Russian Math. Surveys **61**:6 1101–1166 ©2006 RAS(DoM) and LMS

Uspekhi Mat. Nauk **61**:6 111–178 DOI 10.1070/RM2006v061n06ABEH004370

Szemerédi's theorem and problems on arithmetic progressions

I. D. Shkredov

Abstract. Szemerédi's famous theorem on arithmetic progressions asserts that every subset of integers of positive asymptotic density contains arithmetic progressions of arbitrary length. His remarkable theorem has been developed into a major new area of combinatorial number theory. This is the topic of the present survey.

Contents

§ 1. Introduction

Let k and d be positive integers. By an arithmetic progression of length k with difference d one means a set of the form $n, n+d, n+2d, \dots, n+(k-1)d$, where n is an integer. In 1927 B. L. van der Waerden proved his famous theorem on arithmetic progressions (see [1]), praised by A. Ya. Khinchin [Khinchine, Khintchine, Hinchin] as a pearl of number theory (see [2]).

Theorem 1 (van der Waerden). *Let h and k be positive integers. For any partition of the integers into h subsets $C_1, \dots, C_h$, one of these subsets contains an arithmetic progression of length k.*

This research was supported by grants from the Russian Foundation for Basic Research (no. 06-01-00383), the President of the Russian Federation (no. 1726.2006.1), and INTAS (no. 03-51-5070).

AMS 2000 *Mathematics Subject Classification.* Primary 11B25; Secondary 05D10, 28D05, 28D15.

Although van der Waerden's theorem seems simple and natural, it has played a significant role in the development of two directions in mathematics: additive combinatorics and combinatorial ergodic theory. We shall say more about these directions below in our survey, but for now we note only that the two areas are very intimately connected and form a meeting point of disciplines like additive and analytic number theory, graph theory, and the theory of dynamical systems. Van der Waerden's theorem itself is a fundamental result of Ramsey theory (see [3], [4]). Indeed, if one treats the partition of the set of integers in Theorem 1 into h subsets $C_1, \ldots, C_h$ as a colouring of $\mathbb{Z}$ with h distinct colours, then van der Waerden's theorem asserts that the set of integers contains a *monochromatic* arithmetic progression, that is, a progression whose elements all have the same colour.

The present survey is devoted to a variety of results related in some way to van der Waerden's theorem and its generalizations.

Before discussing the generalizations, we reformulate Theorem 1.

Theorem 2. *Let h and k be positive integers. There is a number $N(h,k)$ such that, for any positive integer $N \geqslant N(h,k)$ and any partition of the set $1,\ldots,N$ into h subsets, at least one of these subsets contains an arithmetic progression of length k.*

In contrast to Theorem 1, all the sets in Theorem 2 are finite. For this reason, Theorem 2 is referred to as a *finite* version of Theorem 1. One can easily show that these results are equivalent (see below).

Apparently, the simplest question arising in connection with van der Waerden's theorem is: how rapidly does $N(h,k)$ tend to infinity? Unfortunately, van der Waerden's original proof gives very weak estimates for $N(h,k)$, even when $h=2$. We shall formulate the existing result more precisely.

We introduce a sequence of functions $f_i\colon \mathbb{N}\to\mathbb{N}$ (the Ackermann hierarchy). Let $f_1(n)=n+1$ and let $f_{i+1}(n)=\underbrace{(f_i\circ\cdots\circ f_i)}_{n}(1)$ for any $i\geqslant 2$. Then, for example, $f_2(n)=2n$, $f_3(n)=2^n$, and $f_4(n)$ is the tower of n twos. By the *Ackermann function* one means the function $A(n)=f_n(n)$, $n\in\mathbb{N}$. It is clear that $A(n)$ tends to infinity more rapidly than any fixed function $f_i(n)$. Moreover, $A(n)$ is not primitive recursive (roughly speaking, $A(n)$ cannot be expressed by using finitely many compositions of ordinary functional operations; for details, see [5]). The original proof of van der Waerden's theorem implies the estimate $N(2,k)\leqslant A(k)$ for any $k\geqslant 2$.

In 1987 S. Shelah obtained the first primitive recursive estimate for $N(2,k)$ (see [6]). Let $S(1)=2$ and let $S(n)=f_4(S(n-1))$ for any $n\geqslant 2$. Then the inequality $N(2,k)\leqslant S(Ck)$ holds for any $k\geqslant 2$, where C is an absolute constant.

Another generalization related to Theorem 1 was expressed in 1936 by P. Erdős and P. Turán.

Let A be an arbitrary subset of integers. By the *Banach upper density* (or simply *upper density*) of a set A we mean the quantity

$$D^*(A)=\limsup_{N\to\infty}\frac{|A\cap\{1,2,\ldots,N\}|}{N}.$$

Erdős and Turán conjectured that an arbitrary set of integers of positive upper density contains an arithmetic progression of arbitrary length. It is clear that this conjecture implies van der Waerden's theorem. Indeed, the Banach upper density has the subadditivity property

$$D^*\left(\bigcup_{i=1}^{m} A_i\right) \leqslant \sum_{i=1}^{m} D^*(A_i),$$

where $A_i \subseteq \mathbb{Z}$, $i = 1, \dots, m$, are arbitrary sets. Hence, if $\mathbb{Z}$ is partitioned into h subsets, then the upper density of one of these subsets is at least $1/h > 0$. Let us state the Erdős–Turán conjecture more precisely.

Conjecture 1 (Erdős, Turán). *Let $A \subseteq \mathbb{Z}$ be an arbitrary set with $D^*(A) > 0$. Then A contains an arithmetic progression of length k for any integer $k \geqslant 3$.*

It follows from Conjecture 1 that every set of positive density contains infinitely many arithmetic progressions of length k.

The Erdős–Turán conjecture has another (equivalent) formulation. For convenience, we denote the set $\{1, 2, \dots, N\}$ by $[N]$.

Conjecture 1′ (Erdős, Turán). *Let $k \geqslant 3$ be an integer and let $0 < \delta \leqslant 1$. Then there is a positive integer $N(k, \delta)$ such that for any $N \geqslant N(k, \delta)$ an arbitrary set $A \subseteq [N]$ with $|A| \geqslant \delta N$ contains an arithmetic progression of length k.*

For completeness we prove here the equivalence of Conjectures 1 and 1′. More or less the same reasoning proves the equivalence of Theorems 1 and 2.

Obviously, Conjecture 1′ implies Conjecture 1. Let us prove the converse. Suppose that Conjecture 1′ fails for some integer $k \geqslant 3$ and some $\delta \in (0, 1]$. In other words, for any positive integer N there is a set $A \subseteq [N]$ with $|A| \geqslant \delta N$ such that A contains no arithmetic progressions of length k. Let $N_1 = 1$ and $b_1 = 0$ and let

$$N_i := b_{i-1} + N_{i-1}, \qquad b_i := b_{i-1} + N_{i-1} + N_i + 1, \tag{1}$$

for $i \geqslant 2$. We obtain an increasing sequence of positive integers $1 = N_1 < N_2 < N_3 < \cdots$ and a sequence of sets $A_1, A_2, A_3, \dots$ such that $A_i \subseteq [N_i]$, $|A_i| \geqslant \delta N_i$ for any i, and none of the sets A_i contain arithmetic progressions of length k. Let $\widetilde{A}_i = A_i + b_i$. It is clear that the sets $\widetilde{A}_i$ are disjoint and do not contain arithmetic progressions of length k. Let $A = \bigsqcup_i \widetilde{A}_i$. Using (1), we see that A also does not contain arithmetic progressions of length k. We have $b_i \leqslant 3N_i$ for any $i \geqslant 1$. Further,

$$\frac{|A \cap [b_i + N_i]|}{b_i + N_i} \geqslant \frac{|\widetilde{A}_i \cap [b_i, b_i + N_i]|}{4N_i} \geqslant \frac{\delta N_i}{4N_i} = \frac{\delta}{4}\,. \tag{2}$$

It follows from (2) that the upper density of the set A is at least $\delta/4 > 0$. This contradicts Conjecture 1.

The Erdős–Turán conjecture turned out to be extremely difficult. The simplest case $k = 3$ was proved by K. F. Roth only in 1953 (see [7]). It should be noted that the case of progressions of length three is special, because in this situation one can use the more or less familiar technique connected with the circle method.

We reformulate the Erdős–Turán conjecture once more.

Let N be a positive integer. We set

$$a_k(N) = \frac{1}{N} \max\{|A| : A \subseteq [N], A \text{ contains no arithmetic progressions of length } k\}.$$

We make a remark concerning the function $a_k(N)$. Let $k \geqslant 3$ be an integer, let N and M be arbitrary positive integers, and let A be a subset of $[N+M]$ without arithmetic progressions of length k. Then the sets $A_1 = A \cap \{1, \dots, N\}$ and $A_2 = A \cap \{N+1, \dots, N+M\}$ also do not contain arithmetic progressions of length k. We obtain the obvious inequality $a_k(N+M)(N+M) \leqslant a_k(N)N + a_k(M)M$, which implies the existence of the limit $\lim_{N\to\infty} a_k(N)$. The Erdős–Turán conjecture means that for any $k \geqslant 3$ we have

$$a_k(N) \to 0 \quad \text{as} \quad N \to \infty. \tag{3}$$

Roth proved the following theorem, where log denotes the logarithm to base 2.

Theorem 3 (Roth). *Let $N \geqslant 3$ be an integer. Then*

$$a_3(N) \ll \frac{1}{\log\log N}.$$

Thus, Roth obtained more than the Erdős–Turán conjecture needed for $k = 3$. His theorem gives a quantitative estimate for the rate of vanishing of $a_3(N)$. We present the proof of Theorem 3 in § 2.

Roth's result was later improved by E. Szemerédi in [8] and D. R. Heath-Brown in [9]. These authors independently obtained the following estimate for $a_3(N)$.

Theorem 4 (Szemerédi, Heath-Brown). *Let $N \geqslant 3$ be an integer. Then*

$$a_3(N) \ll \frac{1}{(\log N)^c},$$

where for the constant c one can take $1/20$.

At present, the best result on an upper bound for $a_3(N)$ is due to J. Bourgain [10] (see also his paper [11] on subsets of $\mathbb{R}^k$ containing no arithmetic progressions).

Theorem 5 (Bourgain). *Let $N \geqslant 3$ be an integer. Then*

$$a_3(N) \ll \sqrt{\frac{\log\log N}{\log N}}. \tag{4}$$

The theorems of Roth, Szemerédi, Heath-Brown, and Bourgain involve estimates of $a_3(N)$ (see also the interesting paper [12], in which the Erdős–Turán conjecture for $k = 3$ is proved by using methods of graph theory). As was noted above, the conjecture in (3) is much more simple for $k = 3$ than in the case $k \geqslant 4$. If $k \geqslant 4$, then the usual analytic methods are no longer effective. It was only in 1969 that Szemerédi proved the Erdős–Turán conjecture in the case $k = 4$ (see [13]), and then in 1975 he obtained a complete solution of this problem for any $k \geqslant 4$ (see [14]). Let us formulate this beautiful result.

Theorem 6 (Szemerédi). *Let A be an arbitrary subset of the positive integers with $D^*(A) > 0$. Then A contains an arithmetic progression of length k for any integer $k \geqslant 3$.*

In his proof Szemerédi uses complicated combinatorial arguments. The proof is based on the so-called regularity lemma, which is at present the most important tool for studying graphs. Szemerédi's proof is discussed in more detail in § 4.

An alternative proof of Theorem 6 was proposed by Furstenberg in [15] (a simpler proof is presented in [16]). His approach uses ergodic theory methods. Furstenberg showed that Szemerédi's theorem is equivalent to an assertion on multiple recurrence for almost all points in an arbitrary dynamical system.

Theorem 7 (Furstenberg). *Let X be a set, let $\mathscr{B}$ be a σ-algebra of measurable sets on X, and let μ be a finite measure on X with $\mu(X) > 0$. Suppose that T is a map of X into itself which preserves the measure μ and that E is an arbitrary measurable subset of X with $\mu(E) > 0$.*[1] *Then there is an integer $n > 0$ such that*

$$\mu(E \cap T^{-n}E \cap T^{-2n}E \cap \cdots \cap T^{-(k-1)n}E) > 0.$$

We discuss Furstenberg's theorem in § 5.

It should be noted that, using his method, Furstenberg and his students obtained a series of deep generalizations of Szemerédi's theorem (see, for instance, [17]–[21]) which have not yet been proved by combinatorial methods.

We cite here only one of the results in the paper [19].

Theorem 8 (Bergelson, Leibman). *Let X be a set, let $\mathscr{B}$ be a σ-algebra of measurable sets on X, and let μ be a finite measure on X with $\mu(X) > 0$. Let $k \geqslant 2$, let $T_1, \ldots, T_k$ be invertible commuting self-maps of X that preserve the measure μ, and let $p_1(n), \ldots, p_k(n)$ be polynomials having rational coefficients, taking integral values at all integral points n, and satisfying $p_i(0) = 0$, $i = 1, \ldots, k$. Then*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(T_1^{-p_1(n)}E \cap T_2^{-p_2(n)}E \cap \cdots \cap T_k^{-p_k(n)}E) > 0$$

for any measurable set E with $\mu(E) > 0$.

Theorem 8 shows that the linear functions $n, 2n, 3n, \ldots, (k-1)n$ in Theorem 7 can be replaced by arbitrary integral polynomials $p_1(n), p_2(n), \ldots, p_k(n)$ satisfying the condition $p_i(0) = 0$, $i = 1, \ldots, k$.

Unfortunately, Szemerédi's methods give very weak upper bounds for $a_k(N)$. The ergodic approach gives no bounds at all. It was only in 2001 that W. T. Gowers [22] obtained the first effective result on the rate of vanishing of the quantity $a_k(N)$ for $k \geqslant 4$. For a weaker estimate of $a_4(N)$, see [23].

Theorem 9 (Gowers). *The inequality $a_k(N) \ll 1/(\log\log N)^{c_k}$ holds for all integers $N \geqslant 3$ and $k \geqslant 4$, where the constant is $c_k = 2^{-2^{k+9}}$.*

[1] *Russian Editors' note*: It is also assumed that k is a given integer with $k \geqslant 3$; see Theorem 25 below.

Corollary 1. *Let k and N be positive integers and let the set $[N]$ be coloured with at most $(\log\log N)^{c_k}$ colours. Then the set $[N]$ contains a monochromatic arithmetic progression of length k.*

Although the proofs of the theorems of Szemerédi, Furstenberg, and Gowers use different methods, the theorems have much in common. In each of these approaches the proof is an iterative procedure, and the crucial idea is in the dichotomy between the structure and the randomness. More precisely, at each step of the iterative procedure (of an algorithm proving the theorem) we test our object X of interest (a set or a dynamical system) to find out whether it has 'random' properties. The structure of the iterative procedure has a specific feature: if X has random properties (for example, if the system X is a weakly mixing dynamical system or if the set X has 'small' Fourier coefficients), then one can rather easily establish the existence of arithmetic progressions in X. However, if X has no random properties, then some part of X (a subset or a quotient system) has some 'structural properties'. Applying our considerations to this part of X, we either prove the desired result by finding in X a subobject with random properties, or single out a more and more 'structurized' subobject of X at every step of the iterative procedure, so that it eventually becomes quite clear that the subobject has configurations we are interested in.

Gowers' result is a significant step toward the proof of another famous conjecture of Erdős and Turán on arithmetic progressions.

Conjecture 2 (Erdős, Turán). *Let $A = \{n_1 < n_2 < \cdots\}$ be an infinite sequence of positive integers such that*

$$\sum_{i=1}^{\infty} \frac{1}{n_i} = \infty.$$

Then A contains an arithmetic progression of arbitrary length.

It can readily be proved (and will be proved in § 5) that Conjecture 2 is equivalent to the condition that the series $\sum_{l=1}^{\infty} a_k(4^l)$ converges for any integer $k \geqslant 3$. Hence, to prove Conjecture 2, it suffices to obtain the estimate $a_k(N) \ll 1/(\log N)^{1+\varepsilon}$ for any $k \geqslant 3$ and for some $\varepsilon > 0$.

The proof of Theorem 9 contains many new and beautiful ideas. The methods in Gowers' paper were developed by several authors (see, for instance, [24]–[33]). The most striking result obtained in these papers is undoubtedly the theorem of B. Green and T. Tao on progressions in the primes.

Theorem 10 (Green, Tao). *For any integer $k \geqslant 3$ the set of primes contains an arithmetic progression of length k.*

In fact, Green and Tao proved a stronger result.

Let A be an arbitrary subset of the set $\mathscr{P}$ of primes and let $\pi(N)$ be the number of primes not exceeding N. The upper density of A with respect to $\mathscr{P}$ is defined to be $\limsup_{N\to\infty} |A \cap [N]|/\pi(N)$.

Theorem 11 (Green, Tao). *Let $A \subseteq \mathscr{P}$ be an arbitrary set of positive upper density with respect to $\mathscr{P}$ and let $k \geqslant 3$. Then A contains an arithmetic progression of length k.*

Theorem 11 for $k = 3$ was proved by Green in [34]. He proved an even stronger result. We write $\log_{[1]} = \log N$ and let $\log_{[l]} N = \log(\log_{[l-1]} N)$ for any $l \geqslant 2$. Thus, $\log_{[l]} N$ is the iterated logarithm (the result of taking the logarithm of the number N successively l times).

Theorem 12 (Green). *Let N be a sufficiently large positive integer and let A be an arbitrary subset of $\mathscr{P} \cap [N]$ such that*

$$|A| \gg \frac{N\sqrt{\log_{[5]} N}}{\log N \sqrt{\log_{[4]} N}}.$$

Then A contains an arithmetic progression of length three.

We derive a simple consequence of Theorem 11 (see [24]). As is known (see, for instance, [35]), the set A_1 of primes whose residue modulo four is equal to one has positive upper density with respect to $\mathscr{P}$. Moreover, every element of A_1 can be represented as the sum of two squares (see, for example, [36], Ch. V, Question 9c). Applying Theorem 11 to A_1, we see that there is an arithmetic progression of arbitrary length all of whose elements are sums of two squares.

In conclusion we note that Conjecture 2 implies both Theorem 10 and Theorem 11.

Let us briefly describe the structure of the survey. In § 2 we consider the simplest case $k = 3$ of Conjecture 1′ and prove Roth's theorem, Theorem 13. In § 3 we present results of F. A. Behrend, R. A. Rankin, and others on lower bounds for the quantity $a_k(N)$. The next section, § 4, is devoted to Szemerédi's theorem, simple corollaries to it, and the regularity lemma. In § 5 we discuss the ideas at the basis of the proof of Gowers' result in the case $k = 4$ and also properties of the Gowers norm that are used in the proof of the general Theorem 9. Since the proof of Theorem 9 is very complicated, in our presentation we follow a simpler paper [23] in which Gowers proved a weaker estimate for $a_4(N)$ than in Theorem 9. In § 6 we sketch Furstenberg's proof of Szemerédi's theorem and give a brief survey of results obtained by methods of ergodic combinatorial number theory. In the next section we consider the simplest two-dimensional generalization of Theorem 6, and in § 8 the Green–Tao result (Theorem 10) on progressions in the primes. In § 9 we discuss further generalizations of Roth's theorem and also of Schur's theorem (see [37]). The main results in this area of combinatorial number theory were obtained by Roth [38], R. Rado [39]–[41], and also by P. Frankl, R. L. Graham, and V. Rödl [42]. In § 10 we present two theorems of E. Croot on critical sets without arithmetic progressions, several results on arithmetic progressions in sums, and a theorem on rainbows. Finally, we discuss in conclusion several unsolved problems related to Szemerédi's theorem and arithmetic progressions.

We finish the Introduction with a proof of van der Waerden's theorem (see [43]).

Let $k \geqslant 1$ and $r \geqslant 0$ be integers. We denote the arithmetic progression a, $a + r, \ldots, a + (k-1)r$ by the symbol $[a, r, k]$. Let $\mathbf{c}\colon [N] \to [m]$ be a colouring of the segment $1, \ldots, N$ of positive integers with m colours.

Definition 1. By a *fan with radius* k, *degree* d, *and initial point* a we mean a d-tuple $([a, r_1, k], \dots, [a, r_d, k])$ of arithmetic progressions belonging to $[N]$. Each progression of the form $[a + r_i, r_i, k - 1]$, $1 \leqslant i \leqslant d$, is called a *spoke* of the fan.

A fan $([a, r_1, k], \dots, [a, r_d, k])$ is said to be *polychromatic* if there are $d+1$ distinct colours $c_0, c_1, \dots, c_d$ such that the initial point a has colour c_0 and all the elements of the ith spoke of the fan have colour c_i, $i = 1, \dots, d$.

Proof of Theorem 2. The proof can be carried out by induction on the length k of the arithmetic progression. If $k = 1$, then van der Waerden's theorem obviously holds, and we therefore assume that $k \geqslant 2$. Suppose that the theorem has been proved for $k-1$. In other words, for any positive integer m there is a positive integer $N(m, k-1)$ such that any colouring of $[N(m, k-1)]$ with m colours contains a monochromatic arithmetic progression of length $k-1$.

We use our induction hypothesis to prove the following assertion: for any $d \geqslant 0$ there is a positive integer $N_f(m, k-1, d)$ such that every colouring of $[N(m, k-1, d)]$ contains either a monochromatic arithmetic progression of length k or a polychromatic fan of radius k and degree d; denote this assertion by N_f. We prove N_f by induction on d. If $d = 0$, then the assertion N_f is obvious. We note that a polychromatic fan of degree d can exist only if we have at least $(d + 1)$ colours. Hence, if our assertion is proved for $d = m$, then van der Waerden's theorem will follow.

Suppose that the validity of the assertion N_f has been proved for $d - 1$, $d > 1$. Let $N = N_f(m, k-1, d) := 4kN_1N_2$, where $N_1 = N_f(m, k-1, d-1)$ and $N_2 = N(m^d N_1^d, k-1)$. We note that the existence of the numbers N_1 and N_2 follows from our induction hypotheses. Let $\mathbf{c}$ be an arbitrary colouring of the set $[N]$ with m colours. For any b in$[N_2]$ the set $\{bkN_1 + 1, bkN_1 + 2, \dots, bkN_1 + N_1\}$ is an arithmetic progression of length N_1 belonging to the segment $[N]$ of the positive integers. By the induction hypothesis, the set $\{bkN_1 + 1, bkN_1 + 2, \dots, bkN_1 + N_1\}$ contains either a monochromatic arithmetic progression of length k or a polychromatic fan of radius k and degree $d - 1$. If we have a monochromatic progression of length k, then we have proved van der Waerden's theorem; therefore, we assume that for any $b \in [N_2]$ there is a polychromatic fan of radius k and degree $d - 1$. In other words, for any $b \in [N_2]$ there are some elements $a(b), r_1(b), \dots, r_{d-1}(b) \in \{1, \dots, N_1\}$ and distinct colours $c_0(b), c_1(b), \dots, c_{d-1}(b) \in [m]$ such that $\mathbf{c}(bkN_1 + a(b)) = c_0(b)$ and $\mathbf{c}(bkN_1 + a(b) + jr_i(b)) = c_i(b)$ for any $j = 1, \dots, k-1$, $i = 1, \dots, d-1$. The map $b \to (a(b), r_1(b), \dots, r_{d-1}(b), c_0(b), c_1(b), \dots, c_{d-1}(b))$ determines a colouring of $[N_2]$ with $m^d N_1^d$ colours. We can assume that these colours are indexed by the numbers from 1 to $m^d N_1^d$. By the definition of the number N_2, there is a monochromatic arithmetic progression $[b, s, k-1]$ (of length $k-1$) belonging to $[N_2]$. Suppose that this progression has the colour $(a, r_1, \dots, r_{d-1}, c_0, c_1, \dots, c_{d-1})$. Without loss of generality we can assume that $s > 0$, because otherwise we can consider the arithmetic progression $b, b - s, \dots, b - (k-2)s$ instead of $[b, s, k-1]$.

Let $b_0 = (b-s)kN_1 + a$. Since $N = 4kN_1N_2$, it follows that $b \in [N]$. We consider the fan of radius k and degree d with initial point b_0:

$$([b_0, skN_1, k], [b_0, skN_1 + r_1, k], \dots, [b_0, skN_1 + r_{d-1}, k]). \tag{5}$$

Let us prove that the fan (5) is polychromatic. Consider the first spoke of the fan. For $j = 1, \dots, k-1$,

$$\mathbf{c}(b_0 + jskN_1) = \mathbf{c}((b + (j-1)s)kN_1 + a) = c_0(b + (j-1)s) = c_0.$$

Hence, the first spoke of the fan is monochromatic. The other spokes of the fan are also monochromatic. Indeed, for any $j = 1, \dots, k-1$, $t = 1, \dots, d-1$ we have

$$\mathbf{c}(b_0 + j(skN_1 + r_t)) = \mathbf{c}((b + (j-1)s)kN_1 + a + jr_t) = c_t(b + (j-1)s) = c_t.$$

If some spoke of the fan is coloured with the colour of the initial point b_0, then we have found a monochromatic arithmetic progression of length k. If the colour of b_0 differs from the colour of every spoke of the fan, then we obtain a polychromatic fan of radius k and degree d. In other words, we have proved the assertion N_f, and hence van der Waerden's theorem as well.

§ 2. Roth's theorem

In the present section we prove Roth's theorem, Theorem 3. Let us formulate this beautiful theorem once more.

Theorem 13 (Roth). *Let $\delta > 0$ and let N be an integer with $N \gg \exp\exp(\delta^{-1})$. Assume that A is an arbitrary subset of $[N]$ with $|A| = \delta N$. Then the set A contains an arithmetic progression of length three.*

Before sketching the proof of Theorem 13, we shall give several definitions.

Let f be a complex function on the set of integers, and let E be an arbitrary subset of $\mathbb{Z}$. We write $f\colon E \to \mathbb{C}$ if the function f vanishes outside E.

Suppose that the function f takes finitely many non-zero values. The Fourier transform of the function f is given by the formula

$$\widehat{f}(x) = \sum_n f(n)e^{-2\pi inx}, \qquad x \in \mathbb{S}^1. \tag{6}$$

(The numbers $\widehat{f}(x)$ are also called the Fourier coefficients of the function f.) We write $\int$ instead of $\int_0^1$ and $\sum_n$ instead of $\sum_{n\in\mathbb{Z}}$. The following formulae hold:

$$\int |\widehat{f}(x)|^2\,dx = \sum_n |f(n)|^2, \tag{7}$$

$$\int \widehat{f}(x)\overline{\widehat{g}(x)}\,dx = \sum_n f(n)\overline{g(n)}. \tag{8}$$

Definition 2. Let $\alpha \in (0,1)$ and let N be a positive integer. A function $f\colon [N] \to \mathbb{C}$ is said to be *α-uniform* if

$$\|\widehat{f}\|_\infty \leqslant \alpha N. \tag{9}$$

Let $A \subseteq [N]$, $|A| = \delta N$. In this paper the characteristic function of A will be denoted by χ_A, or simply by the same symbol A. The function $f = \chi_A - \delta\chi_{[N]}$ is called the *balance* function of A. A set A is said to be *α-uniform* if its balance function is α-uniform. We note that $\sum_n f(n) = 0$.

The simplest examples of α-uniform sets are the so-called random sets. We dwell on these sets in a bit more detail. Let $\delta > 0$ be an arbitrary number. Let Ω be the space of sequences of length N formed by zeros and ones and let $\mathscr{F}$ be the σ-algebra of all subsets of Ω. If a sequence $\omega \in \Omega$ has k ones and $N-k$ zeros, then we assign the probability $\delta^k(1-\delta)^{N-k}$ to it. This defines a probability space denoted by $(\Omega, \mathscr{F}, \mathsf{P})$. To any point $\omega = (\omega_1, \dots, \omega_N)$ of the set Ω we assign the set $A(\omega) = \{i : \omega_i = 1\}$. The random sets are the sets of the form $A(\omega)$ for the 'generic' sequences ω. For example, one can easily see that the expectation of the cardinality of the set $A(\omega)$ is equal to δN. Proceeding from this fact, one says that the cardinality of the random set in the above probability model is equal to δN. Similarly, one can find the expectation of the number of arithmetic progressions in the random set. This expectation turns out to be of the order of $\delta^3 N^2$. Using large-deviation estimates for sequences of independent identically distributed random variables, one can also show that all the non-zero Fourier coefficients of the random set do not exceed $N^{1/2+\varepsilon}$, where $\varepsilon > 0$ (for more details, see, for instance, [44]). Hence, the random sets are examples of α-uniform sets for $\alpha > N^{-1/2+\varepsilon}$. One can say that, in a sense, 'almost all' subsets of $[N]$ of cardinality δN are α-uniform. Nevertheless, one can easily give examples of sets in $[N]$ of cardinality δN which are not α-uniform. For instance, one can take an arbitrary arithmetic progression in $[N]$ of length δN.

We discuss the main ideas used in the proof of Roth's theorem. Suppose that a set $A \subseteq [N]$, $|A| = \delta N$, contains no arithmetic progressions of length three. The proof of Roth's theorem is an algorithm. At the first step of this algorithm, two situations are possible: either the set A is α-uniform for some α depending only on δ (namely, the proof takes $\alpha = 2^{-5}\delta^2$), or it is not.

