ISRAEL JOURNAL OF MATHEMATICS **174** (2009), 1–18
DOI: 10.1007/s11856-009-0100-3

SUMSETS IN DIFFERENCE SETS

BY

VITALY BERGELSON*

Department of Mathematics, Ohio State University
Columbus, Ohio, 43210, USA
e-mail: vitaly@math.ohio-state.edu

AND

IMRE Z. RUZSA**

Alfréd Rényi Institute of Mathematics
Budapest, Pf. 127, H-1364, Hungary
e-mail: ruzsa@renyi.hu

ABSTRACT

We study some properties of sets of differences of dense sets in $\mathbb{Z}^2$ and $\mathbb{Z}^3$ and their interplay with Bohr neighbourhoods in $\mathbb{Z}$. We obtain, inter alia, the following results.

(i) If $E \subset \mathbb{Z}^2$, $\overline{d}(E) > 0$ and $p_i, q_i \in \mathbb{Z}[x]$, $i = 1, \dots, m$ satisfy $p_i(0) = q_i(0) = 0$, then there exists $B \subset \mathbb{Z}$ such that $\overline{d}(B) > 0$ and

$$E - E \supset \bigcup_{i=1}^{m} \big(p_i(B) \times q_i(B)\big).$$

(ii) If $A \subset \mathbb{Z}$ with $\overline{d}(A) > 0$, then for any r, s, t such that $r + s + t = 0$ the set $rA + sA + tA$ is a Bohr neighbourhood of 0.

(iii) For any $0 < \alpha < 1/2$ there exists a set $E \subset \mathbb{Z}^3$ with $\overline{d}(E) > 0$ such that $E - E$ does not contain a set of the form $B \times B \times B$, where $B \subset \mathbb{Z}$ and $\overline{d}(B) > 0$.

* First author was supported by NSF grant DMS-0600042.
** Second author was supported by Hungarian National Foundation for Scientific Research (OTKA), Grants No. K 61908, T 42750, T 43623.
Received October 20, 2007

1. Introduction

In this paper, we consider some additive properties of dense sets of integers and lattice points in dimensions 2 and 3.

We define the **upper asymptotic density** of a set $E \subset \mathbb{Z}^d$ by the formula

$$\overline{d}(E) = \limsup_{n\to\infty} \frac{|E \cap [-n,n]^d|}{(2n+1)^d}.$$

By taking the lower limit instead we obtain the concept of **lower density** $\underline{d}(E)$, and if $\overline{d}(E) = \underline{d}(E)$, we call this value the **asymptotic density** $d(E)$.

We define the **upper Banach density** by

$$d^*(E) = \lim_{n\to\infty} \max_{t\in\mathbb{Z}^d} \frac{|(E-t) \cap [1,n]^d|}{n^d}.$$

The first author proved the following results ([1, Corollaries 3.1.1 and 3.1.2]).

STATEMENT 1.1: *Let $E \subset \mathbb{Z}^2$ and suppose that $d^*(E) > 0$. Then there exists $B \subset \mathbb{Z}$ such that $\overline{d}(B) > 0$ and*

$$E - E \supset B \times B.$$

STATEMENT 1.2: *Let $A \subset \mathbb{Z}$ and suppose that $d^*(A) > 0$. Then there exists $B \subset \mathbb{Z}$ such that $\overline{d}(B) > 0$ and*

$$A - A \supset B + B.$$

In the same paper the following questions are raised.

– Given a set $E \subset \mathbb{Z}^3$ with $\overline{d}(E) > 0$, can one find $B \subset \mathbb{Z}$, $\overline{d}(B) > 0$ such that

$$E - E \supset B \times B \times B? \tag{1.1}$$

– Given a set $A \subset \mathbb{Z}$ with $\overline{d}(A) > 0$, can one find $B \subset \mathbb{Z}$, $\overline{d}(B) > 0$ such that

$$A - A \supset B + B + B? \tag{1.2}$$

In this paper, we answer the first problem in the negative, present some results related to the second problem and improve upon Statements 1.1 and 1.2 in different directions. These improvements are as follows.

