the form $\sum_{y\in Y_A} \varepsilon_y y$ which implies there are at least two sums which coincide, i.e. $\sum_{y\in Y_A} \varepsilon_y y = \sum_{y\in Y_A} \varepsilon'_y y$. If the representations are not disjoint delete the common terms. $\square$

LEMMA 4. *Let p, q be integers greater than 1, $(p,q) = 1$ and let $g = q^2$. Let $\alpha_1 = \beta_1 = 1$ and for $i > 0$ let*

$$\alpha_{i+1} = [24\log_2 g\alpha_i\beta_i], \quad \beta_{i+1} = [24\log_2 p\alpha_i\beta_i], \quad p_i = p^{\alpha_i}, \quad q_i = g^{\beta_i}.$$

For $i > 0$ let $A_i = [5\log_2 p_i] + 1$. Then for every n there are sets $U_n = \{u_1 < u_2 < \ldots < u_n\}$, $V_n = \{v_1 < v_2 < \ldots < v_n\}$ for which

$$(4)\quad u_i v_i \in P(Y_{A_i}) = P\left(\{p_i^k q_i^m \mid k \in \mathbf{N};\ 1 \leqq m \leqq A_i\}\right),\ v_i - u_i = p^{k_i} g^{m_i};$$

$$u_i, v_i \text{ are disjoint } (i = 1, 2, \ldots, n)$$

and for $1 \leqq i < j \leqq n$,

$$(5)\qquad \{p^{k_j - k_i} g^{m_j - m_i} u_i, p^{k_j - k_i} g^{m_j - m_i} v_i, u_j, v_j\}$$

is a d-set.

PROOF. In the first step we construct elements $u_1, v_1, u_2, v_2, \ldots, u_n, v_n$ for which (4) is true and in the second step we shall show that (5) holds.

Let $i \geqq 1$ and consider $P(Y_{A_i})$. By Lemma 3 we have a number z up to $q_i^{4A_i}$ which has at least two disjoint representations by elements of Y_{A_i}. One of the representations contains at least two terms. Choose one of them and denote it by $p_i^{k'_i} q_i^{m'_i}$ (since $p_i = p^{\alpha_i}$; $q_i = g^{\beta_i}$ we have $k_i = \alpha_i k'_i$ and

$m_i = \beta_i m_i'$). Let now $u_i = z - p^{k_i} g^{m_i}$ and let v_i be the other representation of z. Clearly u_i and v_i are disjoint and so (4) holds.

Now we turn to the proof of (5). We prove it by induction on n. By (4) for $n = 1$ condition (5) is trivial. Let $n > 1$ and assume that the sets $U_{n-1} = \{u_1, \dots, u_{n-1}\}$ and $V_{n-1} = \{v_1, \dots, v_{n-1}\}$ (constructed above) have been defined. By the inductive hypothesis we only have to check that for every i, $1 \leqq i \leqq n$, $A = \{p^{k_n - k_i} g^{m_n - m_i} u_i, p^{k_n - k_i} g^{m_n - m_i} v_i, u_n v_n\}$ is a d-set. Note that $\max\{u_i, v_i\} \leqq \max\{u_{n-1}, v_{n-1}\} \leqq q_{n-1}^{4A_{n-1}}$. Thus if $p^r g^s$ is any term in the representation of u_i or v_i then

$$g^s \leqq p^r g^s \leqq \max\{u_i, v_i\} \leqq q_{n-1}^{4A_{n-1}} \tag{6}$$

and

$$p^r \leqq p^r g^s \leqq \max\{u_i, v_i\} \leqq q_{n-1}^{4A_{n-1}}. \tag{7}$$

By (6), the definition of β_{n-1} and A_{n-1} we have $g^s < g^{4\beta n-1\left([5\log_2 p_{n-1}]+1\right)} < g^{24\log_2 p \cdot \alpha_n - 1\beta_{n-1}}$ and thus