If A is α-uniform, then one can readily show that A contains very many arithmetic progressions of length three. More precisely, the order of the number of arithmetic progressions of length three in A is equal to that of $\delta^3 N^2$. We note that the random sets contain exactly this number of arithmetic progressions. The set A also contains the so-called trivial (or 'degenerate') progressions, that is, progressions with zero difference. It is clear that the number of such progressions does not exceed the cardinality of A. By assumption, the number N is not less than $\exp\exp(\delta^{-1})$. This leads to the conclusion that if A is α-uniform, then the number of arithmetic progressions in A is greater than $|A| = \delta N$. Hence, A contains at least one non-trivial progression (in fact, there are many such progressions in A). We arrive at a contradiction to the assumption that there are no arithmetic progressions of length three in A.

Suppose now that the set A is not α-uniform. One can show that the α-uniform condition for the set A is equivalent to the uniform distribution of A in long arithmetic progressions. More precisely, the cardinality of the intersection of an α-uniform set A of cardinality δN with any sufficiently long progression P is approximately equal to $\delta|P|$. The exact meaning of the words *sufficiently long* will become clear from the proof of Roth's theorem. Hence, if the set A is not α-uniform, then there is a progression P such that $|A \cap P| = (\delta + \theta)|P|$, where $|\theta| > 0$. More precise arguments enable one to prove that P can be chosen so that θ is positive and can be expressed explicitly in terms of the density δ.

Next we consider the new set $A' = A \cap P$ and apply our algorithm to it. We note that since $A' \subseteq A$, the set A' contains no arithmetic progressions of length three.

Moreover, the density of A' in P is not less than $\delta + \theta$, where $\theta > 0$. Hence, at every step of the algorithm the density of the sets obtained increases by a positive quantity. On the other hand, the density is always not greater than one. This means that our algorithm must terminate in finitely many steps. Consequently, at some step of the algorithm we obtain an arithmetic progression $\widetilde{P}$ and an α-uniform set $\widetilde{A}$ belonging to $A \cap \widetilde{P}$. As was said above, in this case the set $\widetilde{A}$ contains an arithmetic progression of length three. Thus, A also contains this arithmetic progression. We again arrive at a contradiction to the assumption that there are no arithmetic progressions of length three in A.

We have thus sketched the proof of Roth's theorem. Let us proceed to the proof itself.

Proposition 1. *Let $\delta > 0$, let M be a positive integer, let P be an arithmetic progression of length M, and let A be a subset of P, $|A| = \delta M$, without arithmetic progressions of length three. Assume that*

$$M \geqslant 2^{25}\pi^2\delta^{-4}. \tag{10}$$

Then there is a progression $P' \subseteq P$ such that

1) $|A \cap P'| \geqslant (\delta + 2^{-9}\delta^2)|P'|$,
2) $|P'| \geqslant 2^{-15}\delta^2\sqrt{M}$.

We show that Roth's theorem follows from Proposition 1.

As remarked above, the proof of Roth's theorem is an algorithm. Here is the first step of the algorithm. Suppose that a set $A \subseteq [N]$ contains no arithmetic progressions of length three. Let $P_0 = [N]$, $A_0 = A$, and $\delta_0 = \delta$. We have $N \geqslant 2^{25}\pi^2\delta^{-4}$. Applying Proposition 1, we obtain a progression $P_1 \subseteq [N]$ such that $|A_0 \cap P_1| \geqslant (\delta_0 + 2^{-9}\delta_0^2)|P_1|$ and $|P_1| \geqslant 2^{-15}\delta_0^2\sqrt{|P_0|}$. Then we set $A_1 = A_0 \cap P_1$.

Suppose that at the ith step of the algorithm, $i \geqslant 1$, we have constructed a progression P_i and a set $A_i \subseteq P_i$ without arithmetic progressions of length three and such that

$$|A_{i-1} \cap P_i| \geqslant (\delta_{i-1} + 2^{-9}\delta_{i-1}^2)|P_i|, \qquad |P_i| \geqslant 2^{-15}\delta_{i-1}^2\sqrt{|P_{i-1}|}. \tag{11}$$

We note that the density of the sets A_i in the progressions P_i increases by the quantity $2^{-9}\delta_{i-1}^2$ at every step of the algorithm. At the first step the density of A_0 in P_0 is equal to δ. If

$$N_i := |P_i| \geqslant 2^{25}\pi^2\delta^{-4} \geqslant 2^{25}\pi^2\delta_i^{-4}, \tag{12}$$

then the $(i+1)$st step of the algorithm can be carried out.

We estimate the maximal number of steps of the algorithm. Let $\varepsilon(t) = 2^{-9}t^2$. Also, let $k_1 = \lceil \delta_0/\varepsilon(\delta_0) \rceil$, $k_2 = \lceil \delta_{k_1}/\varepsilon(\delta_{k_1}) \rceil$, $\ldots$, $k_s = \lceil \delta_{k_{s-1}}/\varepsilon(\delta_{k_{s-1}}) \rceil$, $\ldots$. Using the first inequality in (11), we obtain $\delta_{k_1} \geqslant 2\delta$, $\delta_{k_2} \geqslant 2^2\delta$, $\ldots$, $\delta_{k_s} \geqslant 2^s\delta$, $\ldots$. This implies that in at most K steps, where

$$k_1 + k_2 + \cdots < 200\left(\frac{1}{\delta} + \frac{1}{2\delta} + \frac{1}{2^2\delta} + \cdots\right) = \frac{400}{\delta} = K,$$

the density δ_K becomes greater than 1. The contradiction obtained does not yet prove Roth's theorem, because we did not verify the condition (12) at every step

of our algorithm. We have $N \gg \exp\exp(\delta^{-1})$. Applying the second inequality in (11), we obtain

$$N_K \geqslant (2^{-15}\delta_K^2)^K N^{1/2^K} \geqslant (2^{-15}\delta^2)^K N^{1/2^K} = 2^{-6000/\delta}\delta^{800/\delta} N^{1/2^K} \geqslant 2^{25}\pi^2\delta^{-4}.$$

Since $N_i \geqslant N_K$ for any $i \leqslant K$, it follows that the inequality (12) holds at every step of the algorithm. This completes the proof of Theorem 13.

Proof of Proposition 1. Let $P = \{b, b+d, \dots, b+(|P|-1)d\}$. Without loss of generality we can assume that $b = d = 1$. For otherwise we can consider the progression $\widetilde{P} = \{n : n = (p-b)/d,\ p \in P\}$ and the set $\widetilde{A} = \{n : n = (a-b)/d, a \in A\}$. Then $\widetilde{A} \subseteq \widetilde{P}$, and $\widetilde{A}$ contains no arithmetic progressions of length three.

Thus, let $P = \{1, \dots, M\}$.

Case 1. *The set A is α-uniform.* Suppose that A is an α-uniform subset of P, where $\alpha = 2^{-5}\delta^2$. The number of arithmetic progressions in A is equal to

$$\sigma = \int \widehat{A}^2(x)\widehat{A}(-2x)\,dx. \tag{13}$$

We have

$$\begin{aligned}\sigma &= \delta^3 \int \widehat{P}^2(x)\widehat{P}(-2x)\,dx + \delta^2 \int \widehat{P}^2(x)(\widehat{A}-\delta\widehat{P})(-2x)\,dx \\ &\quad + \delta \int \widehat{P}(x)(\widehat{A}-\delta\widehat{P})(x)\widehat{A}(-2x)\,dx + \int (\widehat{A}-\delta\widehat{P})(x)\widehat{A}(x)\widehat{A}(-2x)\,dx \\ &= \sigma^* + \sigma_1 + \sigma_2 + \sigma_3\,. \end{aligned} \tag{14}$$

The modulus of each of the terms σ_1, σ_2, and σ_3 does not exceed $\alpha\delta M^2$. For example, let us estimate σ_3. Since the set A is α-uniform, it follows from the equality (7) and the Cauchy–Bunyakovskii inequality that

$$\begin{aligned}|\sigma_3| &\leqslant \|\widehat{A}-\delta\widehat{P}\|_\infty \int |\widehat{A}(x)|\,|\widehat{A}(-2x)|\,dx \\ &\leqslant \alpha M \left(\int |\widehat{A}(x)|^2\,dx\right)^{1/2} \left(\int |\widehat{A}(-2x)|^2\,dx\right)^{1/2} = \alpha\delta M^2.\end{aligned}$$

A similar estimate applies for σ_1 and σ_2. Hence,

$$\sigma = \delta^3 \int \widehat{P}^2(x)\widehat{P}(-2x)\,dx + 3\alpha\delta M^2\theta_1, \tag{15}$$

where $|\theta_1| \leqslant 1$. The quantity $\int \widehat{P}^2(x)\widehat{P}(-2x)\,dx$ is the number of arithmetic progressions in $[M]$. This number is equal to $(M-2)+(M-4)+\dots+(M-\lfloor M/2\rfloor) \geqslant M^2/8$. Since $\alpha = 2^{-5}\delta^2$, it follows that $\sigma \geqslant \delta^3 M^2/8 - 3\delta^3 M^2/32 = \delta^3 M^2/32$. The number of trivial arithmetic progressions in A is equal to δM. By assumption, we have $M \geqslant 2^{25}\pi^2\delta^{-4}$. Hence, $\delta^3 M^2/32 > \delta M$. The last inequality means that A contains a non-trivial arithmetic progression of length three, a contradiction to the assumption of Proposition 1.

Case 2. *The set A is not α-uniform.* We assume now that A is not an α-uniform subset of P with $\alpha = 2^{-5}\delta^2$. Let f be the balance function of the set A. Since A is not α-uniform, there is an $x_0 \in \mathbb{S}^1$ such that $|\widehat{f}(x_0)| \geqslant \alpha M$. Let $M_1 = \sqrt{M}$. By the Dirichlet theorem, there is a positive integer $q \leqslant M_1$ such that $\|qx_0\| \leqslant 1/M_1$, where $\|\cdot\|$ is the distance to the nearest integer. In other words, $x_0 = p/q + \theta/(qM_1)$, where $|\theta| \leqslant 1$. Let $P_j = \{n \in [M] \mid n \equiv j \pmod q\}$, $j = 1, \dots, q$. We have $|P_j| = \lfloor (M-j)/q \rfloor + 1$ and $P_j = \{j + kq,\ k = 0, 1, \dots, \lfloor (M-j)/q \rfloor\}$. Let $\widetilde{P}_j = \{0, 1, \dots, \lfloor (M-j)/q \rfloor\}$.

Let $t = \lceil (\pi 2^5)/(\alpha M_1) \cdot M/q \rceil > 1$. We have $t < M/q$. One can easily show that every progression $\widetilde{P}_j$ can be partitioned into t progressions $\widetilde{P}_j^l$ whose lengths can differ by at most 1. We have

$$\begin{aligned}
\alpha M \leqslant |\widehat{f}(x_0)| &= \left| \sum_n f(n) e^{-2\pi i n x_0} \right| \\
&= \left| \sum_{j=1}^{q} \sum_{l=1}^{t} \sum_{k \in \widetilde{P}_j^l} f(j+kq) e^{-2\pi i (j+kq)\left(\frac{p}{q} + \frac{\theta}{qM_1}\right)} \right| \\
&= \left| \sum_{j=1}^{q} e^{-2\pi i j \left(\frac{p}{q} + \frac{\theta}{qM_1}\right)} \sum_{l=1}^{t} \sum_{k \in \widetilde{P}_j^l} f(j+kq) e^{-2\pi i \frac{k\theta}{M_1}} \right| \\
&\leqslant \sum_{j=1}^{q} \sum_{l=1}^{t} \left| \sum_{k \in \widetilde{P}_j^l} f(j+kq) e^{-2\pi i \frac{k\theta}{M_1}} \right|. \qquad (16)
\end{aligned}$$

Let $\widetilde{P}_j^l = \{c, c+1, \dots, c+r-1\}$, $r = |\widetilde{P}_j^l|$. Then

$$\sum_{k \in \widetilde{P}_j^l} f(j+kq) e^{-2\pi i \frac{k\theta}{M_1}} = e^{-2\pi i \frac{c\theta}{M_1}} \left(\sum_{k \in \widetilde{P}_j^l} f(j+kq) + 2\pi\theta' \frac{r^2}{M_1} \right). \qquad (17)$$

Since the lengths of the progressions $\widetilde{P}_j^l$ can differ by at most 1, it follows that $r = |\widetilde{P}_j^l| \leqslant 4M/(tq)$ for any l. Using the last inequality together with (16) and (17), we see that

$$\sum_{j=1}^{q} \sum_{l=1}^{t} \left| \sum_{k \in \widetilde{P}_j^l} f(j+kq) \right| \geqslant \frac{\alpha M}{2}. \qquad (18)$$

We have

$$\sum_{j=1}^{q} \sum_{l=1}^{t} \sum_{k \in \widetilde{P}_j^l} f(j+kq) = 0. \qquad (19)$$

The inequality (18) and the equality (19) imply the existence of a progression $\widetilde{P}_{j_0}^{l_0}$ such that

$$\sum_{k \in \widetilde{P}_{j_0}^{l_0}} f(j+kq) \geqslant \frac{\alpha M}{4qt}. \qquad (20)$$

Let $P' = \{n : n = j + kq,\ k \in \widetilde{P}^{l_0}_{j_0}\}$. We have $|P'| = |\widetilde{P}^{l_0}_{j_0}| \geqslant M/(4qt)$. Since $t = \lceil \pi 2^5/(\alpha M_1)\cdot M/q\rceil$, it follows that $|P'| \geqslant 2^{-10}\alpha M_1 = 2^{-15}\delta^2\sqrt{M}$, and thus the condition 1) in Proposition 1 holds. Let us now prove that the condition 2) also holds. Using the inequality (20), we obtain

$$2^{-9}\delta^2|P'| = \frac{\alpha|P'|}{16} \leqslant \frac{\alpha M}{4qt} \leqslant \sum_{n\in P'} f(n) = |A\cap P'| - \delta|P'|. \qquad (21)$$

This completes the proof of Proposition 1.

Recently, Green [25] somewhat sharpened Roth's theorem. He proved that an arbitrary sufficiently dense subset of $[N]$ contains an arithmetic progression of length three whose difference can be represented in the form $x^2 + y^2$, where x and y are some positive integers.

Theorem 14 (Green). *Let N be a positive integer. Then there is an effective constant $c > 0$ such that every set $A \subseteq [N]$ with $|A| \gg N/(\log\log N)^c$ contains an arithmetic progression of length three whose difference can be represented in the form $x^2 + y^2$.*

§ 3. Lower bounds for $a_k(N)$

In the previous section we proved Roth's theorem on an upper bound for the quantity $a_3(N)$. We shall now present several results on *lower* bounds for $a_k(N)$, $k \geqslant 3$.

In 1946 Behrend [45] developed an approach in the papers [46] and [47] of R. Salem and D. C. Spencer and obtained the following lower bound for $a_3(N)$ (see also [48]).

Theorem 15 (Behrend). *Let $\varepsilon > 0$ be arbitrary. Then there is a number $N_\varepsilon \in \mathbb{N}$ such that*

$$a_3(N) \geqslant \exp\bigl(-(1+\varepsilon)C\sqrt{\ln N}\,\bigr)$$

for any positive integer $N \geqslant N_\varepsilon$, where C is a positive absolute constant.

Proof. Let N_ε be a positive integer such that $\ln(4\ln N)/\sqrt{\ln N} < \varepsilon$ for any $N \geqslant N_\varepsilon$, $N \in \mathbb{N}$. Also, let m and n be positive integer parameters and let $\Lambda = \{0, 1, \dots, m-1\}^n$. We consider the n-dimensional sphere $S_t = \{\mathbf{x} \in \Lambda : x_1^2 + \dots + x_n^2 = t\}$, where $0 \leqslant t \leqslant n(m-1)^2$. It is clear that every sphere S_t contains no arithmetic progressions of length three, in the sense that the equality $\mathbf{x} + \mathbf{y} = 2\mathbf{z}$ is possible for $\mathbf{x}, \mathbf{y}, \mathbf{z} \in S_t$ only if $\mathbf{x} = \mathbf{y} = \mathbf{z}$.

We note that the total number of spheres does not exceed nm^2. Since every point of the set Λ belongs to some sphere, it follows from the Dirichlet principle that there is a $t_0 \in \mathbb{N}$ for which the cardinality of S_{t_0} is not less than $m^n/(nm^2) = m^{n-2}/n$.

Thus, we have found a rather dense set S_{t_0} in Λ that contains no arithmetic progressions of length three. Let us construct from S_{t_0} a new set $A \subseteq \mathbb{N}$ which also does not contain arithmetic progressions. To this end, we consider the map $\varphi\colon \Lambda \to \mathbb{Z}$ given by the formula $\varphi(\mathbf{x}) = \sum_{i=1}^n x_i(2m)^{i-1}$. Let $A = \varphi(S_{t_0})$. Since the equality $\varphi(\mathbf{x}) + \varphi(\mathbf{y}) = 2\varphi(\mathbf{z})$ holds if and only if $\mathbf{x} + \mathbf{y} = 2\mathbf{z}$, it follows that the

set A contains no arithmetic progressions of length three. We note that the set A is contained here in the segment $\{1, 2, \dots, (2m)^n\}$ of positive integers.

We now choose the parameters m and n. Namely, let $n = \lceil \sqrt{2 \log N} \rceil$ and let m satisfy the conditions $(2(m-1))^n \leqslant N < (2m)^n$. We have

$$|A| = |S_{t_0}| \geqslant \frac{m^{n-2}}{n} \geqslant N \frac{N^{-2/n}}{2^n n} \geqslant N \exp\left(-\left(2\sqrt{2 \ln 2} + \frac{\ln(4 \ln N)}{\sqrt{\ln N}} \right) \sqrt{\ln N} \right)$$
$$\geqslant N \exp\bigl(-\bigl(2\sqrt{2 \ln 2} + \varepsilon \bigr) \sqrt{\ln N} \bigr). \tag{22}$$

Since $a_3(N) \geqslant |A|/N$, this proves Theorem 15.

Although the result obtained by Behrend in 1946 is simple, it remains the best result so far.

Rankin [49] generalized Behrend's theorem to the case of arbitrary $k \geqslant 3$.

Theorem 16 (Rankin). *Let $\varepsilon > 0$ be arbitrary and let $k \geqslant 3$ be an integer. Then*

$$a_k(N) \geqslant \exp(-(1+\varepsilon) C_k (\ln N)^{1/(k-1)})$$

for all sufficiently large N, where C_k is a positive absolute constant depending only on k.

Developing the main idea of Theorem 15, Rankin constructs in his paper a sequence of spheres S_t. In this construction he uses the asymptotic formula in [50] for the number of solutions of the equation $x_1^2 + \dots + x_n^2 = t$, $0 \leqslant x_i \leqslant m-1$, $i = 1, \dots, n$.

The first example of an infinite sequence of positive integers without arithmetic progressions of length three was proposed by Erdős and Turán in [51] (see also [52]). This sequence consists of numbers whose ternary expansion does not contain the digit 2:

$$0, 1, 3, 4, 9, 10, 12, 13, 27, \dots . \tag{23}$$

Unfortunately, this sequence has a very small density. The segment of integers from 1 to N contains roughly $N^{\log 2/\log 3}$ elements of this sequence. Rankin [49] proposed a way of constructing an infinite subset of positive integers whose density is equal to that used in Theorem 16 (see also [52]). For simplicity, we confine ourselves to the case $k = 3$.

Proposition 2. *Let $\varepsilon > 0$ be arbitrary. Then there is an infinite set $A^* \subseteq \mathbb{N}$ without an arithmetic progression of length three and such that*

$$\frac{|A^* \cap [M]|}{M} \geqslant \exp\bigl(-(1+\varepsilon) C \sqrt{\ln M} \bigr) \tag{24}$$

for all sufficiently large M, where C is the constant in Theorem 15.

Proof. Let N_ε be the number in Behrend's theorem, Theorem 15. We also assume that s is a positive integer such that $3^{s-2} < N \leqslant 3^{s-1}$. Consider the disjoint half-open intervals $[2 \cdot 3^{t-1}, 3^t)$, $t \geqslant s$. By Behrend's theorem, for any $t \geqslant s$ there is a set A_t belonging to $[2 \cdot 3^{t-1}, 3^t)$, containing no arithmetic progressions of length three, and having cardinality

$$|A_t| \geqslant 3^{t-1} \exp(-(1+\varepsilon) C (\ln 3^{t-1})^{1/2}). \tag{25}$$

Let $A^* = \bigsqcup_{t=s}^{\infty} A_t$. Then A^* is an infinite set. One can easily see that A^* contains no arithmetic progressions of length three. Let $M \in \mathbb{N}$ be sufficiently large, namely, $M \geqslant 3^s$. We set $p = [\ln M / \ln 3]$. Then

$$\frac{|A^* \cap [M]|}{M} \geqslant \frac{|A_p|}{M} \geqslant \frac{1}{9}\exp(-(1+\varepsilon)C(\ln M)^{1/2}) \geqslant \exp(-(1+2\varepsilon)C(\ln M)^{1/2}). \tag{26}$$

This completes the proof of Proposition 2.

§ 4. Szemerédi's theorem

In this section we discuss some ideas used in the proof of Szemerédi's theorem. In the present author's opinion, the core of the proof is what is called *the regularity lemma*, which asserts (roughly speaking) that every graph with n vertices and cn^2 edges ($0 < c \leqslant 1$ is an absolute constant) can be partitioned into a small number of subgraphs having 'random' properties. We cannot present here a proof of the regularity lemma nor, all the more so, a proof of Theorem 6. Rather, we confine ourselves to formulating the lemma and indicating how it is used to prove the estimate $a_3(n) = o(1)$. Thus, we consider only the simplest case $k = 3$ of Szemerédi's theorem.

Let us proceed to the formulation of the regularity lemma.

Let $G = (V, E)$ be a finite non-oriented graph without loops and multiple edges and let $A, B \subseteq V$ be two non-empty disjoint subsets of V. Let $e(A, B)$ be the number of edges (a, b) in G such that $a \in A$ and $b \in B$. By the *edge density* of the pair (A, B) we mean the ratio

$$d(A, B) := \frac{e(A, B)}{|A|\,|B|}\,.$$

Definition 3. A pair (A, B) is said to be ε-*uniform* if

$$|d(A', B') - d(A, B)| < \varepsilon$$

for any $A' \subseteq A$ and $B' \subseteq B$ such that $|A'| > \varepsilon|A|$ and $|B'| > \varepsilon|B|$.

Here is an example of ε-uniform pairs. Let A and B be two disjoint sets and let $0 < p \leqslant 1$. By a *bipartite random graph* we mean any graph $G_p = (V, E)$, where $V = A \sqcup B$ and any edge (a, b), $a \in A$, $b \in B$, belongs to E with probability p. Thus, the graph G_p contains no edges going from A to A or from B to B. It is clear that for G_p we have $d(A, B) = p$ almost surely and the pair (A, B) is ε-uniform almost surely for any fixed $\varepsilon > 0$.

Definition 4. A partition of the set V of vertices of a graph G into sets C_0, C_1, …, C_k is said to be ε-*uniform* if

1) $|C_0| < \varepsilon|V|$,
2) $|C_1| = |C_2| = \cdots = |C_k|$,
3) all but possibly $\varepsilon\binom{k}{2}$ of the pairs of the form (C_i, C_j), $1 \leqslant i < j \leqslant k$, are ε-uniform.

Lemma 1 (Szemerédi's regularity lemma). *Let $0 < \varepsilon \leqslant 1$ and let l be a positive integer. There are two positive integers $n_0(\varepsilon, l)$ and $k_0(\varepsilon, l)$ such that for any graph G with at least $n_0(\varepsilon, l)$ vertices there is an ε-uniform partition of the vertices of G into k classes, where $l < k < k_0(\varepsilon, l)$.*

The classes C_i are referred to as *groups* or *clusters*. The regularity lemma asserts that, roughly speaking, the set of vertices of every sufficiently large graph can be partitioned into a not very large number of clusters C_i, $i = 1, \dots, k$, and an 'exceptional' set C_0 in such a way that 'almost all' pairs (C_i, C_j) behave like bipartite random graphs. The number l in the regularity lemma is needed for the cardinality of the clusters $|C_i|$, $i = 1, \dots, k$, to be sufficiently small. This is sometimes necessary in order to assert, for instance, that the number of edges between clusters is significantly greater than the number of edges beginning and ending in the same cluster C_i. It should be noted that Szemerédi's estimates for the numbers $n_0(\varepsilon, l)$ and $k_0(\varepsilon, l)$ are extremely weak (for this reason, see the paper [53]).

For the proof of the regularity lemma, see [14] and [54] (see also the nice survey [55]).

We make another remark. Suppose that $\varepsilon > 0$ and the number of edges in the graph $G = (V, E)$ is equal to $\varepsilon' |V|^2$, where ε' is a sufficiently small number depending on ε (for example, $\varepsilon' = \varepsilon^3/100$). In this case the assertion of the regularity lemma becomes trivial, because every partition of the set V into k clusters such that the conditions 1) and 2) hold must have the property 3) as well. Thus, the regularity lemma can be applied only for sufficiently dense graphs, for example, for graphs such that $|E| > c|V|^2$, where $c > 0$ is an absolute constant. Nevertheless, there are papers in which the regularity lemma is extended to the case of graphs with few edges (see, for instance, [56]).

We show how to obtain the estimate $a_3(n) = o(1)$ (see [12]) by using the regularity lemma. In the proof we follow the survey [57].

Theorem 17 (Ruzsa–Szemerédi). *$a_3(n) = o(1)$ as $n \to \infty$.*

Proof. Let $0 < \delta \leqslant 1$ be a fixed number, suppose that a set $A \subseteq [n]$ with $|A| = \delta n$ contains no arithmetic progressions of length three, and let X, Y, and Z be three disjoint copies of the segment $[1, 3n]$. Consider the set S of all the triples $(x, y, z) \in X \times Y \times Z$ such that

$$y - x = z - y = \frac{z - x}{2} \in A. \tag{27}$$

We construct a graph $G = (V, E)$ from the set A. Let $V = X \sqcup Y \sqcup Z$, $|V| = 9n$, and let a pair of vertices in V be joined by an edge if and only if there is a triple in S containing this pair. It is clear that $|E| \geqslant 3|A|n$. A triple $(x, y, z) \in X \times Y \times Z$ in G is referred to as a *triangle* if all the vertices x, y, z are joined by edges. If the triple (x, y, z) belongs to S, then the triangle is said to be *simple*.

It is easy to see that if the graph G contains a non-simple triangle, then the set A contains an arithmetic progression of length three. Indeed, let (x, y, z) be a non-simple triangle. Suppose, for example, that in this triangle we have $y - x \neq z - y$. Let $a := y - x$ and $b := z - y$. Then $a \in A$, $b \in B$, and $(a + b)/2 = (z - x)/2 \in A$. Hence, the set A contains an arithmetic progression of length three.

Thus, to prove the theorem, it suffices to show that G contains a non-simple triangle. Let $m = 9n = |V|$ and $\beta = |E|/\binom{m}{2}$. Since $|E| \geqslant 3|A|n = 3\delta n^2$,

it follows that β is a positive number independent of n. Let $\varepsilon = \beta/15$ and let $l = [\varepsilon^{-1}] + 1$. Applying the regularity lemma with the parameters ε and l to the graph G, we obtain a partition of the set V into clusters $C_1, \dots, C_k$ and C_0 satisfying the inequalities 1)–3). We have $|C_0| < \varepsilon |V|$. Therefore, the number of edges in G with an end at C_0 is at most εm^2. The properties 1) and 2) imply the inequality $m/(2k) \leqslant |C_i| \leqslant m/k$, $i = 1, \dots, k$. Hence, the number of edges beginning and ending in the same cluster is at most $k\binom{m/k}{2}$. Moreover, by the property 3) there are at most $\varepsilon\binom{k}{2}$ pairs (C_i, C_j), $1 \leqslant i < j \leqslant k$, that are not ε-uniform. Thus, the number of edges which are not contained in the ε-uniform pairs (C_i, C_j) with $d(C_i, C_j) \geqslant \beta/6$ is at most

$$\varepsilon m^2 + k\binom{m/k}{2} + \varepsilon\binom{k}{2}\left(\frac{m}{k}\right)^2 + \frac{\beta}{6}\binom{k}{2}\left(\frac{m}{k}\right)^2 < \frac{\beta}{3}\binom{m}{2}. \tag{28}$$

Deleting these edges, we obtain a graph G' and three *distinct* clusters C_p, C_q, and C_r contained in G', and every pair of these clusters is ε-uniform with edge density at least $\beta/6$.

We assert that the cluster C_r contains an element x that belongs to at least $(\beta/10)^3 |C_p|\,|C_q|$ triangles.

Since $d(C_p, C_r), d(C_q, C_r) \geqslant \beta/6$, there are at least $(1 - 2\varepsilon)|C_r|$ vertices x in C_r that are joined to at least $(\beta/6 - \varepsilon)|C_i|$ vertices of the cluster C_i, where $i = p$ or $i = q$. Let N_x^i be the set of vertices in C_i joined to x, $i = p, q$. We have $\beta/6 - \varepsilon = \beta/10 > \varepsilon$. It follows from the definition of the ε-uniform property that there are at least $(\beta/10)^3 |C_p|\,|C_q|$ edges joining the vertices in N_x^p and N_x^q. It is clear that corresponding to any such edge there is a triangle containing x.