THEOREM 1.3: *Let $A \subset \mathbb{Z}$ and suppose that $d^*(A) > 0$. Then there exists $B \subset \mathbb{Z}$ such that $B = -B$, $0 \in B$, B has asymptotic density, $d(B) > 0$ and*

$$A - A \supset B + B.$$

THEOREM 1.4: *Let $E \subset \mathbb{Z}^2$ and suppose that $\overline{d}(E) > 0$. Let $p_i, q_i \in \mathbb{Z}[x]$, $i = 1, \ldots, m$ satisfy $p_i(0) = q_i(0) = 0$ for all i. Then there exists $B \subset \mathbb{Z}$ such that $\overline{d}(B) > 0$ and*

$$E - E \supset \bigcup_{i=1}^{m} \big(p_i(B) \times q_i(B)\big).$$

Here for a polynomial p and a set $B \subseteq Z$, we write $p(B) = \{p(n) : n \in B\}$.

The proofs, based on some results in ergodic theory, are given in Sections 2–3.

Concerning the first question we show the following.

THEOREM 1.5: *For every $0 < \alpha < 1/2$ there is a set $E \subset \mathbb{Z}^3$ with $d(E) > \alpha$ such that there is no $B \subset \mathbb{Z}$, $\overline{d}(B) > 0$ satisfying* (1.1).

The results concerning the second question will be explained in Section 4 and proved in Sections 5–8.

2. Two summands with a density

In this section, we prove Theorem 1.3. It will be derived from the following result about dynamical systems.

THEOREM 2.1: *Let $(X, \mathcal{B}, \mu, T)$ be a probability space with a measure-preserving transformation. For every $Y \in \mathcal{B}$ with $\mu(Y) > 0$ there exists a sequence $B \subset \mathbb{N}$ of positive density such that*

$$\mu\big(Y \cap T^{b_1}Y \cap T^{-b_1}Y \cap \cdots \cap T^{b_k}Y \cap T^{-b_k}Y\big) > 0 \tag{2.1}$$

for every $b_1, \ldots, b_k \in B$.

For the proof we need the following result of Bourgain [6].

LEMMA 2.2: *Let T be an ergodic measure-preserving transformation on a probability space, f, g bounded measurable functions. The sequence*

$$N^{-1} \sum_{n=1}^{N} f(T_1^n x) g(T_2^n x),$$

where T_1, T_2 are powers of T, converges almost everywhere.

Proof of Theorem 2.1. Removing, if necessary, a subset of measure 0 from Y we may assume that every set of the form

$$Y \cap T^{n_1}Y \cap T^{-n_1}Y \cap \cdots \cap T^{n_k}Y \cap T^{-n_k}Y \tag{2.2}$$

is either empty, or has positive measure.

From Bourgain's aforementioned theorem we know that

$$f(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} I_Y(x) I_{T^{-n}Y}(x) I_{T^nY}(x)$$

exists almost everywhere. Clearly $f(x) \ge 0$. We are going to show that that $J = \int f\, d\mu > 0$.

By boundedness we can exchange limit and integration. Hence

$$J = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int I_Y I_{T^{-n}Y} I_{T^nY}\, d\mu = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int I_Y I_{T^{-n}Y} I_{T^{-2n}Y}\, d\mu$$

and the positivity is implied by the ergodic Roth Theorem (see Theorem 4.27 in [8], also Section 4.2 in [2]).

Take an x_0 such that $f(x_0) > 0$. Then the sequence

$$B = \big\{b : x_0 \in Y \cap T^{-b}Y \cap T^bY\big\}$$

has positive density (exactly $f(x_0)$). Thus every set of type (2.2) with $n_i \in B$ is non-empty (contains x_0), consequently of positive measure by the above assumption. ∎

To deduce Theorem 1.3, we will use a variant of Furstenberg's correspondence principle. For a proof of the particular version that we are giving here see Bergelson and McCutcheon [4], Proposition 7.2. See also Furstenberg [8, p. 152].