$$s \leqq [24\log_2 p \cdot \alpha_{n-1}\beta_{n-1}] = \beta_n. \tag{8}$$

Furthermore by (7) $p^r < g^{4\beta_{n-1}\left([5\log_2 p_{n-1}]+1\right)}$ and so

$$r \leqq [24\log_2 g \cdot \alpha_{n-1}\beta_{n-1}] = \alpha_n. \tag{9}$$

Assume now contrary to the assertion that A is not a d-set and suppose without loss of generality that $p^{k_n - k_i} g^{m_n - m_i} \cdot u_i$ contains a term which occurs as a term of u_n (the other five cases are similar), i.e. if $p^r g^s$ is a term of u_i then

$$p^{k_n - k_i + r} g^{m_n - m_i + s} = p_n^t q_n^s = p^{\alpha_n t} g^{\beta_n z}, \tag{10}$$

so by $(p, q) = 1$ we have

$$k_n - k_i + r = \alpha_n \cdot t; \quad m_n - m_i + s = \beta_n \cdot z. \tag{11}$$

Recall that $k_n = \alpha_n \cdot C$; $m_n = \beta_n \cdot D$ ($C, D \in \mathbf{N}$), whence

$$r - k_i = \alpha_n(t - C); \quad s - m_i = \beta_i(z - D). \tag{12}$$

Here $1 \leqq r$, $k_i \leqq \alpha_n$; $1 \leqq s$, $m_i \leqq \beta_n$, thus we have $t = C$; $r = k_i$ and $s = m_i$; $z = D$. But as we have defined $u_i, p^{k_i} g^{m_i}$ is not a term of u_i; a contradiction. $\square$

Corollary to Lemma 4. *Let $c_1 = 48\log_2 q$, $c_2 = 24\log_2 p$, $c = c_1 c_2$. Then for every n there exists a d-set*

$$D = \{x_1, y_1, x_2, y_2, \ldots, x_n, y_n\}$$

for which $y_1 - x_1 = y_2 - x_2 = \ldots = y_n - x_n = p^{k_n} q^{2m_n}$, $D \subset P(Y_{K_n})$ where $K_n \leqq 2\beta_{n+1}$. Furthermore we have for $k > 1$

$$\alpha_k \leqq \frac{1}{c_2} c^{2^{k-1}} \quad \textit{and} \quad \beta_k \leqq \frac{1}{c_1} c^{2^{k-1}}. \tag{13}$$

Proof. Let U_n, V_n be the sets defined in Lemma 4. For $1 \leqq i \leqq n$ let $y_i = v_i \cdot p^{k_n - k_i} (q^2)^{m_n - m_i}$; $x_i = u_i \cdot p^{k_n - k_i} (q^2)^{m_n - m_i}$. We get

$$y_i - x_i = (v_i - u_i) \cdot p^{k_n - k_i} (q^2)^{m_n - m_i}$$
$$= p^{k_i} (q^2)^{m_i} \cdot p^{k_n - k_i} (q^2)^{m_n - m_i} = p^{k_n} (q^2)^{m_n}.$$

By (5) we get that D is a d-set. As we have seen in the proof of Lemma 4, $m_n \leqq \beta_{n+1}$. Thus we have $K_n \leqq 2\beta_{n+1}$.

Now we prove (13) by induction on k. For $k = 2$ this is the definition of α_2, β_2. Assume that (13) is true for $k \geqq 2$. By the inductive hypothesis and the definition of α_k and β_k we get

$$\alpha_{k+1} \leqq c_1 \alpha_k \beta_k \leqq c_1 \frac{c^{2^{k-1}}}{c_2} \frac{c^{2^{k-1}}}{c_1} = \frac{c^{2^k}}{c_2}$$

and

$$\beta_{k+1} \leqq c_2 \alpha_k \beta_k \leqq c_2 \frac{c^{2^{k-1}}}{c_2} \frac{c^{2^{k-1}}}{c_1} = \frac{c^{2^k}}{c_1}. \qquad \square$$

Lemma 5. *Let $A = \{0 < a_1 < a_2 < \ldots < a_N < \ldots\}$ be a sequence of integers. Assume*

$$U = \{x_1, x_2, \ldots, x_k, y_1, y_2, \ldots, y_k\} \subset P(A),$$

U is a d-set and for every j, $1 \leqq j \leqq k$, $y_j - x_j = d > 0$ for some fixed d. Then $P(A)$ contains an arithmetic progression of length $k+1$.