Completing the proof of the theorem, we note that there are at most three simple triangles having two common vertices. Hence, there are at most $3|C_p| = 3|C_q|$ simple triangles containing x. Moreover, the inequality $|C_i| \geqslant m/(2k)$, $i = 1, \dots, k$, implies the estimate

$$\left(\frac{\beta}{10}\right)^3 |C_p|\,|C_q| > 3|C_p| \tag{29}$$

for any sufficiently large m and a fixed k. Therefore, G contains a non-simple triangle. Thus, the set A contains an arithmetic progression of length three, which completes the proof of Theorem 17.

Recently, analogues of the regularity lemma were obtained in [58]–[60] for hypergraphs, and it was shown how these analogues imply the estimate $a_k(n) = o(1)$ for any $k \geqslant 3$.

We formulate a sharpening of Szemerédi's theorem due to P. Varnavides [61].

Theorem 18 (Varnavides). *Let $k \geqslant 3$ be an integer and let $\delta > 0$. Assume that N is a sufficiently large integer and $A \subseteq [N]$ is a set such that*

$$|A| \geqslant \delta N. \tag{30}$$

Then

$$\frac{1}{N^2} \sum_{x, r \in [N]} A(x)A(x + r) \cdots A(x + (k - 1)r) \geqslant c(k, \delta) \tag{31}$$

for some constant $c(k, \delta) > 0$ depending only on k and δ.

Remark 1. Let $\rho \in (0,1]$. We denote by $N_k(\rho)$ the minimal positive integer such that an arbitrary set $T \subseteq [m]$ with $|T| \geqslant \rho m$ contains an arithmetic progression of length k for any positive integer $m \geqslant N_k(\rho)$. As we shall see in the proof of Theorem 18, for the constant $c(k,\delta)$ in (31) one can take $\delta^2/(16N_k(\delta/2)^3)$.

Proof of Theorem 18. Let $N > 2N_k(\rho/2)$. Let $K = N_k(\rho/2)$ and denote by P_d, $d \geqslant 1$, the family of arithmetic progressions of length K with difference d that are contained in $[N]$. It is clear that $|P_d| \leqslant N$ for any $d \geqslant 1$. Let $P = \bigsqcup_{d\geqslant 1} P_d$.

A progression $p \in P$ is said to be *good* if $|A \cap p| \geqslant \delta/2 \cdot K$. Let G be the set of all good progressions. We assert that $|G| \geqslant \delta^2 N^2/(32K)$. Consider the case in which

$$d < \frac{\delta N}{8K}. \tag{32}$$

We note that every element $x \in [Kd, N - Kd]$ belongs to exactly K progressions in P_d. Applying the inequality (32), we see that

$$\begin{aligned}\sigma &= \sum_{p\in P_d} |A\cap p| = \sum_{x\in[N]} A(x) \sum_{p\in P_d} p(x) \geqslant \sum_{x\in[Kd,N-Kd]} A(x) \sum_{p\in P_d} p(x)\\ &= K|A\cap[Kd,N-Kd]| \geqslant K(|A| - 2Kd)\\ &= K(\delta N - 2Kd) \geqslant 3\delta KN/4.\end{aligned} \tag{33}$$

On the other hand,

$$\begin{aligned}3\delta KN/4 \leqslant \sigma &\leqslant \sum_{p\in(P_d\cap G)} |A\cap p| + \sum_{p\in(P_d\setminus G)} |A\cap p|\\ &\leqslant |P_d\cap G|K + \delta KN/2.\end{aligned} \tag{34}$$

Therefore, for any $1 \leqslant d < \delta N/(8K)$ we have the inequality $|P_d \cap G| \geqslant \delta N/4$. Hence, $|G| \geqslant \delta^2 N^2/(32K)$.

By definition, for any progression in G one can find $x, r \in [N]$ such that $A(x) \times A(x+r)\cdots A(x+(k-1)r) > 0$. Unfortunately, the numbers $x, r \in [N]$ can be the same for distinct progressions in G. Let us estimate the number of progressions $p \in P$ containing a fixed progression $x_0, x_0 + r_0, \ldots, x_0 + (k-1)r_0$. It is clear that there are at most $K - 2$ progressions of this kind in P_{r_0}. Moreover, if a progression $p \in P_d$, $d \neq r_0$, contains the progression $x_0, x_0 + r_0, \ldots, x_0 + (k-1)r_0$, then d divides r_0. Let $d = r_0/t$, where $t > 1$. Since $p \in P_d$ contains the progression $x_0, x_0 + r_0, \ldots, x_0 + (k-1)r_0$, it follows that $Kd > 2r_0 = 2td$, and hence $t < K/2$. Therefore, the number of progressions $p \in P$ containing the progression $x_0, x_0 + r_0, \ldots, x_0 + (k-1)r_0$ is at most $K - 2 + (K-2)K/2 \leqslant K^2/2$.

We have $2|G|/(K^2N^2) \geqslant \delta^2/(16K^3)$. This implies that the constant $c(k,\delta)$ in (31) can be taken to be $\delta^2/(16K^3) = \delta^2/(16N_k(\delta/2)^3)$, and Theorem 18 is proved.

Croot [62] considered the problem of the number of arithmetic progressions modulo N and somewhat sharpened Theorem 18.

Theorem 19 (Croot). *Let $k \geqslant 3$ be an integer and let $\delta > 0$. Assume that N is a sufficiently large integer and $A \subseteq \mathbb{Z}_N$ is a set such that*

$$|A| \geqslant \delta N. \tag{35}$$

Then

$$\frac{1}{N^2} \sum_{x,r \in \mathbb{Z}_N} A(x)A(x+r)\cdots A(x+(k-1)r) \geqslant c(k,\delta) \tag{36}$$

for some constant $c(k,\delta) > 0$ depending only on k and δ. Moreover, the constant $c(k,\delta)$ can be taken to be $\delta/(16N_k(\delta/2)^2)$.

Let $k \geqslant 3$ be an integer and let $\delta > 0$. Also, let N be a sufficiently large positive integer and let $A \subseteq \mathbb{Z}_N$ be a set such that $|A| \geqslant \delta N$. The following question arises in connection with Theorems 18 and 19: how many arithmetic progressions of length k belong to the set A for sufficiently large N? We denote by $\mu_k(A)$ the number of arithmetic progressions of length k in A, divided by N^2. Using the bound in Theorem 5 and the estimates in Theorems 18 and 19, we see that $\mu_3(A) \geqslant \exp(C\delta^{-2}\log(1/\delta))$, where $C > 0$ is an absolute constant. For an arbitrary $k \geqslant 4$ it follows from Theorems 18 and 19 and from Theorem 9 that $\mu_k(A) \gg \exp(-\exp(\delta^{-c_k}))$, where $c_k > 0$ is the absolute constant in Theorem 9. Using results of Behrend [45] and Rankin [49], Croot [62] obtained *lower* bounds for the quantity $\mu_k(A)$.

Theorem 20 (Croot). *Let $k \geqslant 3$ be an integer, let $\delta \in (0,1)$, and let N be a sufficiently large positive integer. Then there is a set $A \subseteq \mathbb{Z}_N$ such that $|A| \geqslant \delta N$ and A contains at most $N^2 \exp\Big(-\Big(\dfrac{1}{2C_k}\log\dfrac{1}{4\delta}\Big)^{k-1}\Big)$ arithmetic progressions of length k, where C_k is the absolute constant in Theorem 16.*

In other words, Theorem 20 gives a lower bound for $\mu_k(A)$:

$$\mu_k(A) \geqslant \exp\Big(-\Big(\frac{1}{2C_k}\log\frac{1}{4\delta}\Big)^{k-1}\Big).$$

Proof. Let $L_k(x) = \exp(2C_k(\log x)^{1/(k-1)})$, where C_k is the absolute constant in Theorem 16, and take an x such that $4L_k(x) < 1/\delta \leqslant 4L_k(x+1)$. Let $N > 4x$. We apply Theorem 16 with $\varepsilon = 1$. By this theorem, there is a set $S \subseteq [x]$, $|S| \leqslant xL_k(x)^{-1}$, containing no arithmetic progressions of length k. Let

$$A = \{s + 2mx : s \in S, : 0 \leqslant m \leqslant M = [N/(4x)]\}.$$

Then $A \subseteq [N/2] \subseteq \mathbb{Z}_N$. We note that

$$\frac{|A|}{N} = \frac{|S|(M+1)}{N} > \frac{|S|}{4x} > \frac{1}{4L_k(x)} > \delta.$$

Thus, the density of A in $\mathbb{Z}_N$ is not less than δ. Suppose that the numbers $a_1, \ldots, a_k \in A$ form an arithmetic progression of length k. Since $A \subseteq [N/2]$, these numbers form an arithmetic progression of length k in $\mathbb{Z}_N$ if and only if they form an arithmetic progression of length k in $[N]$. It follows from the properties

of the set A that $a_i = s + 2m_i x$, $i = 1, \dots, k$, where $s \in S$ and the numbers $m_1, \dots, m_k$ form an arithmetic progression in $[M]$. Therefore, A contains at most $|S|M^2$ arithmetic progressions of length k. We have

$$|S|M^2 \leqslant \frac{xN^2}{4x^2} \leqslant \frac{N^2}{x+1} \leqslant \frac{N^2}{\exp\left(\left(\frac{1}{2C_k}\log\frac{1}{4\delta}\right)^{k-1}\right)}.$$

This completes the proof of Theorem 20.

§ 5. Gowers' bounds for $a_k(N)$

In this section we discuss upper bounds for $a_k(N)$ obtained by Gowers [22] (see Theorem 9). The proof of Theorem 9 is very complicated, and therefore we confine ourselves to the discussion of the main ideas used in Gowers' method.

In his remarkable paper Gowers gave a definition of α-uniform functions of degree d. We discuss the combinatorial meaning of α-uniform functions of degree d somewhat later, but for now we give the rigorous definitions.

Let $d \geqslant 0$ be a positive integer and let $\{0,1\}^d = \{\omega = (\omega_1, \dots, \omega_d) : \omega_j \in \{0,1\}, j = 1, \dots, d\}$ be the usual d-dimensional cube. For any $\omega \in \{0,1\}^d$ we let $|\omega| = \omega_1 + \dots + \omega_d$. If $h = (h_1, \dots, h_d) \in \mathbb{Z}_N^d$, then $\omega \cdot h := \omega_1 h_1 + \dots + \omega_d h_d$. We denote by $\mathscr{C}$ the operator of complex conjugation. If n is a positive integer, then $\mathscr{C}^n$ denotes the n-fold application of the operator of complex conjugation.

Definition 5. Let 2^d complex functions $(f_\omega)_{\omega\in\{0,1\}^d}$ defined on $\mathbb{Z}_N$ be given. By the *Gowers inner product* of the functions $(f_\omega)_{\omega\in\{0,1\}^d}$ we mean the quantity

$$\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d} := \frac{1}{N^{d+1}} \sum_{x\in\mathbb{Z}_N,\ h\in\mathbb{Z}_N^d} \prod_{\omega\in\{0,1\}^d} \mathscr{C}^{|\omega|} f_\omega(x + \omega\cdot h). \tag{37}$$

We derive some properties of the Gowers inner product.

Let $d \geqslant 1$. Assume first that the functions $(f_\omega)_{\omega\in\{0,1\}^d}$ do not depend on the last digit ω_d (in other words, $f_\omega = f_{\omega_1,\dots,\omega_{d-1}}$). Then the formula (37) can be rewritten as follows:

$$\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d} = \frac{1}{N^{d+1}} \sum_{x\in\mathbb{Z}_N,\ h'\in\mathbb{Z}_N^{d-1},\ h_d\in Z_N} \prod_{\omega'\in\{0,1\}^{d-1}} \mathscr{C}^{|\omega'|}(f_{\omega'}(x+\omega'\cdot h')\overline{f_{\omega'}(x+h_d+\omega'\cdot h')}),$$

where $\omega' = (\omega_1, \dots, \omega_{d-1})$ and $h' = (h_1, \dots, h_{d-1})$. Therefore,

$$\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d} = \frac{1}{N^{d+1}} \sum_{h'\in\mathbb{Z}_N^{d-1}} \left| \sum_{y\in\mathbb{Z}_N} \prod_{\omega'\in\{0,1\}^{d-1}} \mathscr{C}^{|\omega'|} f_{\omega'}(y+\omega'\cdot h') \right|^2. \tag{38}$$

Hence, the Gowers inner product has the following non-negativity property for any $d \geqslant 1$ and any function $f\colon \mathbb{Z}_N \to \mathbb{C}$:

$$\langle (f)_{\omega\in\{0,1\}^d}\rangle_{U^d} \geqslant 0. \tag{39}$$

The inequality (39) enables one to define the *Gowers uniform norm* (or simply the *Gowers norm*) of any function $f\colon \mathbb{Z}_N \to \mathbb{C}$ by the formula

$$\|f\|_{U^d} := \langle (f)_{\omega\in\{0,1\}^d}\rangle_{U^d}^{1/2^d} = \left(\frac{1}{N^{d+1}} \sum_{x\in\mathbb{Z}_N,\ h\in\mathbb{Z}_N^d} \prod_{\omega\in\{0,1\}^d} \mathscr{C}^{|\omega|} f_\omega(x+\omega\cdot h)\right)^{1/2^d}. \tag{40}$$

We shall see below that the formula (40) defines a norm only for $d \geqslant 2$. The Gowers norm for $d = 1$ is in fact a seminorm.

If the functions $(f_\omega)_{\omega\in\{0,1\}^d}$ really depend on the last digit ω_d, then the sum (37) must be rewritten as follows:

$$\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d} = \frac{1}{N^{d+1}} \sum_{h'\in\mathbb{Z}_N^{d-1}} \left(\sum_{y\in\mathbb{Z}_N} \prod_{\omega'\in\{0,1\}^{d-1}} \mathscr{C}^{|\omega'|} f_{\omega',0}(y+\omega'\cdot h')\right) \times \left(\overline{\sum_{y\in\mathbb{Z}_N} \prod_{\omega'\in\{0,1\}^{d-1}} \mathscr{C}^{|\omega'|} f_{\omega',1}(y+\omega'\cdot h')}\right).$$

Using the Cauchy–Bunyakovskii inequality and the formula (38), we obtain

$$|\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d}| \leqslant \langle (f_{\omega',0})_{\omega\in\{0,1\}^d}\rangle_{U^d}^{1/2} \cdot \langle (f_{\omega',1})_{\omega\in\{0,1\}^d}\rangle_{U^d}^{1/2}.$$

Using the Cauchy–Bunyakovskii inequality and the formula (38) now for every digit $\omega \in \{0,1\}^d$, we see that

$$|\langle (f_\omega)_{\omega\in\{0,1\}^d}\rangle_{U^d}| \leqslant \prod_{\omega\in\{0,1\}^d} \|f_\omega\|_{U^d}. \tag{41}$$

The inequality (41) is called the *Cauchy–Bunyakovskii–Gowers inequality*. By (41), the fact that the inner product (37) is multilinear, and Newton's binomial formula, we easily obtain

$$|\langle (f+g)_{\omega\in\{0,1\}^d}\rangle_{U^d}| \leqslant (\|f\|_{U^d} + \|g\|_{U^d})^{2^d}, \tag{42}$$

which implies the triangle inequality for the Gowers norm:

$$\|f+g\|_{U^d} \leqslant \|f\|_{U^d} + \|g\|_{U^d}, \qquad d \geqslant 1. \tag{43}$$

As was mentioned above, the Gowers norm is not a norm for $d = 1$. Indeed, $\|f\|_{U^1} = \frac{1}{N}\bigl|\sum_{x\in\mathbb{Z}_N} f(x)\bigr|$, and $\|f\|_{U^1}$ vanishes for any function such that $\sum_{x\in\mathbb{Z}_N} f(x) = 0$. On the contrary, the Gowers norm is a norm for $d \geqslant 2$. Let us prove this.

Let $\widehat{f}(r) = \sum_{x\in\mathbb{Z}_N} f(x)e^{-2\pi ixr/N}$ be the rth Fourier coefficient of the function f. Then the inversion formula $f(x) = 1/N \cdot \sum_{r\in\mathbb{Z}_N} \widehat{f}(r)e^{2\pi ixr/N}$ holds, and it implies the equality

$$\|f\|_{U^2} = \left(\sum_{r\in\mathbb{Z}_N} |\widehat{f}(r)|^4\right)^{1/4}. \tag{44}$$

This means that $\|f\|_{U^2} = 0$ if and only if $\widehat{f} \equiv 0$ or, which is the same, $f \equiv 0$. Thus, we have proved that $\|f\|_{U^2}$ is a norm.

We now assert that $\|f\|_{U^d}$ is a norm for any $d \geqslant 2$. Let ν_{const} stand for the function identically equal to one. We have $\|\nu_{\mathrm{const}}\|_{U^d} = 1$. Let $f\colon \mathbb{Z}_N \to \mathbb{C}$ be a function. Consider the family of functions $(f_\omega)_{\omega\in\{0,1\}^d}$, where $f_\omega := \nu_{\mathrm{const}}$ if $\omega_d = 1$ and $f_\omega := f$ if $\omega_d = 0$. Applying the inequality (41) to the family $(f_\omega)_{\omega\in\{0,1\}^d}$, we obtain

$$\|f\|_{U^{d-1}} \leqslant \|f\|_{U^d} \tag{45}$$

for any $d \geqslant 2$. The inequality (45) is called the *monotonicity inequality* for the Gowers norm. It follows from (45) that if the expression $\|f\|_{U^d}$ vanishes for $d \geqslant 2$, then the norm $\|f\|_{U^2}$ of the function f vanishes, and hence $f \equiv 0$.

Other properties of the Gowers norm can be found in the recent paper [32] (see also [63]).

Using the notion of the norm $\|\cdot\|_{U^d}$, Gowers gave a definition of the α-uniform functions of degree d.

Definition 6. Let $d \geqslant 2$ be an integer and let $\alpha \in [0,1]$. A function $f\colon \mathbb{Z}_N \to \mathbb{C}$ is said to be *α-uniform of degree d* if

$$\|f\|_{U^d} \leqslant \alpha.$$

In § 2 we gave another definition of α-uniform functions (see Definition 2). One can easily show that the usual notion of α-uniformity in § 2 coincides with Definition 6 for $d = 2$ (see [23] and [22]). We thus see that Gowers' approach significantly generalizes the classical Definition 2.

Let $f = A - \delta$ be the balance function of a set A. The set A is said to be α-uniform of degree d if the balance function of A is α-uniform of degree d.

We discuss the combinatorial meaning of the notion of α-uniform set of degree d.

Let $d \geqslant 0$ and let $a_0, a_1, \ldots, a_d \in \mathbb{Z}_N$ be some residues. Then the corresponding *d-dimensional cube* is defined to be the set of 2^d points in $\mathbb{Z}_N$ of the form $a_0 + \varepsilon_1 a_1 + \cdots + \varepsilon_d a_d$, where $\varepsilon_i \in \{0,1\}$. Let $A \subseteq \mathbb{Z}_N$ be a set; A contains a d-dimensional cube if all the points of this cube belong to A. Using the Cauchy–Bunyakovskii inequality, one can easily show that every set $A \subseteq \mathbb{Z}_N$ of cardinality δN always contains at least $\delta^{2^d} N^{d+1}$ d-dimensional cubes, and the equality is attained at the 'random subsets' of $\mathbb{Z}_N$ that have density δ. On the other hand, the following result holds.

Theorem 21 (combinatorial meaning of α-uniform sets of degree d). *Let $d \geqslant 2$ and let $A \subseteq \mathbb{Z}_N$, $|A| = \delta N$, be an α-uniform set of degree d. Then A contains at most $(\delta + \alpha)^{2^d} N^{d+1}$ cubes.*

Proof. We have $A = \delta + f$. Applying the triangle inequality for the Gowers norm (43), we obtain $\|A\|_{U^d} \leqslant \|\delta\|_{U^d} + \|f\|_{U^d}$. It is clear that $\|\delta\|_{U^d} = \delta$ and that $N^{d+1}\|A\|_{U^d}^{2^d}$ is the number of d-dimensional cubes in A. Since the set A is α-uniform of degree d, it follows from the definition that $\|f\|_{U^d} \leqslant \alpha$. Hence, $\|A\|_{U^d} \leqslant \delta + \alpha$ and we obtain the desired result.

Theorem 21 shows that the number of d-dimensional cubes in α-uniform sets of degree d for small values of the parameter α is approximately equal to the corresponding number of cubes contained in random sets. In this sense the α-uniform sets are close to random sets. Moreover, in many cases the α-uniform sets behave

like random sets. For example, Gowers showed in the first part of his arguments that the α-uniform sets and the random sets contain approximately the same number of arithmetic progressions.

Theorem 22. *Let $k \geqslant 3$ and let $A \subseteq \mathbb{Z}_N$, $|A| = \delta N$, be an α-uniform set of degree $k-1$. Then*

$$\sum_{r \in \mathbb{Z}_N} |(A+r) \cap (A+2r) \cap \dots \cap (A+kr) - \delta^k N^2| \leqslant 2^k \alpha^{1/2^{k-1}} N^k.$$

Gowers derives the following corollary from Theorem 22.

Corollary 2. *Let $k \geqslant 3$ and let $A \subseteq \mathbb{Z}_N$, $|A| = \delta N$, be an α-uniform set of degree $k-1$. Assume that $\alpha \leqslant (\delta/2)^{2k}$ and $N \geqslant 32k^2\delta^{-k}$. Then A contains an arithmetic progression of length k.*

Corollary 2 completes the first part of Gowers' proof. Since the subsequent arguments are very complicated, we confine ourselves for simplicity to the case $k = 4$ and follow here the simpler paper [23]. In this paper Gowers proved the weaker estimate

$$a_4(N) \ll 1/(\log\log\log N)^c, \tag{46}$$

where c is a constant.

Here is a sketch of the proof. Suppose that $A \subseteq \mathbb{Z}_N$, $|A| = \delta N$. If A is α-uniform with a sufficiently small α and if the number N is sufficiently large, $N \geqslant 32k^2\delta^{-k}$, then the set A contains a progression by Corollary 2, and we obtain the desired result. Thus, we can assume that the original set A is not α-uniform for any $\alpha = \alpha_0$. In the second part of his paper Gowers proves that for any set which is not α_0-uniform there is an arithmetic progression P of length at least N^β such that $|A \cap P| \geqslant (\delta + \varepsilon)|P|$, where β and ε depend only on δ and α_0. Gowers then applies the same arguments to the new set $A' = A \cap P$, whose density in P is at least $\delta + \varepsilon$, and so on. After several iterations either we find an α_0-uniform subset of A, or the density of A in some progression becomes sufficiently close to one. If A contains an α_0-uniform subset, then this subset (and hence the set A) contains an arithmetic progression by Corollary 2. If the density of A in some progression exceeds $3/4$, then the existence of an arithmetic progression of length four in A becomes obvious. Arguing as in the proof of Roth's theorem and estimating the number of iterations in terms of β and ε, we obtain the inequality (46).

We make a convention concerning the notation. In what follows, the symbols α_i will denote parameters depending on α_0 polynomially. In [22] all the parameters α_i are expressed in terms of α_0; however, for our purposes, it is more convenient not to express these parameters explicitly.

Gowers' proof depends heavily on a remarkable theorem of G. A. Freiman [64] (see also [65], [66], and [67]). Let $D \geqslant 1$ be an integer and let $E, F \subseteq \mathbb{Z}^D$. The Minkowski sum $E + F$ of these sets is defined to be the set $E + F = \{e + f : e \in E, f \in F\}$. The difference between two sets is defined similarly. If the cardinality of $E + E$ does not exceed $|E|$ too much (for instance, if $|E + E| \leqslant C|E|$ for some constant $C > 1$), then E is called a set with small sumsets. Freiman's theorem describes the structure of all sets of this kind. It is clear that every arithmetic progression is a set with small sumsets. One can readily see that every set of the

form $P_1 + \cdots + P_s$, where P_i is an arithmetic progression, is also a set with small sumsets. Such sets are called *d-dimensional arithmetic progressions*. Moreover, any large subset of $P_1 + \cdots + P_s$ is a set with small sumsets. Freiman's theorem asserts that there are no other examples of sets with small sumsets. We present the exact formulation.

Theorem 23 (Freiman). *Let $C > 0$, let $D \geqslant 1$, and let $A \subseteq \mathbb{Z}^D$ be a set such that $|A + A| \leqslant C|A|$. Then there exist numbers d and K, depending only on C and D, and a d-dimensional arithmetic progression Q such that $|Q| \leqslant K|A|$ and $A \subseteq Q$.*

Freiman's theorem remains valid also for the difference $A - A$.

Let us return to Definition 6. As was proved in Theorem 21, a set $A \subseteq \mathbb{Z}_N$, $|A| = \delta N$, is α-uniform of degree d if and only if it contains approximately $\delta^{2^d} N^{d+1}$ d-dimensional cubes. There is another characterization of α-uniform sets. As we know, if $d = 2$, then A is α-uniform of degree two if and only if the Fourier coefficients of this set are small. It turns out that A is α-uniform of degree three if and only if 'almost all' sets of the form $A \cap (A + k)$, $k \in \mathbb{Z}_N$, have small Fourier coefficients (in terms of α). In this survey we cannot dwell at length on a characterization of α-uniform sets in terms of the Fourier coefficients of their subsets; we note only that such a characterization exists for any degree $d \geqslant 2$ (see [22]). We formulate the exact result in the case $d = 2$. Suppose that $f\colon \mathbb{Z}_N \to \mathbb{Z}_N$ is some function and that $k \in \mathbb{Z}_N$ is an arbitrary residue. The *difference function* $\Delta(f; k)\colon \mathbb{Z}_N \to \mathbb{Z}_N$ is defined to be the function $\Delta(f; k)(s) = f(s)f(s - k)$.

Assertion 1. *Let $A \subseteq \mathbb{Z}_N$, $|A| = \delta N$, be a set which is not α_0-uniform of degree 3. Let B be the set of integers k for which there is an $r = r(k)$ such that $|\Delta(f; k)\widehat{\ }(r)| \geqslant \alpha_1 N$. Then $|B| \geqslant \alpha_2 N$.*

It follows from the definition of the set B that the function $\varphi\colon B \to \mathbb{Z}_N$ such that $|\Delta(f; k)\widehat{\ }(\varphi(k))| \geqslant \alpha_1 N$ for any k in B is well defined on B. The following proposition shows that φ has a property similar to linearity.

Proposition 3. *There are at least $\alpha_3 N^3$ quadruples $(a, b, c, d) \in B^4$ such that $a + b = c + d$ and $\varphi(a) + \varphi(b) = \varphi(c) + \varphi(d)$.*

Consider the graph $\Gamma \subseteq \mathbb{Z}^2$ of the function φ, that is, $\Gamma = \{(b, \varphi(b)) : b \in B\}$. By Proposition 3, the set Γ admits at least $\alpha_3 N^3$ quadruples $(x, y, z, w) \in \Gamma^4$ such that $x + y = z + w$. These quadruples are said to be *additive*. Gowers proved that the sets having many additive quadruples have quite special properties.

Proposition 4. *Let $c_0 > 0$ and let $M \subseteq \mathbb{Z}^D$ be a set of cardinality m having at least $c_0 m^3$ additive quadruples. Then there are constants c and C depending only on c_0 and there is a set $M' \subseteq M$ with $|M'| \geqslant cm$ such that $|M' - M'| \leqslant Cm$.*

Combining Proposition 4 and Freiman's theorem, Theorem 23, Gowers obtains the following result.

Proposition 5. *Let $B \subseteq \mathbb{Z}_N$ be a set of cardinality βN and let the graph of a function $\varphi\colon B \to \mathbb{Z}_N$ have at least $c' N^3$ additive quadruples. Then there are constants γ and η depending only on β and c', an arithmetic progression $P \subseteq \mathbb{Z}_N$*

with $|P| \geqslant N^{\gamma}$, *and numbers* λ *and* μ *such that*

$$\sum_{k\in P} |\Delta(f;k)\widehat{\ }(\lambda k+\mu)|^2 \geqslant \eta|P|N^2. \tag{47}$$

We go through some very rough arguments to clarify the proof of Proposition 5. Let us apply Proposition 4 to the graph Γ. Then Γ has a subset Γ' with small difference. It follows from Freiman's theorem that there is a not very large d-dimensional arithmetic progression Q containing Γ'. On the other hand, every d-dimensional arithmetic progression can be partitioned into several progressions of the form $Q_1 \times Q_2$, where Q_1 and Q_2 are one-dimensional arithmetic progressions in $\mathbb{Z}$. Hence, there is a progression of the form $Q_1 \times Q_2$ that intersects Γ' in a fairly large set. But this means that the values of the function φ on Q_1 coincide very often with the values of some *linear* function. These considerations enable one to obtain the inequality (47). Using the Dirichlet principle, one can show that the lengths of the progressions Q_1 and Q_2 are not less than $|Q|^{1/d}$. This implies the inequality $|P| \geqslant N^{\gamma}$.

Next, Gowers proves the following proposition.

Proposition 6. *Under the assumptions of Proposition 5 there are polynomials* $\psi_0, \psi_1, \dots, \psi_{N-1}$ *of degree two such that*

$$\sum_s \left| \sum_{x\in P+s} f(x) e^{2\pi i \psi_s(x)/N} \right| \geqslant \frac{\eta|P|N}{\sqrt{2}}. \tag{48}$$

Gowers then uses Weil's estimates for trigonometric sums involving polynomials of degree two and derives the following consequence of Proposition 6.