LEMMA 2.3: *Let $E \subset \mathbb{Z}^r$ be a set satisfying $d^*(E) > 0$. Then there exists a probability measure preserving system $(X, \mathcal{B}, \mu, \{T^{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{Z}^r})$ and a set Y with $\mu(Y) > 0$ such that for all $k \in \mathbb{N}$ and $\mathbf{n}_1, \dots, \mathbf{n}_k \in \mathbb{Z}^r$ one has*

$$d^*\big(E \cap (E - \mathbf{n}_1) \cap \cdots \cap (E - \mathbf{n}_k)\big) \ge \mu\big(Y \cap T^{\mathbf{n}_1}Y \cap \cdots \cap T^{\mathbf{n}_k}Y\big). \tag{2.3}$$

Proof of Theorem 1.3. Apply Lemma 2.3 with $r = 1$ and A in the place of E. Then apply Theorem 2.1 for this system; let B_0 be the set obtained. Our set will be

$$B = B_0 \cup (-B_0) \cup \{0\}.$$

Inequalities (2.1) and (2.3) together mean that for every finite $B' \subset B$ we have

$$d^*\{a \in A : a + B' \subset A\} > 0.$$

In particular, for $b_1, b_2 \in B$ we find (lots of) $a \in A$ such that $a + b_1 = a_1 \in A$ and $a - b_2 = a_2 \in A$, whence $b_1 + b_2 = a_1 - a_2 \in A - A$. ∎

With some modifications in the proof one can establish the following slightly more general result.

THEOREM 2.4: *Let $A \subset \mathbb{Z}$ and suppose that $\overline{d}(E) > 0$. Let r, s be given non-zero integers. Then there exists $B \subset \mathbb{Z}$ such that $B = -B$, B has asymptotic density, $d(B) > 0$ and*

$$A - A \supset rB + sB.$$

Here we write

$$rB = \{rb : b \in B\}.$$

In the proof we apply Bourgain's theorem for $T_1 = T^r, T_2 = T^{-s}$ rather than T and T^{-1}.

3. Polynomials

In this section we prove Theorem 1.4. It will be a consequence of the following result.

THEOREM 3.1: *Let $E \subset \mathbb{Z}^2$ and suppose that $\overline{d}(E) > 0$. Let $p_i, q_i \in \mathbb{Z}[x]$, $i = 1, \dots, m$ satisfy $p_i(0) = q_i(0) = 0$ for all i. Then there exists $B \subset \mathbb{Z}$ such that $\overline{d}(B) > 0$ and*

$$\overline{d}\Big(\bigcap_{i=1}^{m} \bigcap_{j=1}^{n} \big(E - \big(p_i(b_j), q_i(b_j)\big)\big) \Big) > 0 \tag{3.1}$$

for every $b_1, \dots, b_n \in B$.

COROLLARY 3.2: *Under the same assumptions we have*

$$E - E \supset \bigcup_{i=1}^{m} \big(p_i(B) \times q_i(B)\big).$$

To prove the Corollary, we apply the previous theorem with a system of polynomials containing the original ones and identically 0 polynomials.

Proof of Theorem 3.1. We use Furstenberg's correspondence principle (Lemma 2.3) for $r = 2$ to find a "model" $(X, \mathcal{B}, \mu, T, U)$, where T, U are commuting measure-preserving transformations on X, and a set $Y \subset X$ with $\mu(Y) = d^*(E)$ satisfying

$$d^*\Big(\bigcap_{i=1}^{k}(E-(m_i,n_i))\Big) \ge \mu\Big(\bigcap_{i=1}^{k} T^{m_i}U^{n_i}Y\Big)$$

for any $m_1, \dots, m_k, n_1, \dots, n_k \in \mathbb{Z}$.

Like in the previous section, we may assume that any intersection of sets of the form T^mU^nY is either empty, or has positive measure.

Put

$$f_N(x) = \frac{1}{N}\sum_{n=0}^{N-1} I_Y\left(T^{p_1(n)}U^{q_1(n)}x\right)\cdots I_Y\left(T^{p_m(n)}U^{q_m(n)}x\right).$$

By the polynomial Szemerédi theorem proved by Bergelson and Leibman ([3, Theorem A]) we know that

$$\limsup_{N\to\infty}\int f_N\,d\mu = \limsup_{N\to\infty}\frac{1}{N}\sum_{n<N}\mu\Big(\bigcap_{i=1}^{m}T^{p_i(n)}U^{q_i(n)}Y\Big) = c > 0.$$