Proof. Since U is a d set we have $P(U) \subset P(A)$. Furthermore

$$\left\{ \sum_{i=1}^{k} x_i + \sum_{j=1}^{t} (y_j - x_j) \,:\, 0 \leqq t \leqq k \right\}$$

$$= \left\{ \sum_{i=1}^{k} x_i + td \,: \leqq t \leqq k \right\} \subset P(U) \subset P(A). \qquad \square$$

LEMMA 6. *Let p, q, a, b be positive integers, $(p,q)=1$ and let $T = b + p^a \cdot \phi(p^a)$ where ϕ is the Euler's function. Let*

$$R_T = \left\{ p^r, q^s \,|\, r \in \mathbf{N};\ 1 \leqq s \leqq T \right\}.$$

Then for every r, $1 \leqq r < p^a q^b$ there is an $x_r \in P(R)$ for which $x_r \equiv r \pmod{p^a q^b}$.

PROOF. Let $w_j = p^{a+j\phi(q^b)}$; $z_j = q^{b+j\phi(p^a)}$. Clearly $w_j \equiv p^a \pmod{q^b}$; $z_j \equiv q^b \pmod{p^a}$. Thus we have

$$\sum_{j=1}^{k} w_j + \sum_{i=1}^{t} z_i = M \cdot p^a q^b + kp^a + tq^b \tag{13}$$

for some integer M. Since $(p,q)=1$, for every integer r there are integers k, $k \leqq q^b$, and a t, $t \leqq p^a$ for which $kp^a + tq^b = r$. This yields that for some positive integers k and t'

$$\sum_{j=1}^{k} w_j + \sum_{i=1}^{t} z_i \equiv kp^a + t'q^b \equiv r \pmod{p^a q^b}$$

as we wanted. Clearly the biggest power of q occurring as a term is at most $b + t' \cdot \phi(p^a) \leqq b + p^a \cdot \phi(p^a)$. $\square$

4. Proof of the Theorem

Let $n = 2q^{4p+3}$. By the Corollary to Lemma 4 and by Lemma 5, there is an arithmetic progression of length n and difference $d = p^{k_n} q^{2m_n}$. Furthermore $H = \left\{ h_0 + kd \,|\, k = 0, 1, \dots, n-1 \right\} \subset P(Y_{K_n})$, where

$$K_n \leqq c^{2^n}. \tag{14}$$

Let us note if $p^k q^s$ is a term of any element of H then s is even and $k_n \leqq \alpha_{n+1}$ and $2m_n \leqq 2\beta_{n+1}$.

Let now $Y^* = dqY_{2p,2}$. By Lemma 2 we conclude that the biggest gap in $P(Y^*) = \{x_1 < x_2 < \ldots < x_n < \ldots\}$ is at most $d \cdot 2q^{4p+2}$. Let us observe if $p^k q^s$ is a term of any element of Z^* then s is odd. This yields that $P(Y^*)$ and H are disjoint.

We prove $P(Y^*)+H$ contains an infinite arithmetic progression with difference d, i.e. $\{x_1+h_0+kd \mid k \in \mathbf{N}_0\} \subset P(Y^*)+H$. Let x_s be an element of $P(Y^*)$ for which

$$x_s \leqq h_0 + x_1 + t < x_{s+1}. \tag{15}$$

Now $2q^{4p+3} \cdot d > x_{s+1} - x_s \geqq h_0 + x_1 + t \cdot d - x_s - h_0 + \left(t - \frac{x_x - x_1}{d}\right) \cdot d$. This yields $0 \leqq t - \frac{x_s - x_1}{d} \leqq 2q^{4p+3}$. Thus there exists a z, $z = t - \frac{x_s - x_1}{d}$ and $h_0 + z \in H$. So $h_0 + x_1 = h_0 + t - \frac{x_s - x_1}{d} + x_s = (h_0 + z) + x_s \in P(Y^*) + H$ as we claimed.