Proposition 7. *Under the assumptions of Proposition 5 there exist a* $\zeta > 0$, *an* $m \leqslant |P|^{\zeta}$, *and progressions* P_{sj} *with* $s \in [N]$ *and* $j \in [m]$ *such that for any* $s \in [N]$ *the progressions* $P_{s1}, \dots, P_{sm}$ *partition the progression* $P + s$ *and*

$$\sum_s \sum_{j=1}^{m} \left| \sum_{x\in P_{sj}} f(x) \right| \geqslant \frac{\eta|P|N}{2\sqrt{2}}. \tag{49}$$

Since f is the balance function of the set A, we have $\sum_s \sum_{j=1}^m \sum_{x\in P_{sj}} f(x) = 0$, and the inequality (49) readily implies that there is a progression P_{sj} for which

$$\sum_{x\in P_{sj}} f(x) \geqslant \frac{\beta|P|}{4m\sqrt{2}} \geqslant \frac{\beta|P_{sj}|}{4\sqrt{2}}.$$

Thus, we have found an arithmetic progression P' satisfying the inequality $|A \cap P'| \geqslant (\delta + \varepsilon)|P'|$, where ε depends only on δ and α_0. As was mentioned above, we eventually find an arithmetic progression in A by repeating this procedure several times.

A simple corollary to Theorem 9 is the following assertion about a colouring of the set $[N]$ with two colours.

Corollary 3. *Let k and N be positive integers and let $N \geqslant 2^{2^{2^{2^{2^{k+9}}}}}$. Assume that the set $[N]$ is coloured with two colours. Then there is a monochromatic arithmetic progression of length k in $[N]$.*

Thus, as can be seen by the reader, Gowers' proof of Theorem 9 makes essential use of Freiman's theorem. Another application of Theorem 23 is due to S. L. G. Choi [68]. He combined Szemerédi's theorem with Freiman's theorem and obtained the following result.

Theorem 24 (Choi). *Let A be an arbitrary subset of $\mathbb{Z}$, let C be a positive constant, and let $k \geqslant 3$ be an integer. Assume that $|A+A| \leqslant C|A|$. Then there is a constant $\alpha = \alpha(C,k)$ such that the set A contains $\alpha|A|^2$ arithmetic progressions of length k.*

We complete the section with some remarks about the second Erdős–Turán conjecture, Conjecture 2.

Conjecture 3 (Erdős, Turán). *Let $A = \{n_1 < n_2 < \cdots\}$ be an infinite sequence of positive integers such that*

$$\sum_{i=1}^{\infty} \frac{1}{n_i} = \infty. \tag{50}$$

Then A contains an arithmetic progression of arbitrary length.

There is a close relationship between estimates for the quantity $a_k(N)$ and Conjecture 3.

Assertion 2. *Conjecture 3 is valid if and only if the condition*

$$\sum_{l=1}^{\infty} a_k(4^l) < \infty \tag{51}$$

holds for any $k \geqslant 3$.

Proof. To prove this, we need a simple lemma about the numbers $a_k(N)$.

Lemma 2. *Let $k \geqslant 3$ be an integer and let x and y be positive integers with $x \leqslant y$. Then*

$$a_k(y) \leqslant \frac{x}{y}\left(\left\lfloor \frac{y}{x} \right\rfloor + 1\right) a_k(x) \leqslant 3a_k(x). \tag{52}$$

Proof. Let $M \subseteq [y]$ be a maximal set without arithmetic progressions of length k, and consider the disjoint segments

$$\Delta_1 = [1,x], \quad \Delta_2 = [x+1, 2x], \quad \ldots, \quad \Delta_s = [(s-1)x+1, sx],$$

where $s = \lfloor y/x \rfloor + 1$. It is clear that the segment $[y]$ is contained in the union of these disjoint segments. Since M contains no arithmetic progressions of length k, it follows that $|M \cap \Delta_j| \leqslant x a_k(x)$ for any $j = 1, \ldots, s$. Hence,

$$a_k(y)y = |M| = \sum_{j=1}^{s} |M \cap \Delta_j| \leqslant \sum_{j=1}^{s} x a_k(x) = x\left(\left\lfloor \frac{y}{x} \right\rfloor + 1\right) a_k(x). \tag{53}$$

Remark 2. In Lemma 2 we have established the simplest inequality for the functions $a_k(N)$. Further results about $a_k(N)$ can be found in the paper [7] of Roth and in the recent paper [62] of Croot.

We now prove Assertion 2.

Sufficiency. Suppose that the series $\sum_{l=1}^{\infty} a_k(4^l)$ converges for any $k \geqslant 3$. We assume that the Erdős–Turán conjecture fails. In other words, for some $k_0 \geqslant 3$ one can find an infinite sequence $A = \{n_1 < n_2 < \dots\}$ of positive integers that contains no arithmetic progressions of length k_0 and for which the series (50) diverges. Let $N = 16$. We partition the set $\mathbb{N}$ of positive integers into the subsets $C_0 = [1, N)$, $C_1 = [N, 4N)$, $C_2 = [4N, 4^2N), \dots$, $C_l = [4^{l-1}N, 4^lN), \dots$. Since A contains no arithmetic progressions of length k_0, it follows that $|A \cap C_l| \leqslant 4^{l-1}N \cdot a_{k_0}(4^{l-1}N)$ for any $l = 0, 1, 2, \dots$. We have

$$\sum_{i=1}^{\infty} \frac{1}{n_i} = \sum_{l=0}^{\infty} \sum_{n_i \in C_l} \frac{1}{n_i} \leqslant 4 \sum_{l=0}^{\infty} \frac{|A \cap C_l|}{4^{l-1}N} \leqslant 4 \sum_{l=0}^{\infty} a_{k_0}(4^{l+1}) = 4 \sum_{l=1}^{\infty} a_{k_0}(4^l) < \infty\,. \tag{54}$$

The inequality (54) contradicts the inequality (50). This completes the proof of sufficiency.

Necessity. Let the series $\sum_{l=1}^{\infty} a_{k_0}(4^l)$ diverge for some $k_0 \geqslant 3$. As in the proof of the equivalence of Conjectures 1 and 1′, we construct two sequences of positive integers. Let $N_1 = 1$ and $b_1 = 1$ and let

$$N_l := b_{l-1} + N_{l-1}, \qquad b_l := b_{l-1} + N_{l-1} + N_l + 1, \tag{55}$$

for any $l \geqslant 2$. We obtain an increasing sequence of positive integers $1 = N_1 < N_2 < N_3 < \cdots$. For any $l = 1, 2, \dots$ there is a set $A_l \subseteq [N_l]$ containing no arithmetic progressions of length k_0 and such that $|A_l| = a_{k_0}(N_l)N_l$. Let $\widetilde{A}_l = A_l + b_l$. It is clear that the sets $\widetilde{A}_l$ are disjoint and contain no arithmetic progressions of length k_0. Let $A = \bigsqcup_i \widetilde{A}_i$, $A = \{n_1 < n_2 < \dots\}$. Using (55), we see that A also contains no arithmetic progressions of length k_0. We have $b_l \leqslant 3N_l$ for any $l \geqslant 1$. Similarly, $N_{l+1} \leqslant 4N_l$ for $l \geqslant 1$. Hence $N_l \leqslant 4^l$, $l \geqslant 1$. By Lemma 1,

$$\sum_{i=1}^{\infty} \frac{1}{n_i} = \sum_{l=1}^{\infty} \sum_{n_i \in [b_l, b_l + N_l]} \frac{1}{n_i} \geqslant \sum_{l=1}^{\infty} \frac{a_{k_0}(N_l)N_l}{b_l + N_l} \geqslant \frac{1}{4} \sum_{l=1}^{\infty} a_{k_0}(N_l) \geqslant \frac{1}{12} \sum_{l=1}^{\infty} a_{k_0}(4^l). \tag{56}$$

It follows from (56) that the series $\sum_{i=1}^{\infty} 1/n_i$ diverges. Applying the Erdős–Turán conjecture, we see that the set A contains an arithmetic progression of length k_0, a contradiction. This completes the proof of Assertion 2.

§ 6. Ergodic approach to Szemerédi's theorem

Let X be a set equipped with a σ-algebra $\mathscr{B}$ of sets. Also, let T be a measurable self-map of X preserving the measure μ. Everywhere below we assume that $\mu(X) = 1$. The quadruple $(X, \mathscr{B}, \mu, T)$ is called a *dynamical system with invariant measure*. The well-known **Poincaré recurrence theorem** [69] asserts that for every measurable set $E \subseteq X$, $\mu(E) > 0$, there is an integer $n > 0$ such that $\mu(E \cap T^{-n}E) > 0$.

In [15] (see also [17] and [16]) Furstenberg generalized the Poincaré theorem to the case of several powers of the map T.

Theorem 25 (Furstenberg). *Let X be a space with a σ-algebra $\mathscr{B}$ of measurable sets and let μ be a measure on X. Let T be a self-map of X preserving the measure μ and let $k \geqslant 3$. Then for any measurable set E with $\mu(E) > 0$ there is an integer $n > 0$ such that*

$$\mu(E \cap T^{-n}E \cap T^{-2n}E \cap \dots \cap T^{-(k-1)n}E) > 0.$$

In this section we show that Theorem 25 is equivalent to Szemerédi's theorem. Hence, Furstenberg obtained an alternative proof of the first Erdős–Turán conjecture, Conjecture 1, by ergodic theory methods. Let us formulate Szemerédi's theorem once again.

Theorem 26 (Szemerédi). *Let A be an arbitrary subset of the set of positive integers and let $D^*(A) > 0$. Then A contains arithmetic progressions of length k for any integer $k \geqslant 3$.*

One can readily show (see [17] or [16]) that Theorem 25 follows from Theorem 26. Indeed, let N be a sufficiently large positive integer (we shall specify this number below). For any $x \in X$ we consider the set $\Lambda(x) = \{l \in [N] : T^l x \in E\}$. We have

$$\int_X |\Lambda(x)|\, d\mu = N\mu(E). \tag{57}$$

Let $M = \{x \in X : |\Lambda(x)| \geqslant N\mu(E)/2\}$. It follows from the inequality (57) that $\mu(M) \geqslant \mu(E)/2$. As was shown in the Introduction, Conjecture 1 (in other words, Theorem 26) is equivalent to Conjecture 1′. Let $N = N(k, \mu(E)/2)$. For any $x \in M$ the set $\Lambda(x)$ contains an arithmetic progression $\{a(x) + b(x)m\}_{m=0}^{k-1}$ of length k. Thus, to any point x in M we have assigned a pair of numbers $(a(x), b(x))$. Since for any $x \in X$ we have $(a(x), b(x)) \in [N]^2$, there is a set $M' \subseteq M$ such that $\mu(M') \geqslant \mu(E)/(2N^2)$, and to any point of this set we have assigned the same pair (a, b). In this case, $\mu\bigl(\bigcap_{m=0}^{k-1} T^{-(a+bm)}E\bigr) \geqslant \mu(M') > 0$. The map T preserves the measure μ, and hence

$$\mu\biggl(\bigcap_{m=0}^{k-1} T^{-(a+bm)}E\biggr) = \mu\biggl(T^{-a}\bigcap_{m=0}^{k-1} T^{-(bm)}E\biggr) = \mu\biggl(\bigcap_{m=0}^{k-1} T^{-(bm)}E\biggr) > 0,$$

as was to be proved.

Thus, we have proved that Theorem 25 follows from Theorem 26. In fact, Theorem 25 is equivalent to Theorem 26. To prove their equivalence, Furstenberg established the following beautiful result, the so-called *Furstenberg correspondence principle* (see [15]).

Theorem 27 (Furstenberg). *Let A be an arbitrary subset of the set of positive integers with $D^*(A) > 0$. Then there exist a dynamical system $(X, \mathscr{B}, \mu, T)$ with invariant measure and a measurable set E with $\mu(E) = D^*(A)$ such that*

$$D^*(A \cap (A + m_1) \cap \dots \cap (A + m_{k-1})) \geqslant \mu(E \cap T^{-m_1}E \cap \dots \cap T^{-m_{k-1}}E) \tag{58}$$

for any integer $k \geqslant 3$ and any positive integers $m_1, m_2, \dots, m_{k-1}$.

Theorem 27 shows the existence of a close relationship between ergodic theory and combinatorial problems on arithmetic progressions.

Assertion 3 (Furstenberg). *Theorems 25 and 27 imply Theorem 26.*

Proof. Let k be a positive integer, $k \geqslant 3$, and let $A \subseteq \mathbb{N}$ be a set with no arithmetic progressions of length k and with positive upper density. By Theorem 27, there exist a dynamical system $(X, \mathscr{B}, \mu, T)$ and a measurable set E of positive measure such that the inequality (58) holds for any positive integers $m_1, m_2, \dots, m_{k-1}$. On the other hand, by Theorem 25, there is an integer $n > 0$ such that

$$\mu(E \cap T^{-n}E \cap T^{-2n}E \cap \dots \cap T^{-(k-1)n}E) > 0. \tag{59}$$

We set $m_1 = n$, $m_2 = 2n$, $\dots$, $m_{k-1} = (k-1)n$. It follows from (58) and (59) that $D^*(A \cap (A+n) \cap \dots \cap (A+(k-1)n)) > 0$. This contradicts the assumption that A contains no arithmetic progressions of length k and thus proves Assertion 3.

In this survey we do not give a complete proof of Theorem 27. The interested reader can find a detailed proof of the correspondence principle in the paper [15] or the book [17]. Nevertheless, we shall try to schematically show how to construct a dynamical system with the desired properties from an arbitrary subset A of positive integers with $D^*(A) > 0$ in Theorem 27.

Thus, let a set $A \subseteq \mathbb{N}$ have positive upper Banach density. Let $\Omega = \{0,1\}^{\mathbb{N}}$ be the space of one-sided infinite sequences of zeros and ones. Let T be the self-map of Ω given by the formula $(T\omega)_i = \omega_{i+1}$. Thus, T is the left shift. Also, let $\omega' \in \Omega$ be the infinite sequence such that $\omega'_i = 1$ if and only if $i \in A$. We take an arbitrary metric on the space Ω and consider the closure X of the orbit of the point ω' under the action of the map T. In other words, let $X = \overline{\{T^i\omega'\}_{i=0}^{\infty}}$. It is clear that X is T-invariant. Let $C_0 = \{\omega \in \Omega : \omega_0 = 1\}$ be an elementary cylinder and let $E = C_0 \cap X$. It follows from the definition of ω' that A contains an arithmetic progression of length k if and only if the set $\bigcap_{m=0}^{k-1} T^{-mn}E$ is non-empty for some integer $n \geqslant 1$ (note that the set $\bigcap_{m=0}^{k-1} T^{-mn}E$ is open).

We do not give a proof of the inequality (58), but only construct a T-invariant probability measure on X such that $\mu(E) = D^*(A) > 0$. Since $D^*(A) > 0$, there is an increasing sequence of positive integers n_k such that $\lim_{k\to\infty} |A \cap [n_k]|/n_k = D^*(A)$. Let $\mu_k := 1/n_k \cdot \sum_{i=0}^{n_k-1} \delta_{T^i\omega'}$, where δ_x stands for the Dirac measure on X with $\delta_x(M) = 1$ if and only if $x \in M$. It is clear that μ_k is a probability measure on X for any k. Moreover, for any set $M \subseteq X$ we have

$$\bigl|\mu_k(T^{-1}M) - \mu_k(M)\bigr| \leqslant \frac{2}{n_k}. \tag{60}$$

Let μ be the $*$-weak limit of the probability measures μ_k (for the definition of $*$-weak limit, see, for instance, [70]). Then it follows from the inequality (60) that μ is a T-invariant probability measure on X. Moreover,

$$\mu(E) = \lim_{k\to\infty} \mu_k(E) = \lim_{k\to\infty} \frac{|A \cap [n_k]|}{n_k} = D^*(A) > 0,$$

as was to be proved.

Using the ergodic approach, Furstenberg, Y. Katznelson, D. Ornstein, and other authors obtained a lot of generalizations of Szemerédi's theorem. We cannot cover all the results in this direction in the present survey, so we confine ourselves to only a few of them. In [18] Furstenberg and Katznelson extended Theorem 25 to the case of several *commuting* maps.

Theorem 28 (Furstenberg, Katznelson). *Let X be a space with a σ-algebra $\mathscr{B}$ of measurable sets and let μ be a measure on X. Let $k \geqslant 2$ and let $T_1, \dots, T_k$ be commuting self-maps of X preserving the measure μ. Then for any measurable set E with $\mu(E) > 0$ there is an integer $n > 0$ such that*

$$\mu(E \cap T_1^{-n} E \cap T_2^{-n} E \cap \dots \cap T_k^{-n} E) > 0.$$

In § 7 we shall prove a quantitative version of Theorem 28 in the case of *two* commuting maps. It should be remarked that Theorem 28 was recently proved in [21] for arbitrary soluble groups.

We also note the remarkable Bergelson–Leibman theorem ([19]; see also [71], [133]), already mentioned in the Introduction.

Theorem 29 (Bergelson, Leibman). *Let X be a set, let $\mathscr{B}$ be a σ-algebra of measurable sets on X, and let μ be a finite measure on X with $\mu(X) > 0$. Let $k \geqslant 2$, let $T_1, \dots, T_k$ be invertible commuting self-maps of X that preserve the measure μ, and let $p_1(n), \dots, p_k(n)$ be polynomials having rational coefficients, taking integral values at all integral values of n, and satisfying $p_i(0) = 0$, $i = 1, \dots, k$. Then*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(T_1^{-p_1(n)} E \cap T_2^{-p_2(n)} E \cap \dots \cap T_k^{-p_k(n)} E) > 0$$

for any measurable set E with $\mu(E) > 0$.

A. Sárközy ([72], [73]) obtained a quantitative version of Theorem 29 in the case when $k = 2$, $T_2 = T_1^2$, $p_1(n) = n$, and $p_2(n) = n^2$ (see also [74] and [25]). For the present, the best result on this topic is due to J. Pintz, W. L. Steiger, and Szemerédi (see [75]).

Theorem 30 (Pintz, Steiger, Szemerédi). *Let A be an arbitrary subset of $[N]$ and let*

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

where $c > 0$ is an absolute constant. Then A contains two elements a and a' such that $a' - a$ is a perfect square.

Using the ergodic technique, Furstenberg and Katznelson [20] re-proved the well-known Hales–Jewett theorem [76] generalizing van der Waerden's theorem (Theorem 1). To formulate the Hales–Jewett theorem, we need several definitions.

Let

$$C_t^n := \{x_1 x_2 \dots x_n : x_i \in \{0, 1, \dots, t-1\}\}$$

be the set of all words of length n composed of the symbols $0, 1, \dots, t-1$.

Definition 7. A *combinatorial line* in C_t^n is a family of t words $X_1, \dots, X_t \in C_t^n$, where $X_i = x_{i1} \dots x_{in}$, such that for some non-empty set $J \subseteq \{1, \dots, n\}$ we have $x_{sj} = s$ if $j \in J$, and for any $j \notin J$ there is a symbol $c_j \in \{0, 1, \dots, t-1\}$ such that $x_{1j} = \dots = x_{tj} = c_j$.

Example. Let $t = 3$ and $n = 5$. Then the words 01012, 11112, 21212 form a combinatorial line in C_3^5.

Theorem 31 (Hales, Jewett). *Let t and r be positive integers. Then there is a positive integer $N = N(t, r)$ such that for any $n \geqslant N$ and for any colouring of C_t^n with r colours there is a monochromatic combinatorial line.*

It is clear that Theorem 31 implies van der Waerden's theorem in the form of Theorem 2 (and hence van der Waerden's theorem in the form of Theorem 1). Indeed, in the number system with base t one can interpret the set C_t^n as the family of numbers from 0 to $t^n - 1$. In this case for any combinatorial line in C_t^n there is a corresponding arithmetic progression of length t in the set $\{0, 1, \dots, t^n - 1\}$.

As already mentioned above, Theorem 25 was the origin of a new area of ergodic theory, namely, combinatorial ergodic theory. In the present survey we cannot even list the results obtained in this area. We can only indicate the remarkable book [17] and the monograph [77] to the interested reader. We also note that quantitative versions of the Poincaré theorem were obtained in the recent papers [78]–[86] and versions of Theorem 27 can be found in [87].

As was noted above, the original proof of Szemerédi's theorem, Furstenberg's proof, and Gowers' proof are closely related. We would like to close this section with a discussion of the main ideas at the basis of Furstenberg's approach.

In the first (preliminary) part of his proof Furstenberg shows the validity of Theorem 25 for two special classes of dynamical systems: weakly mixing and compact dynamical systems. We give rigorous definitions of these classes.

A dynamical system is said to be *weakly mixing* if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Big(\mu(A \cap T^{-n} B) - \mu(A)\mu(B) \Big)^2 = 0$$

for any two sets $A, B \in \mathscr{B}$. With every dynamical system $(X, \mathscr{B}, \mu, T)$ one can associate a unitary operator U acting in $L_2(X, \mathscr{B}, \mu)$ by the rule $(Uf)(x) = f(Tx)$ for any function $f\colon X \to \mathbb{C}$. The function identically equal to one is always an eigenvector of this operator. A dynamical system is said to be *compact* if the spectrum of U is discrete. We give another definition. A system is said to be *compact* if the closure in $L_2(X, \mathscr{B}, \mu)$ of the orbit $\{U^n f\}_{n=0}^{\infty}$ is compact for any function $f \in L_2(X, \mathscr{B}, \mu)$. One can show that a dynamical system cannot be weakly mixing and compact simultaneously (for details, see [17]).

As noted above, Theorem 25 holds for the dynamical systems in these two classes. Nevertheless, the reasons why the theorem holds for these classes are quite different. Weakly mixing dynamical systems have strong random properties, which enables one to prove an assertion stronger than Theorem 25 for these systems.

Theorem 32 (Furstenberg). *Let $(X, \mathscr{B}, \mu, T)$ be a weakly mixing dynamical system and let $A_0, A_1, \dots, A_k \in \mathscr{B}$, $k \geqslant 1$, be arbitrary measurable sets. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \bigl(\mu(A_0 \cap T^{-n} A_1 \cap T^{-2n} A_2 \cap \dots \cap T^{-kn} A_k) - \mu(A_0)\mu(A_1)\dots\mu(A_k)\bigr)^2 = 0.$$

On the other hand, it is quite clear that Theorem 25 holds for periodic maps T, that is, for maps such that T^p is the identity map for some $p \in \mathbb{N}$. Compact dynamical systems form a very broad generalization of periodic systems. Consider the following example. Let X be the unit circle $\mathbb{S}^1$, let μ be the Lebesgue measure on X, and let $\mathscr{B}$ be the σ-algebra of Lebesgue measurable sets. Let the map T_α be the rotation of the circle $\mathbb{S}^1$ given by the formula $T_\alpha x = x + \alpha \pmod 1$, where α is an arbitrary real number (not necessarily rational). In this case the dynamical system $(X, \mathscr{B}, \mu, T_\alpha)$ is compact. It is clear that the map T_α is not periodic if the number α is irrational; however, it is also clear that T_α is 'almost periodic', in the sense that, for example, for any interval I and any $\varepsilon > 0$ there is a number n such that $\mu(T^n I \triangle I) < \varepsilon$. Using 'almost periodic' properties of compact dynamical systems, Furstenberg proved that Theorem 25 holds for them. We note that the stronger Theorem 32 fails for these systems (see [16]). The point is that compact systems fail to have nice random properties.

At the second stage of his proof, Furstenberg gives a characterization of weakly mixing dynamical systems by using factor systems. Let $\mathscr{B}_1 \subseteq \mathscr{B}$ be a sub-σ-algebra of the σ-algebra $\mathscr{B}$, and assume that the map T is measurable with respect to this sub-σ-algebra, in other words, $T^{-1}A \in \mathscr{B}_1$ for any $A \in \mathscr{B}_1$. Then the dynamical system $(X, \mathscr{B}_1, \mu, T)$ is called a *factor system* of the dynamical system $(X, \mathscr{B}, \mu, T)$. If the system $(X, \mathscr{B}_1, \mu, T)$ is compact, then we say that the factor system $(X, \mathscr{B}_1, \mu, T)$ is compact. A factor system is said to be *non-trivial* if it contains sets whose measure differs both from zero and from one.

Theorem 33 (Furstenberg). *A dynamical system $(X, \mathscr{B}, \mu, T)$ is weakly mixing if and only if it has no non-trivial compact factor systems.*

It follows from Theorem 33 that every dynamical system $(X, \mathscr{B}, \mu, T)$ has a non-trivial factor system for which Theorem 25 holds. Indeed, if $(X, \mathscr{B}, \mu, T)$ is a weakly mixing dynamical system, then Theorem 32 holds for it. However, if $(X, \mathscr{B}, \mu, T)$ is not weakly mixing, then by Theorem 33 it admits a non-trivial compact factor system. Since the factor system is compact, we see that Theorem 25 holds for it.

Let $\mathscr{M}$ be the family of non-trivial factor systems of the original dynamical system $(X, \mathscr{B}, \mu, T)$ that satisfy Theorem 25 (in fact, Furstenberg considers in his proof a somewhat different family of factor systems). As we have just proved, the family $\mathscr{M}$ is non-empty. Furstenberg proved two facts about $\mathscr{M}$. The first is that $\mathscr{M}$ has a maximal element, and the second is that this maximal element coincides with $(X, \mathscr{B}, \mu, T)$. It is clear that these two facts imply Theorem 25.

The proof of the existence of a maximal element in $\mathscr{M}$ is rather simple and uses Zorn's lemma. The proof that a maximal element of the family $\mathscr{M}$ coincides with the original dynamical system is complicated and occupies almost half of the paper [16]. We can only briefly present the corresponding ideas.

First, Furstenberg generalizes the notions of weakly mixing and compact dynamical systems to factor systems of $(X, \mathscr{B}, \mu, T)$. Let $\mathscr{B}_1 \subseteq \mathscr{B}_2$ be two σ-algebras. Furstenberg gives definitions of compactness and weak mixing for the 'larger' factor system $(X, \mathscr{B}_2, \mu, T)$ with respect to the 'smaller' factor system $(X, \mathscr{B}_1, \mu, T)$ and proves an assertion about a lifting, namely, if Theorem 25 holds for the 'smaller' factor system, then it also holds for the 'larger' one. Second, Furstenberg proves that if a factor system $(X, \mathscr{B}_1, \mu, T)$ of the initial system $(X, \mathscr{B}, \mu, T)$ is not weakly mixing, then there is a factor system $(X, \mathscr{B}_1', \mu, T)$, $\mathscr{B}_1 \subsetneq \mathscr{B}_1'$, such that $(X, \mathscr{B}_1', \mu, T)$ is compact with respect to $(X, \mathscr{B}_1, \mu, T)$.

Suppose that a *maximal* factor system $F_0 = (X, \mathscr{B}_0, \mu, T)$ in the family $\mathscr{M}$ differs from $F = (X, \mathscr{B}, \mu, T)$. If F_0 is relatively weakly mixing, then $F \in \mathscr{M}$, and we arrive at a contradiction to the maximality of F_0. However, if F_0 is not relatively weakly mixing, then there is a factor system F_0' which is larger than F_0 and compact with respect to F_0. Thus, $F_0' \in \mathscr{M}$, and we again arrive at a contradiction to the maximality of F_0. This argument completes Furstenberg's proof.

It should be noted that Furstenberg's arguments give no upper bounds for the quantity $N(k, \delta)$. As we had seen above, Furstenberg makes essential use of Zorn's lemma in his proof. Nevertheless, a new proof of Szemerédi's theorem recently appeared (see [43]) which uses methods of ergodic theory and gives quantitative bounds for $N(k, \delta)$ (though very weak).

§ 7. Two-dimensional generalizations of Szemerédi's theorem

Consider the two-dimensional lattice $[1, N]^2$ with basis $\{(1, 0), (0, 1)\}$. Let

$$L(N) = \frac{1}{N^2} \max\big\{|A| : A \subseteq [N]^2 \quad \text{and } A \text{ contains} \\ \text{no triples of the form } (k, m), (k + d, m), (k, m + d),\ d > 0\big\}. \tag{61}$$

Any triple of the form given in (61) will be called a *corner*. M. Ajtai and Szemerédi proved in [88] that the quantity $L(N)$ tends to zero as N tends to infinity. More precisely, they obtained the following result.

Theorem 34. *Let $S = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ be the unit square, let $0 < \delta \leqslant 1$, and let N be a positive integer. Then every set $A \subseteq [N]^2$ with $|A| \geqslant \delta N^2$ contains an affine image of S, that is, a set of the form $aS + b$, where $a \in \mathbb{N}$ and $b \in \mathbb{N}^2$.*

Theorem 34 was later re-proved by Furstenberg in [17] (for details, see § 5).

It is clear that the problem about corners is a two-dimensional generalization of the problem about sets without arithmetic progressions of length three. More precisely, it follows from the equality $\lim_{N\to\infty} L(N) = 0$ that $\lim_{N\to\infty} a_3(N) = 0$.