Clearly $0 \le f(x) \le 1$ for all $x \in X$. Let $f(x) = \limsup_{N\to\infty} f(x)$. By Fatou's lemma we have

$$\int f\,d\mu \ge \int \limsup f_N\,d\mu \ge \limsup\int f_N\,d\mu \ge c > 0.$$

Hence for some x_0 we have

$$f(x_0) = \limsup\frac{1}{N}\sum_{n=0}^{N-1} I_Y\left(T^{p_1(n)}U^{q_1(n)}x\right)\cdots I_Y\left(T^{p_m(n)}U^{q_m(n)}x\right) \ge c > 0.$$

This implies that the set

$$B = \left\{b : x_0 \in T^{p_1(b)}U^{q_1(b)}Y\right\}$$

has positive upper density, even

$$\overline{d}\Big(\bigcap_{i=1}^{m}\bigcap_{j=1}^{n}\Big(E-\big(p_i(b_j),q_i(b_j)\big)\Big)\Big) \ge \mu\Big(\bigcap_{i=1}^{m}\bigcap_{j=1}^{n}T^{p_1(b_j)}U^{q_1(b_j)}Y\Big) > 0 \tag{3.2}$$

whenever $b_1, \dots, b_n \in B$. ∎

4. Differences and triple sums

To formulate our results concerning the second question we introduce certain parametric statements (that may hold for some values of the parameters and fail for others). We will consider the more general inclusion

$$A - A \supset rB + sB + tB \tag{4.1}$$

with $r, s, t \in \mathbb{Z}$. We will consider three variants, the density version, the effective density version and the finite version.

Density version. For integers r, s, t and $\alpha \in (0, 1)$, $D(r, s, t, \alpha)$ means that for every set $A \subset \mathbb{Z}$ with $\overline{d}(A) = \alpha$ one can find $B \subset \mathbb{Z}$, $\overline{d}(B) > 0$ satisfying (4.1).

Effective density version. For integers r, s, t and $\alpha, \beta \in (0, 1)$, $E(r, s, t, \alpha, \beta)$ means that for every set $A \subset \mathbb{Z}$ with $\overline{d}(A) > \alpha$ one can find $B \subset \mathbb{Z}$, $\overline{d}(B) > \beta$ satisfying (4.1).

Finite version. For integers r, s, t, real $\alpha, \beta \in (0, 1)$, and positive integer n, $F(r, s, t, \alpha, \beta, n)$ means that for every set $A \subset \{1, 2, \ldots, n\}$ with $|A| > \alpha n$ one can find $B \subset \mathbb{Z}$, $|B| > \beta n$ satisfying (4.1).

Concerning the (most interesting) density case we have only a partial answer.

Theorem 4.1: *Let r, s, t be non-zero integers such that $r + s + t = 0$ and let $\alpha \in (0, 1/2)$. The statement $D(r, s, t, \alpha)$ is false.*

The proof will consist of showing that $rB + sB + tB$ is a Bohr neighbourhood of 0 (the main result of this part), while $A - A$ may not be one. The proof is given in Sections 4–6.

If $r + s + t \neq 0$, then $rB + sB + tB$ may not be a Bohr neighbourhood of 0 for obvious reasons. For certain triplets, namely when $r + s = 0$, one can show the weaker property that it is a Bohr neighbourhood of some integer. If we could establish the existence of an A such that $A - A$ is nowhere dense in the Bohr topology, this would disprove the density version for some further values. We do not even have a conditional argument for the case $r = s = t = 1$, though the results below suggest a negative answer.

Theorem 4.2: *Let r, s, t be non-zero integers, $\alpha \in (0, 1/2)$ and $\beta \in (0, 1)$. The statement $E(r, s, t, \alpha, \beta)$ is false.*

THEOREM 4.3: *Let r, s, t be non-zero integers, $\alpha \in (0, 1/2)$ and $\beta \in (0, 1)$. There is an $n_0 = n_0(r, s, t, \alpha, \beta)$ such that $F(r, s, t, \alpha, \beta, n)$ is false for $n > n_0$.*

Observe that the bound $1/2$ in these theorems is best possible. For instance, if $\overline{d}(A) > 1/2$, then the sets A and $A + x$ cannot be disjoint, thus $A - A = \mathbb{Z}$. This simple observation does not immediately resolve the case $\alpha = 1/2$. If $\overline{d}(A) = 1/2$, then easy arguments show that there is an integer m such that all integers n not contained in $A - A$ satisfy $n \equiv m \pmod{2m}$. One can see that $E(r, s, t, 1/2, \beta)$ is true with $\beta = 1/\big(2(|r| + |s| + |t|)\big)$; we do not know whether it holds with an absolute constant β. To clarify the transition of behaviour in the finite case around $\alpha = 1/2$ may not be easy (but does not seem to be very important).