Let $a = k_n$, $b = 2m_n$. By Lemma 6, there is a set $P(R_T) = \{x_1, x_2, \dots, x_{d-1}\}$ such that $x_r \equiv r \pmod{d}$, $r = 1, 2, \dots, d-1$ and

$$T \leqq q^{2m_n} + P^{k_n} \cdot \phi(p^{k_n}). \tag{16}$$

By the definition of R_T we have that $P(R_T)$, $P(Y^*)$ and H are disjoint. We claim that $P(R_T) + P(Y^*) + H$ contains every sufficiently large number. But this is trivial; for every r all but finitely many elements of the arithmetic progression $\{r + m \cdot d, m = 0, 1, \dots\}$ belong to $P(R_T) + P(Y^*) + H$, so that every large number belongs to $P(R_T) + P(Y^*) + H$ as well. So we conclude $R_T \cup Y^* \cup U_n \cup V_n$ is complete.

In the rest of the proof we give an upper bound for $K(p,q)$.

Denote by $K_1 = K_1(p,q)$, $K_2 = K_2(p,q)$ and $K_3 = K_3(p,q)$ the greatest s for which $p^k q^s$ is a term of an element of $P(Y^*)$, $P(R_T)$ and H, resp.

1. *An upper bound for* $K_1 = K_1(p,q)$. Since $Y^* = dqY_{2p,2}$ we have that if $p^k q^s \in Y^*$ then $K_1 = \max s \leqq 2m_s + 1$. Recall that $m_n \leqq \beta_{n+1}$ and by (14) we have

$$K_1 \leqq 2\beta_{n+1} + 2p + 1 < 2c^{2^{2q^{4p+3}}} < 3c^{2^{2q^{4p+3}}}.$$

2. *An upper bound for* $K_2 = K_2(p,q)$. By the Corollary of Lemma 4, $K_2 = K_n \leqq 2\beta_{n+1} \leqq 2c^{2^{2q^{4p+3}}}$.

3. *An upper bound for* $K_3 = K_3(p,q)$. By Lemma 6,

$$K_3 \leqq 2m_n + p^{k_n}\phi(p^{k_n}) < 2m_n + 2p^{k_n} < 2c^{2^{2q^{4p+3}}} + p^{2c^{2^{2q^{4p+3}}}} < 2p^{2c^{2^{2q^{4p+3}}}}$$

Since this last upper bound is the biggest one we get

$$K(p,q) \leqq 2p^{2c^{2^{2q^{4p+3}}}}. \qquad \square$$

References

[1] B. J. Birch, Note on a problem of Erdős, *Proc. Cambridge Philos. Soc.*, **55** (1959), 370–373.

[2] J. W. Cassels, On the representation of integers as the sums of distinct summands taken from a fixed set, *Acta Sci. Math. (Szeged)*, **21** (1960), 111–124.

[3] R. L. Graham, Complete sequences of polynomial values, *Duke Math. J.*, **32** (1964), 275–286.

[4] S. Burr, P. Erdős, R. L. Graham and W. Wen-Ching Li, Complete sequences of sets of integer powers, *Acta Arithm.*, (1996).

[5] N. Hegyvári and G. Rauzy, On the completeness of certain sets (to appear in *Publ. Math. Debrecen*).

[6] P. Erdős, Some problems and results in additive number theory (reprint).

(Received May 27, 1998; revised June 18, 1999)

ELTE TFK
DEPARTMENT OF MATHEMATICS
UNIVERSITY EÖTVÖS LORÁND
BUDAPEST, H–1055 MARKÓ U. 29
HUNGARY