Indeed, suppose that $\lim_{N\to\infty} L(N) = 0$ but $\lim_{N\to\infty} a_3(N) = a > 0$ (the fact that the limit $\lim_{N\to\infty} a_3(N)$ exists was proved in the Introduction). The last equality means that for any sufficiently large N there is a set $A_N \subseteq [N]$ without progressions of length three and whose cardinality is not less than $aN/2$. Consider the square $Q_N = \{1, \dots, 2N\}^2$ and the set $\widetilde{A}_N = \bigsqcup_{i=1}^{N} \big((A_N + i) \times \{i\}\big) \subseteq Q_N$. The two-dimensional set $\widetilde{A}_N$ is the union of N shifts of the set A_N in the direction of the upper right-hand corner of the square Q_N. One can easily see that if $\widetilde{A}_N$ contains a corner, then A_N contains an arithmetic progression of length three, which is not

the case. Moreover, the cardinality of $\widetilde{A}_N$ is equal to $|A_N|N \geqslant aN^2/2$. This implies that $\limsup_{N\to\infty} L(N) > 0$, contradicting the equality $\lim_{N\to\infty} L(N) = 0$.

Gowers (see [22]) posed the problem of the rate of convergence of $L(N)$ to 0. Developing the approach in the papers [90] and [91], V. H. Vu [89] proposed the following solution of this problem. Let $\log_{[1]} N = \log N$ and let $\log_{[l]} N = \log(\log_{[l-1]} N)$ for $l \geqslant 2$. Thus, $\log_{[l]} N$ is the result of taking the logarithm l times of the number N. Further, let k be the largest positive integer such that $\log_{[k]} N \geqslant 2$. We then set $\log_* N = k$. Vu proved that

$$L(N) \leqslant \frac{100}{(\log_* N)^{1/4}}.$$

The following result was obtained in [28] and [29].

Theorem 35 ([28], [29]). *Let $\delta > 0$, let $N \gg \exp\exp\exp(\delta^{-C})$ for some effective constant $C > 0$, and let $A \subseteq \{1,\dots,N\}^2$ be an arbitrary subset with cardinality at least δN^2. Then A contains a triple of the form (k,m), $(k+d,m)$, $(k,m+d)$, where $d > 0$.*

This theorem was later sharpened (see [30] and [31]).

Theorem 36 ([30], [31]). *Let $\delta > 0$, let $N \gg \exp\exp(\delta^{-c})$, where $c > 0$ is an absolute constant, and let $A \subseteq \{1,\dots,N\}^2$ be an arbitrary subset with cardinality at least δN^2. Then A contains a triple of the form (k,m), $(k+d,m)$, $(k,m+d)$, where $d > 0$.*

Thus, the following upper bound for the quantity $L(N)$ is obtained:

$$L(N) \ll \frac{1}{(\log\log N)^{C_1}},$$

where $C_1 = 1/c$.

The problem of upper bounds for the function $L(N)$ in the groups $\mathbb{Z}_p^n$, where p is a prime, was considered in [92]. In these groups a *corner* is defined to be a triple of the form (k,m), $(k+d,m)$, $(k,m+d)$ with $d \neq 0$. For simplicity we restrict ourselves to the case $p = 2$ in the present survey.

Theorem 37 (Green). *Let $\delta > 0$, let N and n be positive integers, let $N = 2^n$ and $N \gg \exp\exp(\delta^{-c'})$, where $c' > 0$ is an absolute constant, and let $A \subseteq \mathbb{Z}_2^n \times \mathbb{Z}_2^n$ be an arbitrary subset with cardinality at least δN^2. Then A contains a triple of the form (k,m), $(k+d,m)$, $(k,m+d)$, where $d \neq 0$.*

The proofs of Theorems 35 and 36 are rather complicated due to the very many technical details. In this survey we confine ourselves to a presentation of the simpler Theorem 37. Here the main ideas of the proofs of Theorems 35 and 36 are preserved but the technical complications are minimized.

We present the proof according to [92]. Before proving Theorem 37, we give several definitions.

Let $x, y \in \mathbb{Z}_2^n$. The inner product is given by

$$x \cdot y = x_1 y_1 + \dots + x_n y_n.$$

Let $f\colon \mathbb{Z}_2^n \to \mathbb{C}$ be a function and denote by $\widehat{f}(\xi)$ its Fourier coefficients:

$$\widehat{f}(\xi) = \sum_{x\in\mathbb{Z}_2^n} f(x)e(-(x\cdot\xi)),$$

where $e(x) = e^{2\pi i x}$.

We use several facts about the Fourier transform:

$$\sum_{x\in\mathbb{Z}_2^n} |f(x)|^2 = \frac{1}{N}\sum_{\xi\in\mathbb{Z}_2^n} |\widehat{f}(\xi)|^2, \tag{62}$$

$$\sum_{x\in\mathbb{Z}_2^n} f(s)\overline{g(s)} = \frac{1}{N}\sum_{\xi\in\mathbb{Z}_2^n} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}, \tag{63}$$

$$\sum_{x\in\mathbb{Z}_2^n} |(f*g)(x)|^2 := \sum_{x\in\mathbb{Z}_2^n}\Big|\sum_{y\in\mathbb{Z}_2^n} f(y)\overline{g(y-x)}\Big|^2 = \frac{1}{N}\sum_{\xi\in\mathbb{Z}_2^n} |\widehat{f}(\xi)|^2|\widehat{g}(\xi)|^2\,. \tag{64}$$

Definition 8. Let $A\subseteq\mathbb{Z}_2^n$ and let $\alpha\in(0,1)$ be a parameter. The set A is said to be *α-uniform* if

$$\max_{\xi\neq 0} |\widehat{A}(\xi)| \leqslant \alpha N. \tag{65}$$

The following lemma helps to better understand the definition of α-uniform sets.

Lemma 3. *Let $\tau,\kappa\in(0,1)$ be two parameters, let $S_1,\dots,S_k\subseteq\mathbb{Z}_2^n$, and let $\varepsilon(k) = \kappa 2^k\tau^{-k/2}$. Suppose that $|S_i| = \sigma_i N$ and that every set S_i is $\varepsilon(k-1)$-uniform. Then there are at least $(1-2^k\tau)N^k$ vectors $(x_1,\dots,x_k)\in(\mathbb{Z}_2^n)^k$ such that*

$$\Big|\sum_y S_1(x_1+y)S_2(x_2+y)\cdots S_k(x_k+y) - \sigma_1\cdots\sigma_k N\Big| \leqslant \varepsilon(k)N. \tag{66}$$

Proof. We prove the lemma by induction. The case $k=1$ is trivial. Let $u = (x_1,\dots,x_{k-1})$ be a fixed vector. We write

$$F_u(y) = S_1(x_1+y)S_2(x_2+y)\cdots S_{k-1}(x_{k-1}+y).$$

Then

$$\sum_y S_1(x_1+y)S_2(x_2+y)\cdots S_k(x_k+y) = (F_u*S_k)(-x_k) = (F_u*S_k)(x_k). \tag{67}$$

Using the formulae (64) and (62) and the fact that the sets S_k are $\varepsilon(k-1)$-uniform, we obtain

$$\sum_{x_k}\Big((F_u*S_k)(x_k) - \sigma_k\sum_y F_u(y)\Big)^2 = N^{-1}\sum_{\xi\neq 0}|\widehat{F}_u(\xi)|^2|\widehat{S}_k(\xi)|^2 \leqslant \varepsilon(k-1)^2N^3\,. \tag{68}$$

By induction, there are at least $(1-2^{k-1}\tau)N^{k-1}$ values of u such that

$$\Big|\sum_y F_u(y) - \sigma_1\cdots\sigma_{k-1}N\Big| \leqslant \varepsilon(k-1)\tau N. \tag{69}$$

By (68), for these values of u we have

$$\sum_{x_k}((F_u * S_k)(x_k) - \sigma_1 \cdots \sigma_k N)^2 \leqslant 4\varepsilon(k-1)^2 N^3 . \tag{70}$$

By (70), for a fixed value of u the number of points x_k such that the inequality

$$|(F_u * S_k)(x_k) - \sigma_1 \cdots \sigma_k N| > \varepsilon(k)N$$

holds does not exceed

$$\frac{4\varepsilon(k-1)^2}{\varepsilon(k)^2} N \leqslant 2^{k-1}\tau N.$$

Hence, the total number of points $(x_1, \dots, x_k)$ for which the inequality (66) fails is not greater than

$$2^{k-1}\tau N + 2^{k-1}\tau N = 2^k \tau N.$$

This proves Lemma 3.

Thus, if the sets S_i are α-uniform for a sufficiently small number α, then 'almost all' shifts of this set have intersections with cardinality the same as if the sets S_i were random.

We need a lemma about non-α-uniform sets.

Lemma 4. *Let $A \subseteq \mathbb{Z}_2^n$ and $|A| = \delta N$. Suppose that A is not an α-uniform set; in other words, there is a $\lambda \neq 0$ such that $|\widehat{A}(\lambda)| > \alpha N$. Let $H = \langle\lambda\rangle^{\perp} \subseteq \mathbb{Z}_2^n$ be the space perpendicular to the vector λ. Then*

$$\sum_x (A * H)^2(x) \geqslant (\delta^2 + \alpha^2)|H|^2 N. \tag{71}$$

Proof. By the inequality (64),

$$\begin{aligned}\sum_x (A * H)^2(x) &= \frac{1}{N}\sum_{\xi \in \mathbb{Z}_2^n} |\widehat{A}(\xi)|^2 |\widehat{H}(\xi)|^2 \\ &\geqslant \frac{1}{N}\bigl(|\widehat{A}(0)|^2|\widehat{H}(0)|^2 + |\widehat{A}(\lambda)|^2|\widehat{H}(\lambda)|^2\bigr) \geqslant (\delta^2 + \alpha^2)|H|^2 N,\end{aligned}$$

which proves Lemma 4.

We need another definition of α-uniformity.

Definition 9. Let $f\colon \mathbb{Z}_2^n \times \mathbb{Z}_2^n \to D$ be a function. We introduce the *rectangular* norm $\|f\|$ of f by the formula

$$\|f\|^4 = \sum_{x,x',y,y'} f(x,y)\overline{f(x',y)f(x,y')}f(x',y'). \tag{72}$$

It is clear that

$$\|f\|^4 = \sum_{x,x'}\left|\sum_y f(x,y)\overline{f(x',y)}\right|^2, \tag{73}$$

and therefore the right-hand side of (72) is always non-negative. As was proved in [29], the map $\|\cdot\|$ is indeed a norm.

Definition 10. Let $\alpha \in (0,1)$ and let $E_1 \times E_2 \subseteq \mathbb{Z}_2^n \times \mathbb{Z}_2^n$. A function $f\colon E_1 \times E_2 \to D$ is said to be *α-uniform with respect to the rectangular norm* if

$$\|f\|^4 \leqslant \alpha |E_1|^2 |E_2|^2. \tag{74}$$

Let $A \subseteq \mathscr{P} = E_1 \times E_2$. The set A is said to be *α-uniform with respect to the rectangular norm* if the function $A - \delta_{\mathscr{P}}(A)\mathscr{P}$ is α-uniform with respect to this norm.

We call a quadruple of the form $\{(x,y),(x',y),(x,y'),(x',y')\}$ an *elementary rectangle*. A set $A \subseteq \mathbb{Z}_2^n \times \mathbb{Z}_2^n$ contains an elementary rectangle if A contains the four points of the rectangle.

Let $A \subseteq \mathscr{P} = E_1 \times E_2$. It is clear that the number of elementary rectangles in A is equal to $\|A\|^4$. Using the formula (73) and the Cauchy–Bunyakovskii inequality, we obtain

$$\begin{aligned}\|A\|^4 &= \sum_{x,x'} \left| \sum_y A(x,y)A(x',y) \right|^2 \geqslant \frac{1}{|E_1|^2} \left(\sum_{x,x'} \sum_y A(x,y)A(x',y) \right)^2 \\ &= \frac{1}{|E_1|^2} \left(\sum_y \left| \sum_x A(x,y) \right|^2 \right)^2 \geqslant \frac{1}{|E_1|^2|E_2|^2} \left(\sum_{x,y} A(x,y) \right)^4 \\ &= \delta_{\mathscr{P}}(A)^4 |E_1|^2 |E_2|^2.\end{aligned}$$

Thus, every set $A \subseteq \mathscr{P}$ contains at least $\delta_{\mathscr{P}}(A)^4 |E_1|^2 |E_2|^2$ elementary rectangles. On the other hand, one can easily show that the number of elementary rectangles in any random subset of $\mathscr{P}$ is approximately equal to $\delta_{\mathscr{P}}(A)^4 |E_1|^2 |E_2|^2$.

The sets which are α-uniform with respect to the rectangular norm are characterized by a clear combinatorial property (see [29]).

Theorem 38. *Let a set $A \subseteq E_1 \times E_2 = \mathscr{P}$ be α-uniform with respect to the rectangular norm. Then A contains at most $(\delta + \alpha^{1/4})^4 |E_1|^2 |E_2|^2$ elementary rectangles.*

Proof. Let $f = A - \delta_{\mathscr{P}}(A)\mathscr{P}$ be the balance function of the set A and let $\delta = \delta_{\mathscr{P}}(A)$. Since the set A is α-uniform, it follows that $\|f\|^4 \leqslant \alpha |E_1|^2 |E_2|^2$. We have $A = f + \delta\mathscr{P}$. Hence,

$$\|A(x,y)\|^4 \leqslant (\|f\| + \|\delta\mathscr{P}\|)^4 \leqslant (\delta + \alpha^{1/4})^4 |E_1|^2 |E_2|^2. \tag{75}$$

This proves Theorem 38.

Thus, the number of elementary rectangles in the set A which are α-uniform with respect to the rectangular norm is approximately the same as that for random subsets of $\mathscr{P}$. One can say that the less the number α is, the closer A is to a random set.

Let W be a subspace of $\mathbb{Z}_2^n$. It is clear that W is isomorphic to some space $\mathbb{Z}_2^m$ with $m \leqslant n$. We also assume that $E_1, E_2 \subseteq W$, $E_i = \beta_i |W|$, $\mathscr{P} = E_1 \times E_2$, and $A \subseteq \mathscr{P}$, $\delta_{\mathscr{P}}(A) = \delta$. The first step of the proof of Theorem 37 involves the situation in which the sets E_1 and E_2 are α-uniform and, moreover, the set A itself is α'-uniform with respect to the rectangular norm. We show that if α and α' are sufficiently small (that is, E_1, E_2, and A are close to being random), then A contains sufficiently many corners. The main tool in the proof is the Cauchy–Bunyakovskii inequality.

Theorem 39. *Suppose that the sets E_i are $(2^{-36}\beta_1^{12}\beta_2^{12}\delta^{36})$-uniform and the set A is $2^{-8}\delta^{12}$-uniform with respect to the rectangular norm. Then A contains at least $\delta^3\beta_1^2\beta_2^2|W|^3/2$ triples of the form $\{(x,y),(x+d,y),(x,y+d)\}$.*

Proof. Let $f_1, f_2, f_3\colon \mathscr{P} \to D$ be three arbitrary functions. Consider the functional $T(f_1,f_2,f_3) = \sum_{x,y,z} f_1(x,y)f_2(y+z,y)f_3(x,x+z)$. It is clear that T is linear in each of the arguments. Moreover, the value $T(A,A,A)$ is equal to the number of triples $\{(x,y),(x+d,y),(x,y+d)\}$ in A (here we use a specific feature of $\mathbb{Z}_2^n$). Let $f = A - \delta\mathscr{P}$ and $\eta = 2^{-8}\delta^{12}$. Then $T(A,A,A) = \delta T(\mathscr{P},A,A) + T(f,A,A)$ and $\|f\|^4 \leqslant \eta\beta_1^2\beta_2^2|W|^4$.

Let $g(z) = \sum_x A(x,x+z)$. Then $T(\mathscr{P},A,A) = \sum_z g(z)^2$. Since W is a subspace of $\mathbb{Z}_2^n$, it follows that $g(z)$ vanishes outside W. We have $\sum_z g(z) = \delta\beta_1\beta_2|W|^2$. By the Cauchy–Bunyakovskii inequality,

$$T(A,A,A) \geqslant \delta^3\beta_1^2\beta_2^2|W|^3 + T(f,A,A). \tag{76}$$

Let us estimate the second term on the right-hand side of (76). Using once again the Cauchy–Bunyakovskii inequality, we see that

$$\begin{aligned} T(f,A,A) &= \sum_{y,z} A(y+z,y) \sum_x E_1(y+z)f(x,y)A(x,x+z) \\ &\leqslant \Bigl(\sum_{y,z} E_1(y+z)E_2(y)\Bigr)^{1/2} \\ &\quad\times \Bigl(\sum_{x,x',y,z} E_1(y+z)A(x,x+z)A(x',x'+z)f(x,y)f(x',y)\Bigr)^{1/2}. \end{aligned} \tag{77}$$

We have $\sum_{y,z} E_1(y+z)E_2(y) = \beta_1\beta_2|W|^2$. Further,

$$\begin{aligned} \sigma &= \sum_{x,x',y,z} E_1(y+z)A(x,x+z)A(x',x'+z)f(x,y)f(x',y) \\ &= \sum_{x,x',z} A(x,x+z)A(x',x'+z) \sum_y E_1(y+z)E_2(x+z)E_2(x'+z)f(x,y)f(x',y). \end{aligned}$$

Let $\omega(x,x',y,y') = \sum_z E_1(y+z)E_1(y'+z)E_2(x+z)E_2(x'+z)$. A third use of the Cauchy–Bunyakovskii inequality gives us

$$\begin{aligned} \sigma &\leqslant \Bigl(\sum_{x,x',z} E_1(x)E_1(x')E_2(x+z)E_2(x'+z)\Bigr)^{1/2} \\ &\quad\times \Bigl(\sum_{x,x',y,y'} \omega(x,x',y,y')f(x,y)f(x',y)f(x,y')f(x',y')\Bigr)^{1/2}. \end{aligned} \tag{78}$$

Lemma 3 yields

$$\sum_{x,x',z} E_1(x)E_1(x')E_2(x+z)E_2(x'+z) \leqslant 2\beta_1^2\beta_2^2|W|^3.$$

Using the inequalities (77) and (78), we see that

$$T(f,A,A)^4 \leqslant 2\beta_1^4\beta_2^4|W|^7 \sum_{x,x',y,y'} \omega(x,x',y,y')f(x,y)f(x',y)f(x,y')f(x',y'). \quad (79)$$

By Lemma 3 applied to the case $k=4$ and for $\tau = 2^{-4}\beta_1^4\beta_2^4\eta$ and $\kappa = 2^{-12}\beta_1^{12}\beta_2^{12}\eta^3$, one can find at least $(1-\beta_1^4\beta_2^4\eta)|W|^4$ values of (x,x',y,y') such that

$$|\omega(x,x',y,y') - \beta_1^2\beta_2^2|W|\,| \leqslant \beta_1^4\beta_2^4\eta.$$

Hence,

$$\Big|\sum_{x,x',y,y'} (\omega(x,x',y,y') - \beta_1^2\beta_2^2|W|)f(x,y)f(x',y)f(x,y')f(x',y')\Big| \leqslant 3\beta_1^4\beta_2^4\eta|W|^5. \quad (80)$$

By assumption, the set A is η-uniform with respect to the rectangular norm. Using the last inequality and (79), we obtain

$$|T(f,A,A)| \leqslant 2\beta_1^2\beta_2^2\eta^{1/4}|W|^3. \quad (81)$$

By (76), we eventually have

$$T(A,A,A) \geqslant (\delta^3\beta_1^2\beta_2^2 - 2\beta_1^2\beta_2^2\eta^{1/4})|W|^3 \geqslant \delta^3\beta_1^2\beta_2^2|W|^3/2.$$

This completes the proof of Theorem 39.

Suppose that the condition of $2^{-8}\delta^{12}$-uniformity with respect to the rectangular norm of the set A in Theorem 39 fails. In the next proposition we show that in this case there is a sufficiently large set $\mathscr{Q} \subseteq \mathscr{P}$ in which the density of A is greater than δ by a certain positive quantity.

Proposition 8. *Let $\mathscr{P} = E_1 \times E_2$ and $A \subseteq \mathscr{P}$. We also assume that $\delta_{\mathscr{P}}(A) = \delta$ and A is not η-uniform with respect to the rectangular norm with $\eta > 0$. Then there are sets $F_i \subseteq E_i$, $i = 1,2$, such that $|F_i| \geqslant 2^{-8}\eta|E_i|$ and $\delta_{\mathscr{Q}}(A) \geqslant \delta + 2^{-14}\eta^2$ for $\mathscr{Q} = F_1 \times F_2$.*

When proving Proposition 8, the language of graph theory turns out to be useful. This is not surprising, because there is a close relationship between the α-uniform sets and the so-called quasi-random graphs (for details, see [93] and [94]).

Let $|E_1| = M_1$ and $|E_2| = M_2$. It can happen that $E_1 \cap E_2 \neq \varnothing$. We want to avoid this situation. Let X be a bijective image of E_1 and Y a bijective image of E_2, and let $X \cap Y = \varnothing$. We introduce the bipartite graph G associated with the set A. Let X and Y be the parts of this graph and let a vertex $x \in X$ be joined to a vertex $y \in Y$ if and only if $(x,y) \in A$. It is clear that the graph G has exactly $\delta M_1 M_2$ vertices. We denote the set of edges outgoing from a vertex $x \in X$ by $\mathscr{N}(x)$ and the set of edges incoming to a vertex $y \in Y$ by the same symbol $\mathscr{N}(y)$. We also write $d(x) = |\mathscr{N}(x)|$ for $x \in X$ and $d(y) = |\mathscr{N}(y)|$ for $y \in Y$.

To prove Proposition 8, we need several lemmas.

The first lemma asserts that 'for almost all' $x \in X$ one can assume that $d(x)$ is approximately equal to δM_2. By symmetry, the same holds for $d(y)$.

Lemma 5. *Let $\varepsilon_1, \varepsilon_2 \in (0,1)$, and suppose that at least one of the following holds*:
(a) *there are at least $\varepsilon_1 M_1$ vertices $x \in X$ such that $|d(x) - \delta M_2| > \varepsilon_2 M_2$;*
(b) *there are at least $\varepsilon_2 M_2$ vertices $y \in Y$ such that $|d(y) - \delta M_1| > \varepsilon_1 M_1$.*
Then there are sets $X' \subseteq X$ and $Y' \subseteq Y$ such that $|X'| \geqslant \min(\varepsilon_1/2, \varepsilon_2/2)M_1$, $|Y'| \geqslant \min(\varepsilon_1/2, \varepsilon_2/2)M_2$, and $\delta_{X'\times Y'}(A) \geqslant \delta + \varepsilon_1\varepsilon_2/2$.

Proof. We can assume that the first condition holds, in other words, there are at least $\varepsilon_1 M_1$ vertices $x \in X$ such that $|d(x) - \delta M_2| > \varepsilon_2 M_2$.

Suppose first that there are at least $\varepsilon_2 M_1/2$ vertices $x \in X$ such that $d(x) > (\delta+\varepsilon_2)M_2$. Let X' be the set of these vertices x and let $Y' = Y$. Then $\delta_{X'\times Y'}(A) \geqslant \delta + \varepsilon_2$, which proves the lemma.

Let X_0 be the set of vertices $x \in X$ such that $d(x) < (\delta - \varepsilon_2)M_2$. We set $X' = X \setminus X_0$ and $Y' = Y$. Let $|X'| = \kappa M_1$. Since the number of edges in G is $\delta M_1 M_2$, we see that

$$(\delta - \varepsilon_2)(1-\kappa)M_1M_2 + \kappa M_1 M_2 \geqslant \delta M_1 M_2,$$

and hence $\kappa \geqslant \varepsilon_2$. We have $|X_0| \geqslant \varepsilon_1 M_1/2$. Therefore, $\kappa \leqslant 1 - \varepsilon_1/2$ and

$$\delta_{X'\times Y'}(A) \geqslant \frac{\delta M_1 - (\delta - \varepsilon_2)M_1}{|X'|} = \delta + \varepsilon_2\left(\frac{1}{\kappa} - 1\right) \geqslant \delta + \frac{\varepsilon_1\varepsilon_2}{2}. \tag{82}$$

This completes the proof of Lemma 5.

We see that if (a) or (b) holds, then Proposition 8 is proved (for details, see below). We therefore assume that the quantities $d(x)$ and $d(y)$ are 'almost everywhere' equal to δM_2 and δM_1, respectively.

The following lemma can be regarded as a converse to Theorem 38.

Lemma 6. *Let A be a non-η-uniform set with respect to the rectangular norm. Suppose that $|d(x) - \delta M_2| \leqslant \eta M_2/56$ for any $x \in X$ except possibly for $\eta M_1/56$ vertices, and that $|d(y) - \delta M_1| \leqslant \eta M_1/56$ for any $y \in Y$ except possibly for $\eta M_2/56$ vertices. Then $\|A\|^4 \geqslant (\delta^4 + \eta/2)M_1^2M_2^2$.*

Proof. Let $f(x,y) = A(x,y) - \delta X(x)Y(y)$. Removing the parentheses in the expression for $\|A\|^4 = \|f + \delta(X\times Y)\|^4$, we obtain a sum σ with 16 terms. The leading term in σ is equal to $\delta^4 M_1^2 M_2^2$ and the term $\|f\|^4$ is not less than $\eta M_1^2 M_2^2$ by assumption. The other 14 terms in σ are of the form

$$\sum_{x,x'\in X,\ y,y'\in Y} \delta \cdot f(x',y)g(x,y')h(x',y'), \tag{83}$$

where g and h are some functions with $\|g\|_\infty \leqslant 1$ and $\|h\|_\infty \leqslant 1$. Let us estimate the sum (83). We have

$$\begin{aligned}\sigma' &= \left|\delta \sum_{x,x'\in X,\ y,y'\in Y} f(x',y)g(x,y')h(x',y')\right| \\ &\leqslant \sum_{x,x',y'}\left|\sum_y f(x',y)\right| = \sum_{x,x',y'} |d(x') - \delta M_2|. \end{aligned}\tag{84}$$

Using the assumptions of the lemma about $d(x)$, we obtain the inequality $\sigma' \leqslant \eta M_1^2 M_2^2/28$. Thus,

$$\|A\|^4 \geqslant (\delta^4 + \eta - 14\eta/28)M_1^2M_2^2 = (\delta^4 + \eta/2)M_1^2M_2^2.$$

This completes the proof of Lemma 6.

Lemma 7. *Let $\|A\|^4 \geqslant (\delta^4 + \eta/2)M_1^2M_2^2$. Suppose that $|d(x) - \delta M_2| \leqslant \eta M_2/32$ for any $x \in X$ except possibly for $\eta M_1/32$ vertices, and $|d(y) - \delta M_1| \leqslant \eta M_1/32$ for any $y \in Y$ except possibly for $\eta M_2/32$ vertices. Then there are sets $X' \subseteq X$ and $Y' \subseteq Y$ such that $|X'| \geqslant \eta M_1/32$, $|Y'| \geqslant \eta M_2/32$, and $\delta_{X'\times Y'}(A) \geqslant \delta + \eta/8$.*

Proof. Let $(x, y) \in X \times Y$ and let $e(x, y)$ be the number of edges between $\mathscr{N}(x)$ and $\mathscr{N}(y)$. One can readily see that

$$\sum_{(x,y)\in A} e(x, y) = \|A\|^4 \geqslant (\delta^4 + \eta/2)M_1^2M_2^2.$$

Let X_0 be the set of vertices $x \in X$ for which $|d(x) - \delta M_2| \leqslant \eta M_2/32$ and let Y_0 be the set of vertices $y \in Y$ such that $|d(y) - \delta M_1| \leqslant \eta M_1/32$. By assumption, $|X_0^c| = |X \setminus X_0| \leqslant \eta M_1/32$ and $|Y_0^c| = |Y \setminus Y_0| \leqslant \eta M_2/32$. Therefore, the number of edges beginning in X_0^c and ending in Y_0^c does not exceed $\eta M_1M_2/16$. Hence,

$$\sum_{(x,y)\in A,\ x\in X_0, y\in Y_0} e(x, y) \geqslant (\delta^4 + \eta/4)M_1^2M_2^2. \tag{85}$$

It follows from (85) that there exist an $x \in X_0$ and a $y \in Y_0$ for which $e(x, y) \geqslant (\delta^3 + \eta/4\delta)M_1M_2$. We set $X' = \mathscr{N}(x)$ and $Y' = \mathscr{N}(y)$. Since $x \in X_0$ and $y \in Y_0$, it follows that $|X'| \leqslant (\delta + \eta/32)M_1$ and $|Y'| \leqslant (\delta + \eta/32)M_2$. Therefore,

$$\delta_{X'\times Y'}(A) \geqslant \frac{\delta^3 + \eta/(4\delta)}{(\delta + \eta/32)^2} \geqslant \delta + \eta/8.$$

Again using the fact that $x \in X_0$ and $y \in Y_0$, we see that $|X'| \geqslant \delta M_1/2 > \eta M_1/32$ and $|Y'| \geqslant \delta M_2/2 > \eta M_2/32$. This proves Lemma 7.

Proof of Proposition 8. Suppose that there are at least $\eta M_1/56$ vertices $x \in X$ for which $|d(x) - \delta M_2| > \eta M_2/56$. Then by Lemma 5 there are two sets, F_1 and F_2, such that $|F_i| \geqslant 2^{-8}\eta M_i$, $i = 1, 2$, and $\delta_{\mathscr{Q}}(A) \geqslant \delta + 2^{-14}\eta^2$ for $\mathscr{Q} = F_1 \times F_2$. The proposition is proved for this case. The situation in which there are at least $\eta M_2/56$ vertices $y \in Y$ for which $|d(y) - \delta M_1| > \eta M_1/56$ can be treated similarly.