These results will be easy consequences of some known results about the length of arithmetical progessions in sumsets. Details are given in Section 8.

5. The number of solutions of a linear equation

In this section we prove an auxiliary result.

LEMMA 5.1: *Let r, s, t be non-zero integers satisfying $a + b + c = 0$. For every positive ε there exists a $\delta > 0$ and an N_0 with the following property. Whenever we take a set $X \subset [1, N]$ of integers such that $|X| \geq \varepsilon N$ and $N > N_0$, there are at least δN^2 triplets of distinct integers $x, y, z \in X$ satisfying the equation $rx + sy + tz = 0$.*

For the particular equation $x + y = 2z$, that is, three integers in an arithmetic progression, this is a result of Varnavides [12]. Our proof essentially follows his argument with small changes.

Proof. Take an integer l with the following property: whenever we take a set $Y \subset [1, l]$ of integers such that $|Y| \geq (\varepsilon/2)l$, the equation $rx + sy + sz = 0$ has at least one solution in distinct integers $x, y, z \in Y$. The existence of such an integer follows immediately from Szemerédi's theorem on arithmetic progressions, or one can adapt any method used to prove Roth's theorem on three-term arithmetic progressions.

First, we show that for every set $Y \subset [1, l]$ of integers the equation $rx + sy + sz = 0$ has at least $|Y| - (\varepsilon/2)l$ solutions in Y. This is an easy induction on $m = |Y|$. For $m < (\varepsilon/2)l$ the claim is empty, for $m = 1 + [(\varepsilon/2)l]$ it is

the assumption. If we know the statement for m, to establish it for $m+1$ we take an $m+1$ element set Y, select a solution x, y, z, and apply the induction hypothesis for the m-th element set $Y\backslash\{x\}$.

Since our equation is invariant under linear transformations, the same inequality applies for any set Y contained in an arithmetic progression of length l.

Next, consider all arithmetic progressions of length l and difference $\leq D$ (D will be specified later) which have at least one common element with X; let them be $P_1, \ldots, P_k$. Since the starting point must lie in the interval $[1-(l-1)D, N]$, there are $< N+lD$ possibilites for it; combined with the D possible differences, we see that

$$k < D(N+lD). \tag{5.1}$$

Write $Y_i = P_i \cap X$. Each Y_i contains at least $|Y_i| - (\varepsilon/2)l$ solutions of the equation; this makes altogether

$$\sum \Big(|Y_i| - \frac{\varepsilon}{2}l\Big) = \sum |Y_i| - \varepsilon kl/2$$

solutions. Here a solution may be counted multiply. A solution (x, y, z) is counted as many times as the number of l-term arithmetic progressions containing it. This multiplicity is less than l^2. Indeed, if we fix that x is the i-th term and y is the j-th, where $1 \leq i < j \leq l$, this determines the progression uniquely. Hence, the total number, say R, of solutions satisfies

$$R \geq \frac{1}{l^2}\bigg(\sum |Y_i| - \frac{\varepsilon kl}{2}\bigg). \tag{5.2}$$

We have

$$\sum |Y_i| = lD|X|,$$

since each element of X is contained in exactly lD arithmetic progressions of the prescribed kind; we can arbitrarily fix the difference $d \leq D$ and the position $1 \leq i \leq l$ of an element in the progression. By substituing this into equation (5.2) we obtain

$$R \geq \frac{1}{l}\big(D|X| - (\varepsilon k)/2\big).$$

On substituting (5.1) we arrive at

$$R \geq \frac{D}{l}\Big(|X| - \frac{\varepsilon}{2}(N+lD)\Big). \tag{5.3}$$

Now put $D = [N/(2l)]$. If $N > 6l$, then this yields $D \geq N/(3l)$ and (5.3) implies

$$R \geq \frac{\varepsilon}{12l^2} N^2,$$

thus the lemma is proved with $N_0 = 6l$ and $\delta = \varepsilon/(12l^2)$. ∎

6. Bohr neighbourhoods in a triple sum

In the **Bohr topology** on $\mathbb{Z}$, a basic neighbourhood of 0 is a set of the form

$$U(u_1, \dots, u_k, \varepsilon) = \{n \in \mathbb{Z} : \|nu_i\| < \varepsilon \text{ for } i = 1, \dots, k.\} \tag{6.1}$$

Here ε is an arbitrary positive number, $u_1, \dots, u_k$ are arbitrary reals and $\|x\|$ denotes the distance of x from the nearest integer. (Neighbourhoods of other integers are defined by translation.)

Here we prove the main result of this part.

THEOREM 6.1: *Let r, s, t be non-zero integers satisfying $r+s+t = 0$. Let B be a set of integers having positive upper Banach density and put $S = rB + sB + tB$. The set S is a Bohr neighbourhood of 0.*

The proof is similar to Bogolyubov's [5] for the analogous statement for the set $A + A - A - A$. Similarly to the proof in [5], we will use exponential sums, and consider first sets of residues, then dense finite sets, finally infinite sets of positive density. The main difference is that the symmetry of the set $A + A - A - A$ makes the exponential sum easy to estimate at 0, and in lack of this symmetry we need some extra arguments; here we will use the lemma from the previous section.

First we consider residues. We use $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ to denote the set of residues modulo m. We cannot define topology here, but we will need sets whose definition is similar to definition (6.1) of a Bohr neighbourhood.

Definition 6.2: A *Bohr k, η-subset* of $\mathbb{Z}_m$ is a set defined by

$$V(v_1, \dots, v_k, \eta) = \left\{ n \in \mathbb{Z}_m : \left\| \frac{nv_i}{m} \right\| < \eta \text{ for } i = 1, \dots, k. \right\}. \tag{6.2}$$

Here η is a positive number and $v_1, \dots, v_k \in \mathbb{Z}_m$.

THEOREM 6.3: *Let r, s, t be non-zero integers satisfying $r + s + t = 0$, and let m be a positive integer satisfying $(m, rst) = 1$. Let $X \subset \mathbb{Z}_m$ be a set satisfying $|X| \geq \varepsilon m$ with some $\varepsilon > 0$. Put $S = rX + sX + tX$. The set S contains a Bohr*

k, η-set with an integer k and a positive η that depend on r, s, t, ε but not on m.

Proof. For $x \in \mathbb{Z}_m$ write

$$f(x) = \sum_{n \in X} e(nx/m),$$

where, as usual, $e(t) = e^{2\pi i t}$. Clearly $n \in S$ holds if and only if

$$R(n) = \frac{1}{m} \sum_{x \in \mathbb{Z}_m} f(rx) f(sx) f(tx) e(-nx/m) > 0, \tag{6.3}$$

and furthermore, $R(n)$ is exactly the number of triplets $x, y, z \in X$ such that $n = rx+sy+tz$. In particular, $R(0)$ counts the triplets satisfying $rx+sy+tz = 0$. By representing each element of X by an integer in $[1, m]$ and applying Lemma 5.1 we see that $R(0) \geq R > \delta m^2$ for $m > N_0(\varepsilon)$. For $m \leq N_0$ we can assert that $R(0) \geq |X|$, as $R(0)$ counts also the trivial solutions. Hence we always have $R(0) \geq \delta' m^2$ with, say, $\delta' = \min(\delta, \varepsilon/N_0)$. Consequently for a general n we have

$$\begin{aligned} R(n) &= R(0) - \frac{1}{m} \sum_{x \in \mathbb{Z}_m} f(rx) f(sx) f(tx) \big(1 - e(-nx/m)\big) \\ &\geq \delta' m^2 - \frac{1}{m} \sum_{x \in \mathbb{Z}_m} f(rx) f(sx) f(tx) |1 - e(-nx/m)|. \end{aligned} \tag{6.4}$$