If $|d(x) - \delta M_2| \leqslant \eta M_2/56$ for any $x \in X$ except possibly for $\eta M_1/56$ vertices and a similar condition holds for $d(y)$, then $\|A\|^4 \geqslant \delta^4 + \eta/2$ by Lemma 6. We now apply Lemma 7 and find sets F_1 and F_2 such that $|F_i| \geqslant \eta M_i/32$ and $\delta_{\mathscr{Q}}(A) = \delta_{F_1\times F_2}(A) \geqslant \delta + \eta/8$. This completes the proof of Proposition 8.

For convenience let us combine Theorem 39 and Proposition 8.

Proposition 9 (first stage of the proof of Theorem 37). *Let W be a subspace of $\mathbb{Z}_2^n$, let $E_i \subseteq W$, $|E_i| = \beta_i|W|$, $i = 1, 2$, and let $\mathscr{P} = E_1 \times E_2$. Suppose that $A \subseteq \mathscr{P}$, $\delta_{\mathscr{P}}(A) = \delta$, and A contains no corners. Further, suppose that*

$$|W| > 2\delta^{-3}\beta_1^{-2}\beta_2^{-2} \tag{86}$$

and that E_i is a $2^{-36}\beta_1^{12}\beta_2^{12}\delta^{36}$-uniform subset of W for $i = 1, 2$.

Then there are sets $F_i \subseteq E_i$ *such that* $|F_i| \geqslant 2^{-16}\delta^{12}|E_i|$ *and* $\delta_{F_1\times F_2}(A) \geqslant \delta + 2^{-30}\delta^{24}$.

Proof. Suppose that A is η-uniform with respect to the rectangular norm with $\eta = 2^{-8}\delta^{12}$. By Theorem 39, the set A contains at least $\delta^3\beta_1^2\beta_2^2|W|^3$ triples of the form $\{(x,y),(x+d,y),(x,y+d)\}$. The number of triples with $d=0$ does not exceed $|W|^2$. By (86), A contains a corner. Hence, A is not η-uniform with respect to the rectangular norm with $\eta = 2^{-8}\delta^{12}$. By Proposition 8, there are sets $F_i \subseteq E_i$ such that $|F_i| \geqslant 2^{-16}\delta^{12}|E_i|$ and $\delta_{F_1\times F_2}(A) \geqslant \delta + 2^{-30}\delta^{24}$. This completes the proof of Proposition 9.

Proposition 10 (second stage of the proof of Theorem 37). *Let* $\delta,\tau,\sigma \in (0,1)$, *let* W *be a subspace of* $\mathbb{Z}_2^n$, *let* $F_i \subseteq W$, $|F_i| = \beta_i|W|$, $i=1,2$, *and let* $\mathscr{Q} = F_1\times F_2$. *Assume that* $A \subseteq \mathscr{Q}$, $\delta_{\mathscr{Q}}(A) = \delta + \tau$, *and*

$$|W| > \exp(16\sigma^{-2}(\beta_1\beta_2)^{-1}\tau^{-1}). \tag{87}$$

Then there exist a subspace $W' \subseteq W$ *with* $\dim W' \geqslant \dim W - 16\sigma^{-2}(\beta_1\beta_2)^{-1}\tau^{-1}$ *and points* $t_1,t_2 \in W$ *such that the sets* $E_1' = (F_1 - t_1)\cap W'$, $E_2' = (F_2-t_2)\cap W'$, *and* $\mathscr{P}' = E_1'\times E_2'$, *have the following properties*:

(a) $|\mathscr{P}'| \geqslant \beta_1\beta_2\tau|W'|^2/2$;

(b) E_1' *and* E_2' *are* σ*-uniform subsets of* W';

(c) $\delta_{\mathscr{P}'}(A - (t_1,t_2)) \geqslant \delta + \tau/4$.

Proof. The proof of the proposition is an algorithm. We sketch a description. At the jth step of the algorithm the set $W\times W$ is partitioned into *cells*

$$C^{(i)} = (W^{(i)} + t_1^{(i)})\times(W^{(i)} + t_2^{(i)}), \qquad W\times W = \bigsqcup_{i\in\mathscr{I}_j} C^{(i)}, \tag{88}$$

where $t_1^{(i)}, t_2^{(i)} \in W$, each of the sets $W^{(i)}$ is a subspace of W with dimension at least $\dim W - j$, and $\mathscr{I}_j$ is some set of indices. Upon passage to the $(j+1)$st step of the algorithm, some cells $C^{(i)}$, $i\in\mathscr{I}_j$, are unchanged. For the other indices $i\in\mathscr{I}_j$ one chooses some subspaces $H^{(i)} \subseteq W^{(i)}$ of codimension 1. Since $W^{(i)} = H^{(i)}\sqcup(H^{(i)})^{\perp}$, it follows that the cells $C^{(i)}$ are decomposed into four subcells $\widetilde{C}^{(i_k)}$, $k=1,2,3,4$, where

$$\widetilde{C}^{(i_1)} = (H^{(i)} + t_1^{(i)})\times(H^{(i)} + t_2^{(i)}), \qquad \widetilde{C}^{(i_2)} = ((H^{(i)})^{\perp} + t_1^{(i)})\times(H^{(i)} + t_2^{(i)}),$$
$$\widetilde{C}^{(i_3)} = (H^{(i)} + t_1^{(i)})\times((H^{(i)})^{\perp} + t_2^{(i)}), \qquad \widetilde{C}^{(i_4)} = ((H^{(i)})^{\perp} + t_1^{(i)})\times((H^{(i)})^{\perp} + t_2^{(i)}).$$

We now proceed directly to the proof. Let $\beta = \beta_1\beta_2$. At the first step of the algorithm we set $t_1^{(1)} = t_2^{(1)} = 0$, $\mathscr{I}_1 = \{1\}$, and $C^{(1)} = W\times W$. It is clear that the formulae (88) hold in this case. Suppose that we have carried out j steps of the algorithm and have constructed some cells $C^{(i)}$ and a set $\mathscr{I}_j$ that satisfy the formulae (88). Let $D_1^{(i)} = F_1\cap(W^{(i)} + t_1^{(i)})$, $D_2^{(i)} = F_2\cap(W^{(i)} + t_2^{(i)})$, and

$$\beta_1^{(i)} = \frac{|D_1^{(i)}|}{|W^{(i)} + t_1^{(i)}|}, \qquad \beta_2^{(i)} = \frac{|D_2^{(i)}|}{|W^{(i)} + t_2^{(i)}|}. \tag{89}$$

We also assume that $\beta^{(i)} = \beta_1^{(i)}\beta_2^{(i)}$. Clearly,

$$\sum_{i\in\mathscr{I}_j} |C^{(i)}|\beta^{(i)} = \beta|W|^2. \qquad (90)$$

We say that a cell $C^{(i)}$ is *meagre* if $\beta^{(i)} < \beta\tau/2$. Let us partition all the non-meagre cells into two classes. In the first class we put all the cells for which $D_1^{(i)} - t_1^{(i)}$ and $D_2^{(i)} - t_2^{(i)}$ are σ-uniform subsets of $W^{(i)}$. These cells are said to be *uniform*. The remaining non-meagre cells are said to be *non-uniform*.

Thus, for $i \in \mathscr{I}_j$ the cell $C^{(i)}$ can turn out to be meagre, uniform, or non-uniform. Let $\mathscr{E}_j$, $\mathscr{U}_j$, and $\mathscr{N}_j$ denote the corresponding subsets of $\mathscr{I}_j$. If

$$\sum_{i\in\mathscr{N}_j} |C^{(i)}| < \tau\beta|W|^2/4, \qquad (91)$$

then we terminate the algorithm at the jth step. Otherwise we partition the cells for all $i \in \mathscr{N}_j$ and leave unchanged the cells for $i \in \mathscr{E}_j \sqcup \mathscr{U}_j$. We note that for this algorithm we have $\dim W^{(i)} = n - j$ for any $i \in \mathscr{N}_j$. By (91), there are at least $\frac{\tau}{4}2^{2j}$ values $i \in \mathscr{N}_j$. Without loss of generality we can assume that there are at least $\frac{\tau}{8}2^{2j}$ values of i such that $D_1^{(i)} - t_1^{(i)}$ are not σ-uniform subsets of $W^{(i)}$. By Lemma 4, for any such i there is a subspace $H^{(i)} \subseteq W^{(i)}$ of codimension 1 such that

$$\frac{1}{2}(\delta^2_{H^{(i)}}(D_1^{(i)} - t_1^{(i)}) + \delta^2_{(H^{(i)})^\perp}(D_1^{(i)} - t_1^{(i)})) \geqslant \beta_1^{(i)2} + \sigma^2. \qquad (92)$$

We decompose the cell $C^{(i)}$ into four subcells $C^{(i_k)}$, $k = 1, 2, 3, 4$, as was done above. Let $\beta_1^{(i_k)} = \delta_{C^{(i_k)}}(F_1)$ and $\beta_2^{(i_k)} = \delta_{C^{(i_k)}}(F_2)$, $k = 1, 2, 3, 4$. It follows from (92) that

$$\frac{1}{4}(\beta_1^{(i_1)2} + \beta_1^{(i_2)2} + \beta_1^{(i_3)2} + \beta_1^{(i_4)2}) \geqslant \beta_1^{(i)2} + \sigma^2. \qquad (93)$$

We prove that the algorithm terminates after at most $16\sigma^{-2}\beta^{-1}\tau^{-1}$ steps. Let

$$\mathrm{ind}(\mathscr{I}_j) := \frac{1}{2|W|^2}\sum_{i\in\mathscr{I}_j} |C^{(i)}|(\beta_1^{(i)2} + \beta_2^{(i)2}). \qquad (94)$$

Using the inequality (93) and the estimate

$$\sum_{i\in\mathscr{N}_j} |C^{(i)}| \geqslant \tau\beta|W|^2/4, \qquad (95)$$

we see that $\mathrm{ind}(\mathscr{I}_{j+1}) \geqslant \mathrm{ind}(\mathscr{I}_j) + \frac{1}{16}\tau\beta\sigma^2$. On the other hand, $\mathrm{ind}(\mathscr{I}_j) \leqslant 1$ for all j. Hence, the algorithm must terminate after $16\sigma^{-2}\beta^{-1}\tau^{-1}$ steps.

Suppose that the algorithm has terminated at the Kth step, $K \leqslant 16\sigma^{-2}\beta^{-1}\tau^{-1}$. Since every uniform cell of the partition (88) is not meagre, the cell satisfies the condition (a) of Proposition 10. It is also clear that every uniform cell satisfies

the condition (b). We show that there is a uniform cell for which the property (c) holds as well. We have

$$\sum_{i\in\mathscr{N}_j}|C^{(i)}|<\tau\beta|W|^2/4. \tag{96}$$

Let $\delta^{(i)}=|A\cap C^{(i)}|/|C^{(i)}\cap(F_1\times F_2)|$. It is clear that $\delta_{C^{(i)}}(A)=\delta^{(i)}\beta^{(i)}$. Since $\beta^{(i)}<\beta\tau/2$ for any $i\in\mathscr{E}_K$, it follows that

$$\sum_{i\in\mathscr{E}_K}|C^{(i)}|\delta^{(i)}\beta^{(i)}<\beta\tau|W|^2/2. \tag{97}$$

Applying (90) and the equality $\delta_{\mathscr{Q}}(A)=\delta+\tau$, we see that

$$\sum_{i\in\mathscr{U}_K\sqcup\mathscr{N}_K}|C^{(i)}|\delta^{(i)}\beta^{(i)}>\beta(\delta+\tau)|W|^2-\beta\tau|W|^2/2\geqslant\beta(\delta+\tau/2)|W|^2. \tag{98}$$

Suppose that $\delta^{(i)}<\delta+\tau/4$ for all $i\in\mathscr{U}_K$. By (90) and (96), we arrive at the contradictory inequality

$$\begin{aligned}\beta(\delta+\tau/2)|W|^2&\leqslant\sum_{i\in\mathscr{U}_K}|C^{(i)}|\delta^{(i)}\beta^{(i)}+\sum_{i\in\mathscr{N}_K}|C^{(i)}|\delta^{(i)}\beta^{(i)}\\&<(\delta+\tau/4)\sum_{i\in\mathscr{S}_K}|C^{(i)}|\beta^{(i)}+\tau\beta|W|^2/4=\beta(\delta+\tau/2)|W|^2.\end{aligned} \tag{99}$$

Therefore, there is an $i\in\mathscr{U}_K$ such that $\delta^{(i)}\geqslant\delta+\tau/4$. Let $W'=W^{(i)}$, $t_1=t_1^{(i)}$, and $t_2=t_2^{(i)}$. The subspace W' and the vectors t_1 and t_2 satisfy all the conditions in Proposition 10. This proves Proposition 10.

Before passing to the proof of the main result of this section, we combine the first and second stages of the proof of Theorem 37 into a single proposition.

Proposition 11. *Let W be a subspace of $\mathbb{Z}_2^n$, let $E_i\subseteq W$, $|E_i|=\beta_i|W|$, $i=1,2$, let $\beta=\beta_1\beta_2$, and let $\mathscr{P}=E_1\times E_2$. Assume that $A\subseteq\mathscr{P}$, $\delta_{\mathscr{Q}}(A)=\delta$, the sets E_1 and E_2 are $2^{-36}\beta^{12}\delta^{36}$-uniform, and*

$$|W|>\exp(2^{1681}\delta^{-1272}\beta^{-25}). \tag{100}$$

Assume also that A contains no corners. Then there exist a subspace $W'\subseteq W$ and sets E_1' and E_2' such that:

(a) *the numbers β_1', β_2', and β' determined by the equalities $E_1'=\beta_1'|W|$, $E_2'=\beta_2'|W|$, and $\beta'=\beta_1'\beta_2'$ satisfy the inequality $\beta'\geqslant 2^{-63}\delta^{48}\beta$;*
(b) *E_1' and E_2' are $2^{-36}\beta'^{12}\delta^{36}$-uniform subsets of W';*
(c) *$\dim W'\geqslant\dim W-2^{1681}\delta^{-1272}\beta^{-25}$;*
(d) *$\delta_{\mathscr{P}'}(A-t)\geqslant\delta+2^{-32}\delta^{24}$ for some $t\in W\times W$, where $\mathscr{P}'=E_1'\times E_2'$.*

To prove Proposition 11, it suffices to apply Proposition 9 and then Proposition 10. Here the parameters τ and σ in Proposition 10 are equal to $2^{-30}\delta^{24}$ and $2^{-36}(\beta')^{12}\delta^{36}$, respectively.

Let us proceed directly to the proof of Theorem 37. This proof is an algorithm. At the zeroth step of the algorithm we set $\mathscr{P}_0=\mathbb{Z}_2^n\times\mathbb{Z}_2^n$ and assume that the set $A\subseteq\mathscr{P}_0$ contains no corners.

Suppose that i steps of the algorithm have been carried out, $i \geqslant 0$. Then

(i) the set $\mathscr{P}_i = E_i^{(1)} \times E_i^{(2)}$ is contained in $W_i \times W_i$, where W_i is a subspace of $\mathbb{Z}_2^n$ of codimension d_i;

(ii) $|\mathscr{P}_i| = \beta_i |W_i|^2$ and the sets $E_i^{(1)}$ and $E_i^{(2)}$ are $2^{-36}\beta_i^{12}\delta^{36}$-uniform subsets of W_i;

(iii) $\delta_{\mathscr{P}_i}(A - t) \geqslant \delta + 2^{-32}\delta^{24} i$ for some $t \in \mathbb{Z}_2^n \times \mathbb{Z}_2^n$.

If

$$|W_i| > \exp(2^{1681}\delta^{-1272}\beta_i^{-25}), \tag{101}$$

then we apply Proposition 11 and carry out the $(i+1)$st step of the algorithm. Since

$$d_{i+1} \leqslant d_i + 2^{1681}\delta^{-1272}\beta_i^{-25}$$

and

$$\beta_{i+1} \geqslant 2^{-63}\delta^{48}\beta_i,$$

it follows that $\beta_i \geqslant (\delta/2)^{63i}$ and $d_i \leqslant \delta^{-C_1 i}$, where C_1 is an absolute constant.

Using the condition (d) of Proposition 11, we conclude (as in the proof of Roth's theorem) that the number of steps of the algorithm does not exceed $K = C_2\delta^{-23}$, where C_2 stands for another absolute constant.

Suppose that

$$N \gg \exp(\delta^{-C_3\delta^{-23}}), \tag{102}$$

where C_3 stands for an absolute constant. One can easily see that the inequality (102) ensures the validity of the condition (101) at all steps of the algorithm, and in particular at the Kth and last step. Hence, we can carry out the $(K+1)$st step of the algorithm. The contradiction obtained shows that we have the estimate $N \ll \exp(\delta^{-C_3\delta^{-23}})$ (rather than the inequality (102)), which implies that $\delta \ll (\log\log N)^{-1/24}$. This proves Theorem 37.

In [31] and [81] an application of Theorem 36 to the theory of dynamical systems was obtained. Before formulating our theorem, we give several definitions.

Let X be a set equipped with a σ-algebra $\mathscr{B}$ of measurable sets, and let μ be a finite measure on $\mathscr{B}$. Without loss of generality, we assume that $\mu(X) = 1$.

Definition 11 (Hausdorff measure). We consider the measure $H_h(\,\cdot\,)$ on X defined as follows:

$$H_h(E) = \lim_{\delta \to 0} H_h^\delta(E), \tag{103}$$

where $h(t)$ is a non-negative ($h(0) = 0$) continuous increasing function and $H_h^\delta(E) = \inf\{\sum h(\delta_j)\}$, where inf is taken over at most countable coverings of E by open sets $\{B_j\}$ with $\operatorname{diam}(B_j) = \delta_j < \delta$.

If $h(t) = t^\alpha$, then we obtain the usual Hausdorff measure.

The outer measure $H_h(\,\cdot\,)$ is σ-additive on the σ-algebra of Carathéodory measurable sets (for more details see, for instance, [95]). As is well known, this σ-algebra contains all Borel sets.

We say that the measures μ and H_h *are compatible* if any μ-measurable set is H_h-measurable (in the sense of Carathéodory measurability).

Let S and R be two *commuting* maps of the space X that preserve the measure μ.

Definition 12. The function

$$C_{S,R}(x) = C^h_{S,R}(x) := \liminf_{n\to\infty} \{L^{-1}(n) \cdot \max\{h(d(S^n x, x)), h(d(R^n x, x))\}\},$$

where $L^{-1}(n) = 1/L(n)$, is called the *constant of simultaneous (or multiple) recurrence* of a point x.

Theorem 40 [31], [81]. *Let X be a metric space with $H_h(X) < \infty$ and let S and R be commuting self-maps of X that preserve the measure μ. Suppose that the measures μ and H_h are compatible. Then $C_{S,R}(x)$ is an integrable function (with respect to the measure μ) and*

$$\int_A C_{S,R}(x)\, d\mu \leqslant H_h(A) \tag{104}$$

for any μ-measurable set A. If $H_h(A) = 0$, then $\int_A C_{S,R}(x)\, d\mu = 0$ without the assumption that the measures μ and H_h are compatible.

For *lower* bounds of the function $C_{S,R}(x)$, see the paper [87].

§ 8. Arithmetic progressions formed of primes

The conjecture that the set of primes contains arithmetic progressions of arbitrary length has a history of more than two hundred years. The first remarks concerning progressions in primes can be found in the 1770 correspondence between Lagrange and Waring (see [96]). However, the first results in this area were obtained only in 1938, when N. G. Chudakov, using I. M. Vinogradov's method of trigonometric sums, proved that the set of primes contains arithmetic progressions of length three [97] (see also [98], [99]). As far as progressions of length exceeding three are concerned, the problem remained open until very recently.

Computers were used to seek progressions in the primes. For example, using a computer, P. A. Pritchard, A. Moran, and A. Thyssen [100] discovered the following arithmetic progression of length 22 formed of primes:

$$11410337850553 + 4609098694200k,$$

where $k = 0, 1, \ldots, 21$. This record remained unbroken for almost ten years. In 2004 M. Frind, P. Jobling, and P. Underwood found a progression of primes of length 23 (see [101]). This progression began with 56211383760397 and had the difference 44546738095860.

In 2004 Green and Tao proved the following result about progressions in the primes (see [24]).

Theorem 41 (Green, Tao). *The set of primes contains an arithmetic progression of length k for any integer $k \geqslant 3$.*

In this section we sketch the proof of the Green–Tao theorem.

As is well known, the density of the set of primes $\mathscr{P}$ in the segment $[N]$ is $\pi(N)/N \sim 1/\ln N = o(1)$, $N \to \infty$ (see, for instance, [35]). Therefore, we cannot apply Szemerédi's theorem (Theorem 6) to $\mathscr{P}$. One of the main ideas of

Green and Tao was to find a generalization of Szemerédi's theorem to the so-called *pseudorandom* sets (we give a precise definition below), which can have zero density. Generally speaking, there are many parallels between the Green–Tao approach and the ergodic method. For example, pseudorandom sets are analogues of weakly mixing dynamical systems. These parallels are treated in more detail in the survey [102]. In the second part of their proof Green and Tao apply recent results of D. A. Goldston and C. Y. Yildirim (see [103]–[105]) and show that the set of primes has the desired pseudorandom properties.

Let us begin with a more detailed discussion of the first part of the Green–Tao proof, in which they obtain a generalization of Szemerédi's theorem for pseudorandom sets. Their proof makes essential use of Szemerédi's theorem. More precisely, they show that the generalized theorem they proved is a consequence of the usual Szemerédi theorem.

It is more convenient for our purposes to formulate Szemerédi's theorem for the group $\mathbb{Z}_N$. Let N be an arbitrary prime. We say that a quantity is of the form $o(1)$ if it tends to zero as N tends to infinity, and of the form $O(1)$ if it remains bounded as N tends to infinity. As above, we use the notation $\nu_{\mathrm{const}}\colon \mathbb{Z}_N \to \mathbb{R}^+$ for the constant function identically equal to one, that is, $\nu_{\mathrm{const}} \equiv 1$.

Theorem 42 (Szemerédi's theorem). *Let $k \geqslant 3$ be an integer and let $\delta > 0$. Assume that N is a sufficiently large prime and $f\colon \mathbb{Z}_N \to \mathbb{R}^+$ is a non-negative function such that*

$$0 \leqslant f(x) \leqslant \nu_{\mathrm{const}}(x) \quad \text{for any} \quad x \in \mathbb{Z}_N \tag{105}$$

and

$$\frac{1}{N}\sum_{x\in\mathbb{Z}_N} f(x) \geqslant \delta. \tag{106}$$

Then

$$\frac{1}{N^2}\sum_{x,r\in\mathbb{Z}_N} f(x)f(x+r)\cdots f(x+(k-1)r) \geqslant c'(k,\delta) \tag{107}$$

for some constant $c'(k,\delta) > 0$ depending only on k and δ.

There are two differences between Theorem 18 in § 4 and Theorem 42. The first theorem involves $\mathbb{Z}$, and the second involves $\mathbb{Z}_N$. Moreover, the formulation of the first theorem involves functions, while that of the second involves sets. In fact, the theorems are equivalent, and the constants $c(k,\delta)$ and $c'(k,\delta)$ can be expressed explicitly in terms of each other. The first difference is actually inessential (see, for instance, the proof of Theorem 20). The second difference is not important either. Indeed, for any function f satisfying the inequalities (105) and (106) we can consider the set $A = \{x : f(x) \geqslant \delta/2\}$. Then $|A| \geqslant \delta/2 \cdot N$. Hence, by Theorem 18, the characteristic function of the set A satisfies the inequality (31) with the constant $c(k,\delta/2)$. Then the inequality (107) holds for the function f with the constant $c'(k,\delta) = (\delta/2)^k \cdot c(k,\delta/2)$.

Let us sketch the proof of Theorem 41. In their paper Green and Tao replace the stringent condition (105) imposed on f by a weaker condition. We call a function $\nu(x)$ a *measure* if $1/N \cdot \sum_{x\in\mathbb{Z}_N} \nu(x) = 1 + o(1)$. It is clear that $\nu_{\mathrm{const}}(x)$ is

a measure. Another example of a measure is the von Mangoldt function, which is concentrated on the powers of the primes,

$$\Lambda(x) = \begin{cases} \log p & \text{if } x = p^m,\ p \in \mathscr{P},\ m \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

We have $1/N \cdot \sum_{x \in \mathbb{Z}_N} \Lambda(x) = 1 + o(1)$ (see, for instance, [106]), and hence $\Lambda(x)$ is a measure. It is proved in a generalization of Theorem 42 that the function $\nu_{\mathrm{const}}(x)$ (identically equal to one) in the formula (105) can be replaced by some measure $\nu(x)$ satisfying two conditions: the condition of linear forms and the correlation condition (we give precise definitions below). Here the measure $\nu(x)$ can grow as x tends to infinity. Then the generalization of Theorem 42 is applied to a specific measure $\nu_0(x)$ connected with the primes. The recent papers [103]–[105] of Goldston and Yildirim are used in an essential way to verify that the measure $\nu_0(x)$ satisfies the condition of linear forms and the correlation condition.

Definition 13 (condition of linear forms). Let $\nu\colon \mathbb{Z}_N \to \mathbb{R}^+$ be a measure, and let m_0, t_0, and L_0 be positive integers, $L_0 < N$. The measure ν is said to be satisfy the (m_0, t_0, L_0)-condition of linear forms if the following property holds. Let $m \leqslant m_0$ and $t \leqslant t_0$, and let $(L_{ij})_{1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant t}$ be rational numbers whose numerators and denominators do not exceed L_0 in absolute value. Also, let $b_i \in \mathbb{Z}_N$, $1 \leqslant i \leqslant m$, and let $\psi_i\colon \mathbb{Z}_N^t \to \mathbb{Z}_N$ be the linear forms $\psi_i(\mathbf{x}) = \sum_{j=1}^t L_{ij} x_j + b_i$, where $\mathbf{x} = (x_1, \dots, x_t) \in \mathbb{Z}_N^t$ and the rational numbers L_{ij} can be interpreted as elements of $\mathbb{Z}_N$ in the standard way. Moreover, suppose that the rows $(L_{ij})_{1 \leqslant j \leqslant t} \in \mathbb{Q}^t$ are non-zero for all $i = 1, \dots, m$ and that no row can be obtained from another row by multiplying the latter by a rational number. Finally, let

$$\frac{1}{N^t} \sum_{\mathbf{x} \in \mathbb{Z}_N^t} \nu(\psi_1(\mathbf{x})) \cdots \nu(\psi_m(\mathbf{x})) = 1 + o_{L_0, m_0, t_0}(1). \tag{108}$$

Then we say that the measure ν satisfies the (m_0, t_0, L_0)-*condition of linear forms*.

Definition 14 (correlation condition). Let $\nu\colon \mathbb{Z}_N \to \mathbb{R}^+$ be a measure, let m_0 be a positive integer, and for $m = 1, \dots, m_0$ suppose that there is a function $\tau_m\colon \mathbb{Z}_N \to \mathbb{R}^+$ such that

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} \tau_m^q(x) = O_{m,q}(1) \tag{109}$$

for any $q \geqslant 1$ and

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} \nu(x + h_1)\nu(x + h_2) \cdots \nu(x + h_m) \leqslant \sum_{1 \leqslant i < j \leqslant m} \tau_m(h_i - h_j) \tag{110}$$

for any $h_1, \dots, h_m \in \mathbb{Z}_N$. Then the measure ν is said to satisfy the m_0-*correlation condition*.

Definition 15 (pseudorandom measure). Let $\nu\colon \mathbb{Z}_N \to \mathbb{R}^+$ be a measure. The measure ν is said to be k-*pseudorandom* if it satisfies the $(k2^{k-1}, 3k-4, k)$-condition of linear forms and the 2^{k-1}-correlation condition.

It is clear that the measure $\nu_{\mathrm{const}}(x)$ is k-pseudorandom for any positive integer k. It turns out that k-pseudorandom measures are measures close to $\nu_{\mathrm{const}}(x)$ in the sense of the Gowers uniform norms (see the definition in § 5).

Proposition 12. *Let $\nu(x)$ be a k-pseudorandom measure. Then*

$$\|\nu - \nu_{\mathrm{const}}\|_{U^d} = \|\nu - 1\| = o(1) \tag{111}$$

for any $1 \leqslant d \leqslant k-1$.

Proof. By the monotonicity inequality (45) for the Gowers norms, it suffices to prove (111) for $d = k-1$. In other words, it suffices to show that

$$\sigma := \sum_{x\in\mathbb{Z}_N,\ h\in\mathbb{Z}_N^{k-1}} \prod_{\omega\in\{0,1\}^{k-1}} (\nu(x+\omega\cdot h) - 1) = o(N^k). \tag{112}$$

The left-hand side of (112) is equal to

$$\sum_{A\subseteq\{0,1\}^{k-1}} (-1)^{|A|} \sum_{x\in\mathbb{Z}_N,\ h\in\mathbb{Z}_N^{k-1}} \prod_{\omega\in A} \nu(x+\omega\cdot h). \tag{113}$$

For any fixed set A the expression

$$\sum_{x\in\mathbb{Z}_N,\ h\in\mathbb{Z}_N^{k-1}} \prod_{\omega\in A} \nu(x+\omega\cdot h) \tag{114}$$

can be represented in the form

$$\sum_{\mathbf{z}\in\mathbb{Z}_N^k} \nu(\psi_1(\mathbf{z}))\cdots\nu(\psi_{|A|}(\mathbf{z})), \tag{115}$$

where $\psi_1,\dots,\psi_{|A|}$ are non-zero linear forms, $\psi_\omega(\mathbf{z}) := x + \omega\cdot h$, $\omega\in A$, and $\mathbf{z} = (x, h_1,\dots,h_{k-1})$. It is clear that none of these forms can be obtained from another by multiplying by a rational number. Since the measure ν is k-pseudorandom, it satisfies the $(k2^{k-1}, 3k-4, k)$-condition of linear forms and, in particular, the $(2^{k-1}, k, 1)$-condition of linear forms. Hence, every expression of the form (114) is equal to $N^k + o(N^k)$. Therefore,

$$\sigma = N^k \sum_{A\subseteq\{0,1\}^{k-1}} (-1)^{|A|}(1+o(1)) = o(N^k),$$

as was to be proved.