To estimate this sum, observe first that by Plancherel's formula we have

$$\sum |f(x)|^2 = m|X| \leq m^2, \tag{6.5}$$

hence the Cauchy-Schwarz inequality yields

$$\sum_{x \in \mathbb{Z}_m} |f(sx) f(tx)| \leq \left(\sum |f(sx)|^2 \sum |f(tx)|^2 \right)^{1/2} = \sum |f(x)|^2 \leq m^2.$$

(Here we use the assumption $(s, m) = (t, m) = 1$, though we would not lose much by dropping it.) By comparing this to (6.4) we see that $n \in S$ for every n which has the following property:

$$|f(rx)||1 - e(-nx/m)| < \delta' m \tag{6.6}$$

for all x.

To find such values of n, observe first that $|1 - e(-nx/m)| \leq 2$ always, so (6.6) automatically holds for those values of x for which $|f(rx)| < \delta' m/2$. Now

consider those for which $|f(rx)| \geq \delta' m/2$. By (6.5), the number of such values of x is at most $(2/\delta')^2$; denote them by $u_1, \dots, u_k$. We have a bound $k \leq (2/\delta')^2$ for this number, and this bound is independent of m.

For these values we shall require that

$$|1 - e(nu_j/m)| < \delta'/2.$$

Since $|1 - e(y)| \leq 2\pi\|y\|$ for every real y, it suffices to assume that $\|nu_j/m\| < \eta = \delta'/(4\pi)$ for all $j = 1, \dots, k$. This number η is also independent of m and this concludes the proof. ∎

Our calculations were far from optimal at several places. Since we did not give any estimate for δ, and if we did, it would be of the form $e^{-\varepsilon^{-c}}$ with some constant c, the possible savings, which are powers of ε, would not matter much.

LEMMA 6.4: *Let r, s, t be non-zero integers satisfying $r + s + t = 0$, and let $X \subset [-N, N]$ be a set of integers satisfying $|X| \geq \varepsilon N$ with some $\varepsilon > 0$. Put $S = rX + sX + tX$. There are real numbers $u_1, \dots, u_k$ and $\eta > 0$ such that the set S contains a set of the form*

$$U(u_1, \dots, u_k, \eta) \cap [-N, N]. \tag{6.7}$$

Here k and η depend on r, s, t, δ only.

Proof. Take an integer m satisfying

$$m > (|r| + |s| + |t| + 1)N, \quad (m, rst) = 1.$$

We can find such an m below $(|r| + |s| + |t| + 1)N + |rst|$, so we have $|X| > \varepsilon' m$ with $\varepsilon' = \varepsilon/(|r| + |s| + |t| + 1 + |rst|)$. We apply the previous theorem to this set X, now regarded as a set of residues modulo m, with ε' in the place of ε. We get a k and an η, and residues $v_1, \dots, v_k$ such that for every integer n satisfying $\|nv_i/m\| < \eta$ for all i there are $x, y, z \in X$ such that

$$n \equiv rx + sy + tz \pmod{m}.$$

If $|n| < N$, then $|n - (rx + sy + tz)| < m$, thus the congruence becomes equality. Thus we proved the theorem with $u_i = v_i/m$. ∎

Proof of Theorem 6.1. As $d^*(A) > 0$, there is a constant $\varepsilon > 0$ and a sequence of integers z_N such that the sets

$$X_N = (A - z_N) \cap [1, N]$$

satisfy $|X_N| > \varepsilon N$. An application of the preceeding lemma yields the existence of real numbers $u_1^{(i)}, \dots, u_k^{(i)}$ and an $\eta > 0$ such that every integer n satisfying

$$|n| \le N, \quad \|nu_j^{(i)}\| < \eta \quad \text{for } j = 1, \dots, k$$

belongs to S. (Here we use $r+s+t=0$, so that the translations by z_N cancel.) By periodicity of the fractional part we can assume that $u_j^{(i)} \in [0,1]$ for all i, j.