We state the Green–Tao generalization of Szemerédi's theorem.

Theorem 43 (Szemerédi's theorem for pseudorandom measures). *Let k be a positive integer with $k \geqslant 3$, let $0 < \delta \leqslant 1$, let $\nu\colon \mathbb{Z}_N \to \mathbb{R}^+$ be a k-pseudorandom measure, and let $f\colon \mathbb{Z}_N \to \mathbb{R}^+$ be a function such that*

$$0 \leqslant f(x) \leqslant \nu(x) \quad \textit{for any} \quad x\in\mathbb{Z}_N \tag{116}$$

and

$$\frac{1}{N}\sum_{x\in\mathbb{Z}_N} f(x) \geqslant \delta. \tag{117}$$

Then

$$\frac{1}{N^2}\sum_{x,r\in\mathbb{Z}_N} f(x)f(x+r)\cdots f(x+(k-1)r) \geqslant c'(k,\delta) \tag{118}$$

for the constant $c'(k,\delta) > 0$ in Theorem 42.

To derive Theorem 41 from Theorem 43, one must produce a pseudorandom measure $\nu(x)$ and a function $f(x)$ that satisfy the conditions (116) and (117).

For the function $f(x)$ one could try to take the von Mangoldt function $\Lambda(x)$. As was noted above, we have $1/N\cdot\sum_{x\in\mathbb{Z}_N}\Lambda(x) = 1 + o(1)$, and hence $\Lambda(x)$ satisfies (117). One must now find a k-pseudorandom measure $\nu(x)$ such that $\nu(x) \geqslant c(k)\Lambda(x)$ for some positive constant $c(k)$. If a measure of this kind were constructed, then by Theorem 43 one could find an arithmetic progression of length k in the set of numbers of the form $\{p, p^2, p^3, \dots \mid p \in \mathscr{P}\}$. This does not yet prove Theorem 41, of course, because the terms of an arithmetic progression found in this way might not be prime numbers. In fact, one can readily show that the higher powers of primes, p^2, p^3, ..., give a contribution of the form $o(1)$ to the sum (118). Therefore, if a k-pseudorandom measure $\nu(x)$ were found such that $\nu(x) \geqslant c(k)\Lambda(x)$, $c(k) > 0$, then Theorem 41 would be proved.

Unfortunately, there is no k-pseudorandom measure with the desired property. We sketch a proof of this fact here. One can easily show that every pseudorandom measure is uniformly distributed with respect to all q residue classes $a \pmod q$ for any fixed integer $q > 1$. Let $\varphi(x)$ be the Euler function. As is known, the quantity $\varphi(x)/x$ can be made as small as desired (see, for instance, [36], Ch. II, (3) in § 4 and Question 9e), and therefore there is a q such that $\varphi(q)/q < 1/c(k)$. On the other hand, if $(a,q) > 1$ for a residue class a, then the class a contains no primes. Since the contribution of the higher powers $(p^2, p^3, \dots)$ of the primes to the mean value of the von Mangoldt function is equal to $o(1)$, there is a residue class a for which $1/N\cdot\sum_{x\in\mathbb{Z}_N,\ x\equiv a \pmod q}\Lambda(x) \geqslant 1/\varphi(q) + o(1)$. Since $\varphi(q)/q < 1/c(k)$ and $1/N\cdot\sum_{x\in\mathbb{Z}_N, x\equiv a \pmod q}\nu(x) = 1/q + o(1)$, it follows that the inequality $\nu(x) \geqslant c(k)\Lambda(x)$ cannot hold for arbitrary x.

Thus, we see that the 'non-uniformity' in the distribution of the primes prevents the existence of a pseudorandom measure with the desired properties. To avoid the above difficulties, Green and Tao proposed an approach they called the 'W-trick'.

Let $w(N)$ be a function slowly growing to infinity (the order of growth of $w(N)$ will be chosen below) and let $W = \prod_{p\in[w(N)]} p$ be the product of the primes in the segment $[w(N)]$. We define the modified von Mangoldt function $\widetilde{\Lambda}\colon \mathbb{Z}^+ \to \mathbb{R}^+$ by the formula

$$\widetilde{\Lambda}(x) = \begin{cases} \dfrac{\varphi(W)}{W}\log(Wx+1) & \text{if } Wx+1 \text{ is a prime,} \\ 0 & \text{otherwise.} \end{cases}$$

If $w(N) \ll \log\log N$, then we get that $\sum_{x\in\mathbb{Z}_N}\widetilde{\Lambda}(x) = N(1+o(1))$ by using the Dirichlet theorem on arithmetic progressions. Thus, the inequality (117) holds

for $\widetilde{\Lambda}(x)$ as well as for $\Lambda(x)$. In what follows we assume that $w(N) \ll \log\log N$. The main difference between the function $\widetilde{\Lambda}(x)$ and the von Mangoldt function is that $\widetilde{\Lambda}(x)$ admits a k-pseudorandom measure $\nu(x)$ such that $\nu(x) \geqslant c(k)\Lambda(x)$, $c(k) > 0$. More precisely, the following proposition holds.

Proposition 13. *Let $\varepsilon_k = 1/(2^k(k+4)!)$ and let N be a sufficiently large prime. Then there is a k-pseudorandom measure $\nu\colon \mathbb{Z}_N \to \mathbb{R}^+$ with $\nu(x) \geqslant k^{-1}2^{-k-5}\widetilde{\Lambda}(x)$ for all $x \in [\varepsilon_k N, 2\varepsilon_k N]$.*

We derive Theorem 41 from Proposition 13. Let $f\colon \mathbb{Z}_N \to \mathbb{R}^+$ be the function with $f(x) = k^{-1}2^{-k-5}\widetilde{\Lambda}(x)$ for any $x \in [\varepsilon_k N, 2\varepsilon_k N]$ and $f(x) = 0$ otherwise. By the Dirichlet theorem on arithmetic progressions, we obtain

$$\frac{1}{N}\sum_{x\in\mathbb{Z}_N} f(x) = \frac{k^{-1}2^{-k-5}}{N}\sum_{x\in[\varepsilon_k N, 2\varepsilon_k N]}\widetilde{\Lambda}(x) = k^{-1}2^{-k-5}\varepsilon_k(1+o(1)). \tag{119}$$

Applying now Proposition 13 and Theorem 43, we see that

$$\frac{1}{N^2}\sum_{x,r\in\mathbb{Z}_N} f(x)f(x+r)\cdots f(x+(k-1)r) \geqslant c'(k, k^{-1}2^{-k-5}) + o(1). \tag{120}$$

The contribution of the terms with $r = 0$ to the sum (120) is $O((\log N)^k/N) = o(1)$. Hence, we can assume that the summation in (120) ranges over the indices with $r \neq 0$. Since $\varepsilon_k < 1/k$, it follows that all the numbers $x, x+r, \ldots, x+(k-1)r$ belong to $[N]$. Thus, the summation in (120) ranges over the indices with $x, r \in [N]$, $r \neq 0$, such that all the numbers $x+r, \ldots, x+(k-1)r$ belong to $[N]$, and hence the set of primes contains a non-trivial arithmetic progression of length k. This completes the proof of Theorem 41.

Thus, to prove the Green–Tao result, it suffices to prove Proposition 13 by constructing a dominating k-pseudorandom measure $\nu(x)$.

We need a definition given by Goldston and Yildirim (see [103]–[105]).

Definition 16 (Goldston, Yildirim). Let N be a prime and let R be a real parameter depending on N. Then

$$\Lambda_R(x) := \sum_{d|x, d\leqslant R}\mu(d)\log(R/d) = \sum_{d|x}\mu(d)\log(R/d)_+, \tag{121}$$

where μ is the Möbius function and $\log(x)_+ = \max(\log x, 0)$.

Using Definition 16, Green and Tao construct a dominating k-pseudorandom measure $\nu\colon \mathbb{Z}_N \to \mathbb{R}^+$.

Definition 17. Let $R = N^{k^{-1}2^{-k-4}}$ and $\varepsilon_k = 1/(2^k(k+4)!)$. Then we set

$$\nu(x) := \begin{cases} \dfrac{\varphi(W)}{W}\dfrac{\Lambda_R(Wn+1)^2}{\log R} & \text{if } x \in [\varepsilon_k N, 2\varepsilon_k N], \\ 1 & \text{if } x \in \mathbb{Z}_N \setminus [\varepsilon_k N, 2\varepsilon_k N]. \end{cases}$$

We assert that $\nu(x)$ dominates $\widetilde{\Lambda}(x)$.

Lemma 8. *$\nu(x) \geqslant 0$ for any $x \in \mathbb{Z}_N$. Moreover, $\nu(x) \geqslant k^{-1}2^{-k-5}\widetilde{\Lambda}(x)$ for any sufficiently large N and any $x \in [\varepsilon_k N, 2\varepsilon_k N]$.*

Proof. The first assertion of the lemma is trivial. If the number $Wx+1$ is not prime, then the second assertion of the lemma is also trivial. Let $Wx+1$ be a prime and let N be such that $Wx+1 > R$. In this case the sum (121) contains only one term (with $d=1$). This implies that $\Lambda_R(Wx+1) = \log R$ and $\nu(x)\dfrac{\varphi(W)}{W}\log R \geqslant k^{-1}2^{-k-5}\widetilde{\Lambda}(x)$ and completes the proof of Lemma 8.

It remains to prove that ν is a k-pseudorandom measure. To this end, one must prove that ν satisfies the condition of linear forms and the correlation condition. Developing the approach of the paper [105], Green and Tao proved two propositions which imply the conditions (108), (109), and (110) for the measure ν. We do not dwell on the proofs in detail but simply formulate these propositions and refer the interested reader to the paper [24].

Proposition 14. *Let m and t be positive integers, let $\psi_i(\mathbf{x}) := \sum_{j=1}^t L_{ij} + b_j$, $i \in [m]$, be linear forms with integral coefficients L_{ij} such that $|L_{ij}| \leqslant \sqrt{w(N)}/2$, $i \in [m]$, $j \in [t]$, let $\theta_i = W\psi_i + 1$, and let $B = \prod_{i=1}^t I_i \subseteq \mathbb{R}^t$, where I_i, $i \in [m]$, are intervals with length at least R^{10m}. Suppose that the function $w(N)$ grows to infinity sufficiently slowly. Then*

$$\frac{1}{|B|}\sum_{\mathbf{x}\in B}\Lambda_R(\theta_1(\mathbf{x}))^2\cdots\Lambda_R(\theta_m(\mathbf{x}))^2 = (1+o_{m,t}(1))\left(\frac{W\log R}{\varphi(W)}\right)^m. \tag{122}$$

Proposition 15. *Let $m \geqslant 1$ be an integer, let I be an interval with length at least R^{10m}, let $h_1, \dots, h_m$ be distinct integers with $|h_i| \leqslant N^2$ for $i \in [m]$, and let*

$$\Delta := \prod_{1\leqslant i<j\leqslant m}|h_i - h_j|.$$

Suppose that the function $w(N)$ grows to infinity sufficiently slowly. Then

$$\begin{aligned}\frac{1}{|I|}\sum_{x\in I}&\Lambda_R(W(x+h_1)+1)^2\cdots\Lambda_R(W(x+h_m)+1)^2\\ &\leqslant (1+o_m(1))\left(\frac{W\log R}{\varphi(W)}\right)^m\prod_{p|\Delta}(1+O_m(p^{-1/2})).\end{aligned} \tag{123}$$

§ 9. Rado's theorem on systems of linear equations

In § 2 we proved Roth's theorem, Theorem 3, which asserts that the order of the cardinality of any subset of $[N]$ without arithmetic progressions of length three does not exceed $1/\log\log N$. Roth generalized this result in the paper [38].

Let $U = (u_{ij})$ be an $m \times n$ matrix with all its elements integers. A set A is called a *$\mathscr{U}$-set* if A does not contain n *distinct* elements $x_1, \dots, x_n$ satisfying the m equations

$$\sum_{j=1}^n u_{ij}x_j = 0, \qquad i = 1, \dots, m. \tag{124}$$

Let

$$a_U(N) = \frac{1}{N} \max\{|A| : A \subseteq [N],\ A \in \mathscr{U}\}.$$

Roth's result can be reformulated in terms of the function $a_U(N)$ as follows. If U is the 1×3 matrix $(1, -2, 1)$, then $a_U(N) \ll 1/\log\log N$. The question arises as to what conditions on the matrix U ensure that $a_U(N) \to 0$ as $N \to \infty$.

We formulate the main result of the paper [38].

Theorem 44 (Roth). *Let U be a matrix satisfying the following two conditions*:
(a) $\sum_{j=1}^{n} u_{ij} = 0$ *for any* $i = 1, \dots, m$;
(b) *U has m linearly independent columns with the property that if any one of them is removed, then the remaining $n-1$ columns of U can be partitioned into two sets which each contain m linearly independent columns.*
Then

$$a_U(N) \ll \frac{1}{(\log\log N)^{1/l^2}}.$$

We note that the condition (a) in Theorem 44 is necessary for validity of the condition $a_U(N) \to 0$ as $N \to \infty$. Indeed, if

$$\sum_{j=1}^{n} u_{qj} = D$$

for some $q \in [n]$ and some $D \in \mathbb{Z}$, $D \neq 0$, then all the elements x_j equal to 1 modulo $|D|+1$ do not satisfy the qth equation of the system (124). Hence, $a_U(N) \geqslant 1/(|D|+1) > 0$.

On the contrary, the second condition of Theorem 44 is not necessary. Consider the 2×4 matrix

$$U = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{pmatrix}.$$

Then the matrix U does not satisfy the second condition in Theorem 44, because this condition implies that $n \geqslant 2m+1$. On the other hand, we have $a_U(N) = a_4(N)$, and it follows from Szemerédi's theorem that $a_4(N) \to 0$ as $N \to \infty$. We shall describe the matrices U such that $\lim_{N\to\infty} a_U(N) = 0$ a little later. For now, we state an interesting result of Rado (see [39]–[41]).

Definition 18. Let $U = (u_{ij})$ be an $m \times n$ matrix with all its elements integers. The system of equations (124) is said to be *regular in* $\mathbb{N}$ if for any colouring of $\mathbb{N}$ with finitely many colours there is a monochromatic solution of the system (124).

We note that the numbers $x_1, \dots, x_n$ are not assumed to be distinct.

Theorem 45 (Rado). *Let $U = (u_{ij})$ be an $m \times n$ matrix with all its elements integers. The system of equations (124) is regular in $\mathbb{N}$ if and only if there are columns $C_1, \dots, C_n$ and numbers k_i with $1 \leqslant k_1 < \dots < k_t = n$ such that the new columns*

$$A_i = \sum_{j=k_{i-1}+1}^{k_i} C_j$$

satisfy the following conditions:

(a) A_1 *is the zero column*;

(b) *for any* $i = 1, \dots, t$ *the column* A_i *is a linear combination of the columns* $C_1, \dots, C_{k_{i-1}}$ *with rational coefficients.*

The matrices satisfying the conditions (a) and (b) in Theorem 45 are said to be *regular*.

Rado's theorem can be written out in an especially simple way for $m = 1$.

Theorem 46 (Rado). *Let* n *be a positive integer and let* $c_1, \dots, c_n$ *be non-zero integers. The system of equations*

$$c_1 x_1 + \dots + c_n x_n = 0 \tag{125}$$

is regular in $\mathbb{N}$ *if and only if there is a non-empty set* $I \subseteq [n]$ *such that the sum* $\sum_{i \in I} c_i$ *vanishes.*

For example, the equations $x - 2y + z = 0$ and $x + y - z = 0$ are regular, whereas the equation $x + y - 5z = 0$ is not. If $x - 2y + z = 0$, then the numbers x, y, z form an arithmetic progression. We note that Theorem 46 holds trivially for this equation, because one can set $x = y = z = 1$. Van der Waerden's theorem lets us assert for $k = 3$ that for any colouring of $\mathbb{N}$ with finitely many colours there are *distinct* x, y, z of the same colour which satisfy the equation $x - 2y + z = 0$. As for the equation $x + y - z = 0$, the existence of a monochromatic solution of this equation for any finite colouring of $\mathbb{N}$ was proved earlier by I. Schur in [37].

We do not present here a complete proof of Theorem 46 (nor, all the more so, of Theorem 45). Nevertheless, the proof of the necessity condition in Theorem 46 is rather simple.

Thus, let the sum $\sum_{i \in I} c_i$ be non-zero for any non-empty set $I \subseteq [n]$, and consider a colouring of $\mathbb{N}$ with finitely many colours for which there is no monochromatic solution of equation (125). Let p be a prime, which will be chosen below. We colour the set $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ with $(p-1)$ colours. It is clear that this gives a colouring of the set $\mathbb{N}$ as well.

An arbitrary $q \in \mathbb{Q}^*$ can be uniquely represented in the form

$$q = \frac{p^j a}{b}, \qquad j \in \mathbb{Z}, \quad a \in \mathbb{Z}, \quad b \in \mathbb{N}, \quad p \nmid a, \quad p \nmid b, \quad (a, b) = 1. \tag{126}$$

If q is represented in this form, then we colour it with the colour $S_p(q)$, where $S_p(q) = ab^{-1} \pmod p \in \mathbb{Z}_p^*$. The colouring S_p of the set $\mathbb{Q}^*$ with $(p-1)$ colours has the following property: if $S_p(x) = S_p(y)$, then $S_p(\alpha x) = S_p(\alpha y)$ for any $\alpha \in Q^*$.

Since the number of subsets $I \subseteq [n]$ is finite and for any non-empty $I \subseteq [n]$ we have $\sum_{i \in I} c_i \neq 0$, there is a p such that $\sum_{i \in I} c_i \neq 0 \pmod p$ for any $\varnothing \neq I \subseteq [n]$.

Suppose that $x_1, \dots, x_n \in \mathbb{N}$ is a monochromatic solution of the equation (125) for the colouring S_p. Since the family $\mu x_1, \dots, \mu x_n$ is also monochromatic for any $\mu \in \mathbb{Q}^*$, we can assume without loss of generality that the greatest common divisor of $x_1, \dots, x_n$ is equal to one. Re-indexing the numbers $x_1, \dots, x_n$ if necessary, we can find an s, $1 \leqslant s \leqslant n$, such that p does not divide the numbers $x_1, \dots, x_s$ and

divides the numbers $x_{s+1}, \dots, x_n$. We have

$$\sigma = \sum_{j=1}^{n} c_j x_j \equiv \sum_{j=1}^{k} c_j x_j \equiv 0 \pmod p.$$

Since the numbers $x_1, \dots, x_n$ have the same colour, it follows that $x_1 \equiv \dots \equiv x_n \equiv a \pmod p$, where $a \neq 0 \pmod p$. Therefore,

$$\sigma \equiv a \cdot \sum_{j=1}^{k} c_j \equiv 0 \pmod p.$$

Hence, $\sum_{j=1}^{k} c_j \equiv 0 \pmod p$, a contradiction.

We say a few words about results relating to Theorem 45. There is an analogue of Rado's theorem for arbitrary Abelian groups in the paper [107]. Furstenberg [17] proved Rado's theorem by ergodic theory methods. Results similar to Theorem 45 were obtained for certain non-linear equations in the papers [108]–[110]. For other results generalizing Rado's theorem, see [111] and [17].

We can now describe all the matrices U such that $\lim_{N\to\infty} a_U(N) = 0$. It is clear that these matrices must satisfy the conditions of Theorem 45. Moreover, the first condition of Theorem 44 must hold for these matrices. It turns out that these necessary conditions are also sufficient. More precisely, the following result holds (see [42]) as a consequence of Szemerédi's theorem.

Theorem 47 (Frankl, Graham, Rödl). *Let $U = (u_{ij})$ be a regular $m \times n$ matrix with all its elements integers. Suppose that the system of equations $U\mathbf{x} = 0$ has at least one solution $\mathbf{x}' = (x'_1, \dots, x'_n)$ with pairwise distinct elements x'_i. In this case the following two assertions are equivalent*:

(a) $\sum_{j=1}^{n} u_{ij} = 0$, $i = 1, \dots, m$;

(b) *for any set $E \subseteq \mathbb{N}$ with $D^*(E) > 0$ the system of equations $U\mathbf{x} = 0$ has a solution $\mathbf{x} = (x_1, \dots, x_n)$ with its elements x_i in E and pairwise distinct.*

Proof. (1)⇒(2) Let $\mathbf{x}' = (x'_1, \dots, x'_n)$ be a solution of the system of equations $U\mathbf{x}' = 0$ and suppose that the elements x'_i are pairwise distinct. Let $N = \max x'_i$. Since the set X has positive upper density, it follows from Szemerédi's theorem that there is an arithmetic progression of length N in X. Let $c + jd$ be the elements of this progression, $j = 0, 1, \dots, N$. We also assume that $\mathbf{y} = c \cdot \mathbf{1} + d \cdot \mathbf{x}'$, where $\mathbf{1} = (1, \dots, 1)$. By the property (a), we see that $A\mathbf{y} = 0$. Since $N = \max x'_i$, all the components of the vector $\mathbf{y}$ belong to the set X, as was to be proved.

(2)⇒(1) Let N be a positive integer such that $N > \sum_{i,j} |u_{ij}|$, and let $X = \{Ny + 1 : y \in \mathbb{N}\}$. Then the upper density of the set X is equal to $1/N$. Let $\mathbf{x} = (x_1, \dots, x_n) \in X^n$ be a solution of the system of equations $U\mathbf{x} = 0$. Then for some integers $y_i \in \mathbb{N}$ we have $x_i = Ny_i + 1$ and

$$0 = \sum_{j=1}^{n} u_{ij} x_j = N\left(\sum_{j=1}^{n} u_{ij} y_j\right) + \sum_{j=1}^{n} u_{ij}. \tag{127}$$

It follows from (127) that $\sum_{j=1}^{n} u_{ij} = 0$ for $i = 1, \dots, m$. Indeed, if $\sum_{j=1}^{n} u_{ij} y_j = 0$ for some $i \in [m]$, then $\sum_{j=1}^{n} u_{ij} = 0$ as well. But $\sum_{j=1}^{n} u_{ij} y_j \neq 0$ cannot hold, by the choice of N. Theorem 47 is proved.

§ 10. Other results concerning arithmetic progressions

The problems studied above can be described as follows. Let a set A be contained in some 'base' set B and have a sufficiently large density with respect to B. Then the set A contains a family of points $x_1, \dots, x_m$ which has some prescribed properties.

As the set B we have taken the set of integers (Szemerédi's theorem), a segment $[N]$ (Roth's theorem), a two-dimensional lattice $[N]^2$ or $\mathbb{Z}_2^n \times \mathbb{Z}_2^n$ (Theorems 36 and 37), the set of all primes (Theorem 10), and other sets. As the properties which must be satisfied by a family of points $x_1, \dots, x_m$ in A we have mainly considered the property that the points form an arithmetic progression of some length, or form a corner (Theorems 36 and 37), or satisfy a system of linear equations (Theorems 44, 45, and 47).

In this section we consider problems on arithmetic progressions that do not fit in the above scheme, namely, problems about *critical sets*, problems about arithmetic progressions in *sums*, and, finally, theorems on *rainbows*.

In the papers [62] and [112] Croot studied the problem of the structure of critical sets in $\mathbb{Z}_p$ without arithmetic progressions of length three. Let $\rho > 0$. A set $C \subseteq \mathbb{Z}_p$ with cardinality at least ρp is said to be *ρ-critical* if C contains minimally many arithmetic progressions of length three among all sets with cardinality at least ρp. In [62] Croot proved that every critical set has a strong additive structure. His theorem uses the well-known conjecture on prime numbers.

Conjecture 4. *The segment $[x, x + x^\theta]$ contains a prime for any $\theta > 0$ and any sufficiently large x.*

At present, Conjecture 4 has been proved for any $\theta > 0.525$ (see [113]).

Theorem 48 (Croot). *Let $\rho_0 \in (0,1)$. There exist numbers $\rho \in (0, \rho_0)$ and $d \in (0,1)$ and infinitely many primes p such that for any ρ-critical set $C \subseteq \mathbb{Z}_p$ there is a number b with $1 \leqslant b \leqslant p-1$ for which*

$$|C \cap (C + bj)| \geqslant |C| \biggl(1 - \frac{K}{|\log \rho|}\biggr), \qquad j = 0, 1, \dots, p^d,$$

where $K > 0$ is an absolute constant.

Croot's result can be reformulated as follows. If C is a ρ-critical set, then $C \approx A + B$, where $B = \{0, b, 2b, \dots, [p^d]b\}$. Here the symbol $\approx$ means that the cardinality of the symmetric difference $C \triangle (A + B)$ does not exceed $O(|C| \log^{-1}(1/\rho))$. Hence, the less the number ρ is, the closer the set C is to the sum $A + B$. We note that the set constructed in Theorem 20 of § 4 is of the very form $A + B$, where B is an arithmetic progression.

In the paper [112] Croot obtained an unconditional result on the structure of critical sets.

Theorem 49 (Croot). *Let $\rho \in (0,1)$ and let p be a sufficiently large positive integer. Then every ρ-critical set $C \subseteq \mathbb{Z}_p$ contains an arithmetic progression of length at least $\log^{1/4 + o(1)} p$.*

Let us now consider problems on arithmetic progressions in sums, that is, in sets of the form $A_1 + A_2 + \dots + A_k$. It turns out that such sets contain surprisingly

long arithmetic progressions. Many papers on this topic have appeared recently; for instance, see [114]–[126]. We cannot even begin to mention all the results obtained in this direction, but we touch on just a few of them. In [114] Bourgain obtained the following result.

Theorem 50 (Bourgain). *Let $A, B \subseteq [N]$, $|A| = \gamma N$ and $|B| = \delta N$. Then there is an absolute constant $c > 0$ such that the set $A+B$ contains an arithmetic progression with length at least* $\exp(c(\gamma\delta \log N)^{1/3} - \log\log N)$.

On the other hand, Ruzsa found a lower bound for the length of a maximal arithmetic progression in a set of the form $A + A$ (see [115]).

Theorem 51 (Ruzsa). *Let $\varepsilon > 0$ be an arbitrary number. Then there exists a number $p_0(\varepsilon)$ such that for any prime p with $p > p_0(\varepsilon)$ there is a set $A \subseteq \mathbb{Z}_p$ which is symmetric (that is, $A = -A$) and such that $|A| > (1/2 - \varepsilon)p$ and the sum $A + A$ contains no arithmetic progressions whose length exceeds* $\exp((\log p)^{2/3+\varepsilon})$.

Since for any set $A \subseteq \mathbb{Z}_p$ with $|A| \geqslant p/2$ we have $A + A = \mathbb{Z}_p$, it follows that the constant $1/2$ in Theorem 51 is the best possible. Let us compare Theorems 50 and 51. Suppose that the parameters γ and δ in Theorem 50 do not depend on N. According to this theorem, for any A the set $A+A$ contains arithmetic progressions with length at least $\exp(c(\log N)^{1/3})$, where c is a constant. On the other hand, by Theorem 51, there is an A such that $A + A$ contains no progressions with length greater than $\exp(c(\log N)^{2/3})$. As we see, the estimates in Bourgain's and Ruzsa's theorems are fairly close.

Freiman, H. Halberstam, and Ruzsa [117] considered the problem of arithmetic progressions in sets of the form $A + A + A$ and proved the following result (see also [118]).

Theorem 52 (Freiman, Halberstam, Ruzsa). *Let N be a positive integer, let $\delta > 0$, and let $A \subseteq \mathbb{Z}_N$ be an arbitrary set of cardinality δN. Then the set $A + A + A$ contains an arithmetic progression with length at least $c\delta N^{C\delta^3}$, where $c, C > 0$ are absolute constants.*

In the same paper the three authors proved a theorem similar to Theorem 51, namely, they constructed a set $A \subseteq \mathbb{Z}_N$ such that the sum $A + A + A$ contains no arithmetic progressions with length greater than $2N^{\log(1/\delta)}$.

In [116] Green improved Theorems 50 and 52.

Theorem 53 (Green). *Let $A, B \subseteq [N]$, $|A| = \gamma N$ and $|B| = \delta N$. Then there is an absolute constant $C > 0$ such that the set $A+B$ contains an arithmetic progression with length at least* $\exp(C(\gamma\delta \log N)^{1/2} - \log\log N)$.

Theorem 54 (Green). *Let N be a positive integer, let $\delta > 0$, and let $A \subseteq \mathbb{Z}_N$ be an arbitrary set of cardinality δN. Then the set $A + A + A$ contains an arithmetic progression with length at least $2^{-24}\delta^5(\log(1/\delta))^{-2}N^{\delta^2/(250\log(1/\delta))}$.*

In the last part of this section we consider problems relating to arithmetic progressions in rainbows.

Let $c\colon \mathbb{N} \to \{R, G, B\}$ be an arbitrary colouring of the set of positive integers with some colours R, G, and B. It follows from van der Waerden's theorem that for any such colouring of $\mathbb{N}$ there is a monochromatic arithmetic progression of

length three. The question arises as to whether for an arbitrary colouring c one can always find an arithmetic progression with all its elements coloured differently. Such questions are called *anti-Ramsey* questions (see the first paper on this topic, [127], and also [128]). As we shall see a little later, the answer to the question posed above is negative. Nevertheless, the following result was obtained in the paper [129] (see also [130] and [131]).

An arithmetic progression formed of the numbers a_1, a_2, a_3 is called a *rainbow* if $c(a_i) \neq c(a_j)$ for any $i \neq j$. We introduce the set $R_c(n) := [n] \cap \{i : c(i) = R\}$ and the analogous sets $G_c(n)$ and $B_c(n)$.

Theorem 55 [129]. *Let c be an arbitrary colouring of $\mathbb{N}$. Assume that*

$$\limsup_{n\to\infty}(\min\{|R_c(n)|, |G_c(n)|, |B_c(n)|\} - n/6) = +\infty. \tag{128}$$

Then the colouring c admits a rainbow.

Instead of a colouring of the positive integers, one can consider a colouring of the set $\mathbb{Z}_n$ of residues. By a rainbow in $\mathbb{Z}_n$ we mean residues a_1, a_2, a_3 coloured with three distinct colours and such that $a_1 + a_2 \equiv 2a_3 \pmod n$. Theorem 55 implies the following corollary.

Corollary 4. *Let n be a positive integer, let c be an arbitrary colouring of $\mathbb{Z}_n$, and let $R_c = \{i : c(i) = R\}$, $G_c = \{i : c(i) = G\}$, and $B_c = \{i : c(i) = B\}$. If $\min(|R_c|, |G_c|, |B_c|) > n/6$, then the colouring c admits a rainbow.*

Indeed, let c be an arbitrary colouring of $\mathbb{Z}_n$ and consider the colouring $\overline{c}$ of $\mathbb{N}$ such that $\overline{c}(i) := c(i \pmod n)$. By assumption, $\min(|R_c|, |G_c|, |B_c|) > n/6$. Therefore,

$$\limsup_{n\to\infty}(\min\{|R_c(n)|, |G_c(n)|, |B_c(n)|\} - n/6) = +\infty.$$

Using Theorem 55, we find a rainbow in $\overline{c}$, which gives a rainbow in c.

The following proposition shows that the constant $1/6$ in the inequality (128) cannot be replaced by a smaller constant.

Proposition 16. *There is a colouring c of $\mathbb{N}$ such that*

$$\min\{|R_c(n)|, |G_c(n)|, |B_c(n)|\} = [(n+2)/6]$$

for all $n \in \mathbb{N}$.[2]

Proof. Consider the following colouring c of the set $\mathbb{N}$:

$$c(i) = \begin{cases} R & \text{if } i \equiv 1 \pmod 6, \\ G & \text{if } i \equiv 4 \pmod 6, \\ B & \text{otherwise.} \end{cases} \tag{129}$$

One can easily see that c contains no rainbows, and for any n we have

$$\min\{|R_c(n)|, |G_c(n)|, |B_c(n)|\} = |G_c(n)| = [(n+2)/6].$$

[2] *Russian Editors' note*: The author means a colouring which admits no rainbows.

§ 11. Concluding remarks

We complete the present survey with a formulation of several unsolved problems related to Szemerédi's theorem and to problems on arithmetic progressions. Some of these problems have already been formulated in the previous sections.

The most complicated unsolved problem is still the second Erdős–Turán conjecture, Conjecture 2. As was shown in § 5, this conjecture is closely related to the problem of the behaviour of the function $a_k(N)$. We note that the exact order of growth of $a_3(N)$ remains unknown so far (even in the simplest case $k = 3$). Green and Tao recently announced a result on an upper bound of the form $a_k(N) \ll 1/(\log N)^{C_k}$ for $k \geqslant 4$, where $C_k > 0$ is an absolute constant, and they proved a similar result for $a_4(N)$ in the groups $\mathbb{Z}_p^n$, where p is a prime, $p \neq 2, 3$ (see [33]).

Another problem is to obtain a quantitative version of the Bergelson–Leibman theorem, Theorem 8. This theorem can be reformulated in the language of number theory as follows.

Theorem 56 (Bergelson, Leibman). *Let $\delta > 0$, and let $p_1, p_2, \dots, p_k$ be polynomials such that $p_i(\mathbb{N}) \subseteq \mathbb{N}$ and $p_i(0) = 0$, $i = 1, \dots, k$. Then there exists a positive integer $N(\delta, p_1, \dots, p_k)$ such that for any set $A \subseteq [N]$ with $|A| \geqslant \delta N$ there are positive integers a and d for which all the numbers $a + p_i(d)$, $i = 1, \dots, k$, belong to A.*

We use the words "quantitative version of Theorem 56" to mean a result in which an explicit upper bound for the quantity $N(\delta, p_1, \dots, p_k)$ is established.

If $k = 2$, $p_1(n) \equiv 0$, and $p_2(n) = n^2$, then a quantitative analogue of the Bergelson–Leibman theorem is well known (see Theorem 30). Nevertheless, the actual order of growth of the function $N(\delta, p_1, p_2)$ has not been found, even in this simple case. This leads one to formulate the following problem.

Problem. Let $\varepsilon > 0$ and let $N > N_0(\varepsilon)$ be a sufficiently large positive integer. Is there a set $A \subseteq [N]$ with $|A| > N^{1-\varepsilon}$ such that the difference $A - A$ contains no non-zero squares?

In the paper [132] Ruzsa showed that if $\varepsilon = 0.267$, then there is a set $A \subseteq [N]$ with $|A| > N^{1-\varepsilon}$ such that the difference between any two elements in A is not a non-zero square.

In connection with Theorems 34–36, the problem arises naturally of obtaining multidimensional quantitative analogues of the theorem on corners, and also of a quantitative version of the Hales–Jewett theorem (Theorem 31). A very interesting problem is to obtain an analogue of Theorem 56 when A is the set of primes.

The author expresses his deep gratitude to Doctor of the Physical and Mathematical Sciences N. G. Moshchevitin for his undivided attention to the present work, and also to M. G. Ryumina for her hospitality.

Bibliography

[1] B. L. Van der Waerden, "Beweis einer Baudetschen Vermutung", *Nieuw Arch. Wisk.* **15** (1927), 212–216.

[2] A. Ya. Khinchin, *Three pearls of number theory*, 2nd ed., OGIZ, Moscow–Leningrad 1948; 3rd ed., Nauka, Moscow 1979; Reprint, URSS, Moscow 2004 (Russian); English transl. of 2nd ed., Graylock Press, Rochester, NY 1952; Reprint, Dover, Mineola, NY 1998.

[3] R. L. Graham, *Rudiments of Ramsey theory*, Expository lectures from the CBMS Regional Conference held at St. Olaf College (June 18–22, 1979), CBMS Regional Conference Series in Math., vol. 45, Amer. Math. Soc., Providence, RI 1981.

[4] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey theory*, Wiley, New York 1980.

[5] V. A. Uspenskii, *Lectures on computable functions*, Fizmatgiz, Moscow 1960 (Russian); French transl., V. A. Ouspenski, *Leçons sur les fonctions calculables*, Actualités Scientifiques et Industrielles, vol. 1317, Hermann, Paris 1966.

[6] S. Shelah, "Primitive recursive bounds for van der Waerden numbers", *J. Amer. Math. Soc.* **1**:3 (1988), 683–697.

[7] K. F. Roth, "On certain sets of integers", *J. London Math. Soc.* **28** (1953), 104–109.

[8] E. Szemerédi, "Integer sets containing no arithmetic progressions", *Acta Math. Hungar.* **56**:1–2 (1990), 155–158.

[9] D. R. Heath-Brown, "Integer sets containing no arithmetic progressions", *J. London Math. Soc.* (2) **35**:3 (1987), 385–394.

[10] J. Bourgain, "On triples in arithmetic progression", *Geom. Funct. Anal.* **9**:5 (1999), 968–984.

[11] J. Bourgain, "A Szemerédi type theorem for sets of positive density in $\mathbb{R}^k$", *Israel J. Math.* **54**:3 (1986), 307–316.

[12] I. Z. Ruzsa and E. Szemerédi, "Triple systems with no six points carrying three triangles", *Combinatorics* (Keszthely, 1976), Proceedings of the 5th Hungarian colloquium, Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, Amsterdam 1978, pp. 939–945.

[13] E. Szemerédi, "On sets of integers containing no four elements in arithmetic progression", *Acta Math. Acad. Sci. Hungar.* **20**:1–2 (1969), 89–104.

[14] E. Szemerédi, "On sets of integers containing no k elements in arithmetic progression", *Acta Arith.* **27** (1975), 199–245.

[15] H. Furstenberg, "Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions", *J. Anal. Math.* **31** (1977), 204–256.

[16] H. Furstenberg, Y. Katznelson, and D. Ornstein, "The ergodic theoretical proof of Szemerédi's theorem", *Bull. Amer. Math. Soc.* (*N.S.*) **7**:3 (1982), 527–552.

[17] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, M. B. Porter Lectures, Princeton Univ. Press, Princeton, NJ 1981.

[18] H. Furstenberg and Y. Katznelson, "An ergodic Szemerédi theorem for commuting transformations", *J. Anal. Math.* **34** (1978/1979), 275–291.

[19] V. Bergelson and A. Leibman, "Polynomial extensions of van der Waerden's and Szemerédi's theorems", *J. Amer. Math. Soc.* **9**:3 (1996), 725–753.

[20] H. Furstenberg and Y. Katznelson, "A density version of the Hales–Jewett theorem", *J. Anal. Math.* **57** (1991), 64–119.

[21] A. Leibman, "Multiple recurrence theorem for measure preserving actions of a nilpotent group", *Geom. Funct. Anal.* **8**:5 (1998), 853–931.

[22] W. T. Gowers, "A new proof of Szemerédi's theorem", *Geom. Funct. Anal.* **11**:3 (2001), 465–588.

[23] W. T. Gowers, "A new proof of Szemerédi's theorem for arithmetic progressions of length four", *Geom. Funct. Anal.* **8**:3 (1998), 529–551.

[24] B. Green and T. Tao, "The primes contain arbitrarily long arithmetic progressions", *Ann. of Math.* (to appear); `arXiv:math.NT/0404188`.

[25] B. Green, "On arithmetic structures in dense sets of integers", *Duke Math. J.* **144**:2 (2002), 215–238.

[26] J. Bourgain and M.-C. Chang, "Exponential sum estimates over subgroups and almost subgroups of $\mathbb{Z}_Q^*$, where Q is composite with few prime factors", *Geom. Funct. Anal.* **16**:2 (2006), 327–366.

[27] J. Bourgain, A. A. Glibichuk, and S. V. Konyagin, "Estimates for the number of sums and products and for exponential sums in fields of prime order", *J. London Math. Soc.* (2) **73**:2 (2006), 380–398.

[28] I. D. Shkredov, "On a problem of Gowers", *Dokl. Akad. Nauk* **400**:2 (2005), 169–172 (Russian); English transl., *Dokl. Math.* **71**:1 (2005), 46–48.

[29] I. D. Shkredov, "On a problem of Gowers", *Izv. Ross. Akad. Nauk Ser. Mat.* **70**:2 (2006), 179–221 (Russian); English transl., *Izv. Math.* **70**:2 (2006), 385–425.

[30] I. D. Shkredov, "On a generalization of Szemeredi's theorem", *Dokl. Akad. Nauk* **405**:3 (2005), 315–319 (Russian); English transl., *Dokl. Math.* **72**:3 (2005), 899–902.

[31] I. D. Shkredov, *On a generalization of Szemerédi's theorem*, arXiv: math.NT/0503639.

[32] B. Green and T. Tao, *An inverse theorem for the Gowers U^3 norm*, arXiv: math.NT/0503014.

[33] B. Green and T. Tao, *New bounds for Szemerédi's theorem.* I: *Progressions of length* 4 *in finite field geometries*, arXiv: math.CO/0509560.

[34] B. Green, "Roth's theorem in the primes", *Ann. of Math.* (2) **161**:3 (2005), 1609–1636.

[35] A. A. Bukhshtab [Buchstab], *Theory of numbers*, 2nd corr. ed., Prosveshchenie, Moscow 1966. (Russian)

[36] I. M. Vinogradov, *Elements of number theory*, 5th ed., Gostekhizdat, Moscow–Leningrad 1949; 9th ed., Nauka, Moscow 1981; Reprint, Lan', St. Petersburg 2004 (Russian); English transl. of 5th ed., Dover, New York 1954.

[37] I. Schur, "Über die Kongruenz $x^m + y^m \equiv z^m \pmod p$", *Jahresber. Deutsch. Math.-Verein.* **25**:4–6 (1916), 114–117.

[38] K. F. Roth, "On certain sets of integers, II", *J. London Math. Soc.* **29** (1954), 20–26.

[39] R. Rado, "Verallgemeinerung eines Satzes von van der Waerden mit Anwendungen auf ein Problem der Zahlentheorie", *Sitz. Preuß. Akad. Wiss. Phys.-Math. Kl.* **17** (1933), 589–596.

[40] R. Rado, "Studien zur Kombinatorik", *Math. Z.* **36**:1 (1933), 424–470.

[41] R. Rado, "Some recent results in combinatorial analysis", *Comptes Rendus du Congrès International des Mathématiciens* (Oslo, 1936), vol. 2 1937, pp. 20–21.

[42] P. Frankl, R. L. Graham, and V. Rödl, "Quantitative theorems for regular systems of equations", *J. Combin. Theory Ser. A* **47**:2 (1988), 246–261.

[43] T. Tao, *A quantitative ergodic theory proof of Szemerédi's theorem*, arXiv: math.CO/0405251.

[44] N. Alon and J. H. Spencer, *The probabilistic method*, Wiley, New York 1992.

[45] F. A. Behrend, "On sets of integers which contain no three terms in arithmetic progression", *Proc. Nat. Acad. Sci. USA* **32** (1946), 331–332.

[46] R. Salem and D. C. Spencer, *Proc. Nat. Acad. Sci. USA* **28** (1942), 561–563.

[47] R. Salem and D. C. Spencer, "On sets which do not contain a given number of terms in arithmetical progression", *Nieuw Arch. Wisk.* (2) **23** (1950), 133–143.

[48] I. Laba and M. T. Lacey, *On sets of integers not containing long arithmetic progressions*, arXiv: math.CO/0108155.

[49] R. A. Rankin, "Sets of integers containing not more than a given number of terms in arithmetic progression", *Proc. Roy. Soc. Edinburgh Sect. A* **65**:4 (1960/1961), 332–344.

[50] R. A. Rankin, "Representations of a number as the sum of a large number of squares", *Proc. Roy. Soc. Edinburgh Sect. A* **65**:4 (1960/1961), 318–331.

[51] P. Erdős and P. Turán, "On some sequences of integers", *J. London Math. Soc.* **11** (1936), 261–264.

[52] L. Moser, "On non-averaging sets of integers", *Canadian J. Math.* **5** (1953), 245–252.

[53] W. T. Gowers, "Lower bounds of tower type for Szemerédi's uniformity lemma", *Geom. Funct. Anal.* **7**:2 (1997), 322–337.

[54] E. Szemerédi, "Regular partitions of graphs", *Problèmes combinatoires et théorie des graphes* (Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris 1978, pp. 399–401.

[55] J. Komlós and M. Simonovits, "Szemerédi's regularity lemma and its applications in graph theory", *Combinatorics* (Keszthely, 1993) (D. Miklós, V. T. Sós, and T. Szőnyi, eds.), Paul Erdős is Eighty, Bolyai Soc. Math. Stud., vol. 2, János Bolyai Math. Soc., Budapest 1996, pp. 295–352.

[56] Y. Kohayakawa, "Szemerédi's regularity lemma for sparse graphs", *Foundations of computational mathematics* (Rio de Janeiro), Springer-Verlag, Berlin 1997, pp. 216–230.

[57] R. L. Graham and V. Rödl, "Numbers in Ramsey theory", *Surveys in combinatorics* (New Cross, 1987), London Math. Soc. Lecture Note Ser., vol. 123, Cambridge Univ. Press, Cambridge 1987, pp. 111–153.

[58] B. Nagle, V. Rödl, and M. Schacht, "The counting lemma for regular k-uniform hypergraphs", *Random Structures Algorithms* **28**:2 (2006), 113–179.

[59] W. T. Gowers, "Quasirandomness, counting and regularity for 3-uniform hypergraphs", *Combin. Probab. Comput.* **15**:1–2 (2006), 143–184.

[60] W. T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, `http://www.dpmms.cam.ac.uk/~wtg10/papers.html`.

[61] P. Varnavides, "On certain sets of positive density", *J. London Math. Soc.* **34** (1959), 358–360.

[62] E. Croot, *A structure theorem for positive density sets having the minimal number of 3-term arithmetic progressions*, `arXiv:math.NT/0305318`.

[63] A. Samorodnitsky and L. Trevisan, *Gowers uniformity, influence of variables, and PCPs*, `arXiv:math.CO/0510264`.

[64] G. A. Freiman, *Foundations of a structural theory of set addition*, Kazan. Gosudarstv. Ped. Inst.; Elabuzh. Gosudarstv. Ped. Inst., Kazan' 1966 (Russian); English transl., Transl. Math. Monographs, vol. 37, Amer. Math. Soc., Providence, RI 1973.

[65] Y. Bilu, "Structure of sets with small sumset", *Structure theory of sets addition*, Astérisque, vol. 258 1999, pp. 77–108.

[66] I. Ruzsa, "Generalized arithmetic progressions and sumsets", *Acta Math. Hungar.* **65**:4 (1994), 379–388.

[67] M.-C. Chang, "A polynomial bound in Freiman's theorem", *Duke Math. J.* **113**:3 (2002), 399–419.

[68] S. L. G. Choi, "On arithmetic progressions in sequences", *J. London Math. Soc.* (2) **10**:4 (1975), 427–430.

[69] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*. Tome III. *Invariants intégraux. Solutions périodiques du deuxième genre. Solutions doublement asymptotiques*, Gauthier–Villars, Paris 1899; Reprint, Dover, New York 1957 (French); English transl., *New methods of celestial mechanics*. vol. III: *Integral invariants, periodic solutions of the second type, doubly asymptotic solutions*, NASA TT F-452, Nat. Aeronaut. Space Admin., Washington, DC 1967; Revised reprint of the 1967 English transl., Hist. Modern Phys. Astronom., vol. 13, Amer. Inst. Phys., New York 1993.

[70] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications, vol. 54, Cambridge Univ. Press, Cambridge 1995.

[71] V. Bergelson and A. Leibman, "Set-polynomials and polynomial extension of the Hales–Jewett theorem", *Ann. of Math.* (2) **150**:1 (1999), 33–75.

[72] A. Sárközy, "On difference sets of sequences of integers, I", *Acta Math. Acad. Sci. Hungar.* **31**:1–2 (1978), 125–149.

[73] A. Sárközy, "On difference sets of sequences of integers, III", *Acta Math. Acad. Sci. Hungar.* **31**:3–4 (1978), 355–386.

[74] S. Srinivasan, "On a result of Sárközy and Furstenberg", *Nieuw. Arch. Wisk.* (4) **3**:3 (1985), 275–280.

[75] J. Pintz, W. L. Steiger, and E. Szemerédi, "On sets of natural numbers whose difference set contains no squares", *J. London Math. Soc.* (2) **37**:2 (1988), 219–231.

[76] A. W. Hales and R. I. Jewett, "Regularity and positional games", *Trans. Amer. Math. Soc.* **106**:2 (1963), 222–229.

[77] R. McCutcheon, *Elemental methods in ergodic Ramsey theory*, Lecture Notes in Math., vol. 1722, Springer-Verlag, Berlin 1999.

[78] M. D. Boshernitzan, "Quantitative recurrence results", *Invent. Math.* **113**:3 (1993), 617–631.

[79] N. G. Moshchevitin, "On a theorem of Poincaré", *Uspekhi Mat. Nauk* **53**:1 (1998), 223–224 (Russian); English transl., *Russian Math. Surveys* **53**:1 (1998), 219–220.

[80] I. D. Shkredov, "Recurrence in mean", *Mat. Zametki* **72**:4 (2002), 625–632 (Russian); English transl., *Math. Notes* **72**:4 (2002), 576–582.

[81] I. D. Shkredov, *On multiple recurrence*, `arXiv: math.DS/0406413`.

[82] V. Afraimovich, J. R. Chazottes, and B. Saussol, "Pointwise dimensions for Poincaré recurrence associated with maps and special flows", *Discrete Contin. Dyn. Syst.* **9**:2 (2003), 263–280.

[83] L. Barreira, Y. Pesin, and J. Schmeling, "Dimension and product structure of hyperbolic measures", *Ann. of Math.* (2) **149**:3 (1999), 755–783.

[84] L. Barreira and B. Saussol, "Hausdorff dimension of measures via Poincaré recurrence", *Comm. Math. Phys.* **219**:2 (2001), 443–463.

[85] L. Barreira and B. Saussol, "Product structure of Poincaré recurrence", *Ergodic Theory Dynam. Systems* **22**:1 (2002), 33–61.

[86] B. Saussol, S. Troubetzkoy, and S. Vaienti, "Recurrence, dimensions, and Lyapunov exponents", *J. Statist. Phys.* **106**:3–4 (2002), 623–634.

[87] I. D. Shkredov, "Dynamical systems with low recurrence rate", *Mat. Sb.* **197**:11 (2006), 143–158 (Russian); English transl., *Sb. Math.* **197**:11 (2006), 1697–1712.

[88] M. Ajtai and E. Szemerédi, "Sets of lattice points that form no squares", *Studia Sci. Math. Hungar.* **9** (1974/1975), 9–11.

[89] V. H. Vu, "On a question of Gowers", *Ann. Combin.* **6**:2 (2002), 229–233.

[90] J. Solymosi, "Note on a generalization of Roth's theorem", *Discrete and computational geometry*, Algorithms Combin., vol. 25, Springer, Berlin 2003, pp. 825–827.

[91] G. N. Sárközy and S. Selkow, *On a question of Gowers concerning isosceles right-angle triangles*, `http://citeseer.ist.psu.edu/569945.html` 2003.

[92] B. Green, "Finite field models in additive combinatorics", *Surveys in combinatorics* 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge 2005, pp. 1–27.

[93] F. R. K. Chung, R. L. Graham, and R. M. Wilson, "Quasi-random graphs", *Combinatorica* **9**:4 (1989), 345–362.

[94] F. R. K. Chung and R. L. Graham, "Quasi-random subsets of Z_n", *J. Combin. Theory Ser. A* **61**:1 (1992), 64–86.

[95] V. I. Bogachev, *Foundations of measure theory*, vols. 1, 2, Research Center "Regular and chaotic dynamics", Moscow–Izhevsk 2003. (Russian)

[96] L. E. Dickson, *History of the theory of numbers*, vol. III, Carnegie Inst. of Washington, Washington, DC 1919, 1920, 1923.

[97] N. G. Chudakov, "On the density of the set of even numbers which are not representable as a sum of two odd primes", *Izv. Akad. Nauk SSSR Ser. Math.*, 1938, no. 1, 25–40. (Russian)

[98] J. G. van der Corput, "Über Summen von Primzahlen und Primzahlquadraten", *Math. Ann.* **116**:1 (1939), 1–50.

[99] S. Chowla, "There exists an infinity of 3-combinations of primes in A.P.", *Proc. Lahore Philos. Soc.* **6**:2 (1944), 15–16.

[100] P. A. Pritchard, A. Moran, and A. Thyssen, "Twenty-two primes in arithmetic progression", *Math. Comp.* **64**:211 (1995), 1337–1339.

[101] M. Frind, P. Jobling, and P. Underwood, 23 *primes in arithmetic progression*, `http://primes.plentyoffish.com`.

[102] B. Host and B. Kra, "Nonconventional ergodic averages and nilmanifolds", *Ann. of Math.* (2) **161**:1 (2005), 397–488.

[103] D. A. Goldston and C. Y. Yildirim, "Higher correlations of divisor sums related to primes. I: Triple correlations", *Integers* **3** (2003), paper A5.

[104] D. Goldston and C. Y. Yildirim, *Higher correlations of divisor sums related to primes*, III: *k-correlations*, `arXiv: math.NT/0209102`.

[105] D. A. Goldston, Y. Motohashi, J. Pintz, and C. Y. Yildirim, "Small gaps between primes", *Proc. Japan Acad. Ser. A Math. Sci.* **82**:4 (2006), 61–65.

[106] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Clarendon Press, Oxford 1960.

[107] W. Deuber, "Partitions theorems for Abelian groups", *J. Combin. Theory Ser. A* **19**:1 (1975), 95–108.

[108] S. D. Adhikari, "A note on a question of Erdős", *Exposition. Math.* **15**:4 (1997), 367–371.

[109] T. C. Brown and V. Rödl, "Monochromatic solutions to equations with unit fractions", *Bull. Austral. Math. Soc.* **43**:3 (1991), 387–392.

[110] H. Lefmann, "On partition regular systems of equations", *J. Combin. Theory Ser. A* **58**:1 (1991), 35–53.

[111] W. Deuber, "Partitionen und lineare Gleichungssysteme", *Math. Z.* **133**:2 (1973), 109–123.

[112] E. Croot, "Long arithmetic progressions in critical sets", *J. Combin. Theory Ser. A* **113**:1 (2006), 53–66; `arXiv: math.NT/0403082`.

[113] R. C. Baker, G. Harman, and J. Pintz, "The difference between consecutive primes, II", *Proc. London Math. Soc.* (3) **83**:3 (2001), 532–562.

[114] J. Bourgain, "On arithmetic progressions in sums of sets of integers", *A Tribute to Paul Erdős*, Cambridge Univ. Press, Cambridge 1990, pp. 105–109.

[115] I. Z. Ruzsa, "Arithmetic progressions in sumsets", *Acta Arith.* **60**:2 (1991), 191–202.

[116] B. Green, "Arithmetic progressions in sumsets", *Geom. Funct. Anal.* **12**:3 (2002), 584–597.

[117] G. A. Freiman, H. Halberstam, and I. Z. Ruzsa, "Integer sum sets containing long arithmetic progressions", *J. London Math. Soc.* (2) **46**:2 (1992), 193–201.

[118] I. Z. Ruzsa, "Arithmetical progressions and the number of sums", *Period. Math. Hungar.* **25**:1 (1992), 105–111.

[119] A. Sárközy, "Finite addition theorems, I", *J. Number Theory* **32**:1 (1989), 114–130.

[120] A. Sárközy, "Finite addition theorems, II", *J. Number Theory* **48**:2 (1994), 197–218.

[121] A. Sárközy, "Finite addition theorems, III", Groupe de travail en théorie analytique et élémentaire des nombres, 1989–1990, *Publ. Math. Orsay* **92-01** (1992), 105–122.

[122] V. F. Lev, "Optimal representations by sumsets and subset sums", *J. Number Theory* **62**:1 (1997), 127–143.

[123] V. F. Lev, "Blocks and progressions in subset sums sets", *Acta Arith.* **106**:2 (2003), 123–142.

[124] E. Szemerédi and V. H. Vu, "Long arithmetic progressions in sum-sets and the number of x-sum-free sets", *Proc. London Math. Soc.* (3) **90**:2 (2005), 273–296.

[125] E. Szemerédi and V. H. Vu, "Finite and infinite arithmetic progressions in sumsets", *Ann. of Math.* (2) **163**:1 (2006), 1–35.

[126] J. Solymosi, "Arithmetic progressions in sets with small sumsets", *Combin. Probab. Comput.* **15**:4 (2006), 597–603.

[127] P. Erdős and P. Turán, "On a problem of Sidon in additive number theory, and on some related problems", *J. London Math. Soc.* **16** (1941), 212–215.

[128] T. Jiang, "Anti-Ramsey numbers of subdivided graphs", *J. Combin. Theory Ser. B* **85**:2 (2002), 361–366.

[129] V. Jungić, J. Licht, M. Mahdian, J. Nešetřil, and R. Radoičić, "Rainbow arithmetic progressions and anti-Ramsey results", *Combin. Probab. Comput.* **12**:5–6 (2003), 599–620.

[130] V. Jungić and R. Radoičić, "Rainbow 3-term arithmetic progressions", *Integers* **3** (2003), paper A18.

[131] M. Axenovich and D. Fon-Der-Flaass, "On rainbow arithmetic progressions", *Electron. J. Combin.* **11**:1 (2004), Research paper 1.

[132] I. Z. Ruzsa, "Difference sets without squares", *Period. Math. Hungar.* **15**:3 (1984), 205–209.

[133] V. Bergelson and R. McCutcheon, *An ergodic IP polynomial Szemerédi theorem*, Mem. Amer. Math. Soc., vol. 146, no. 695, Amer. Math. Soc., Providence, RI 2000.

I. D. Shkredov
Moscow State University
E-mail: ishkredov@rambler.ru

Received 27/MAR/06
Translated by A. I. SHTERN