Now define $W_N \subset [0,1]^k$ as the set of vectors $(u_1, \dots, u_k)$ which have the following property: every integer n satisfying $|n| \le N$ and $\|nu_i\| \le \eta/2$ for $i = 1, \dots, k$ belongs to S. This defines a closed set. Clearly $W_N \supset W_{N'}$ for $N > N'$, so we have a decreasing sequence of closed sets. We also know that $W_{N_i} \neq \emptyset$, hence

$$W = \bigcap_{N=1}^{\infty} W_N \neq \emptyset.$$

Let $(u_1, \dots, u_k)$ be any element of W. Then we have

$$B(u_1, \dots, u_k, \eta/2) \subset S$$

as wanted. ∎

7. Proof of Theorems 4.1 and 1.5.

Proof of Theorem 4.1. Theorem 4.1 asserts, for any given $\alpha < 1/2$, the existence of a set A such that $\overline{d}(A) > \alpha$ and $A - A$ contains no subset of the form $S = rB + sB + tB$ with $\overline{d}(B) > 0$. We will find actually a set A with $d(A) > \alpha$ (though this is not really stronger, see Section 8).

By Theorem 6.1 we know that S is a Bohr neighbourhood of 0. Thus it is sufficient to find a set A for which $A - A$ is not. The existence of such a set is implied by the following theorem of Kříž ([9, Theorem 3.1]).

LEMMA 7.1: *For every $\varepsilon > 0$ there is a shift-invariant graph on $\mathbb{Z}$ with chromatic number ∞ and containing an independent set of density $> 1/2 - \varepsilon$.*

We denote this set by A; thus we can achieve $d(A) > \alpha$. The assumptions that this is an independent set and the graph is shift-invariant mean that two integers x, y are not connected if $x - y \in A - A$. Hence the property that the chromatic number is infinite yields that there is no partition of the integers into finitely many subsets, say $\mathbb{Z} = Z_1 \cup \dots \cup Z_l$ such that $Z_i - Z_i \subset A - A$ for all i.

This implies that $A - A$ is not a Bohr neighbourhood of 0. Indeed, if it were, say we had

$$A - A \supset U = U(u_1, \dots, u_k, \eta),$$

then we could find a partition (Z_i) in the following way. Cover the unit cube $[0,1]^k$ by cubes of side $< \varepsilon$, say $T_1, \dots, T_l$. Define Z_i by

$$Z_i = \{n : (\{u_1 n\}, \{u_2 n\}, \dots, \{u_k n\}) \in T_i\}.$$

Clearly $\bigcup Z_i = \mathbb{Z}$ and $Z_i - Z_i \subset U$ for all i. ∎

Proof of Theorem 1.5. Theorem 1.5 asserts the existence of a set $E \subset \mathbb{Z}^3$ with $\overline{d}(E) > \alpha$ such that $E - E$ contains no subset of the form $B \times B \times B$, $\overline{d}(B) > 0$. We will actually find a set with $d(E) > \alpha$.

To see this, take a set A of Theorem 4.1, with any choice of r, s, t, say $r = s = 1$, $t = -2$. Define

$$E = \{(x, y, z) : rx + sy + tz \in A\}.$$

It is easy to see that $d(E) = d(A)$, and the inclusion $E - E \supset B \times B \times B$ would immediately yield $A - A \supset rB + sB + tB$, which we have excluded. ∎

8. The effective and finite versions

In this section we prove Theorems 4.2 and 4.3.

The proof will be based on two results concerning arithmetic progressions in sumsets. The first one is from Ruzsa [11].

LEMMA 8.1: *Let ε be a positive number. For every prime $p > p_0(\varepsilon)$ there is a symmetric set X of residues* mod p *such that $|X| > (1/2 - \varepsilon)p$ and $X + X$ contains no arithmetical progression of length*

$$\exp(\log p)^{2/3+\varepsilon}. \tag{8.1}$$

The other is a result of Freiman, Halberstam and Ruzsa [7]. We quote a (slightly weakened) version of Theorem 3, with some change in the notation.

LEMMA 8.2: *Let p be a prime, Z a set of residues modulo p, $|Z| \geq \gamma p$ with $0 < \gamma < 1$. The set $Z + Z + Z$ contains an arithmetic progression of length at least p^δ, with δ depending only on γ.*

(The theorem gives an explicit value of δ and an estimate for the number of progressions, which we do not need.)

First we deduce a slight generalization.

LEMMA 8.3: *Let p be a prime, Y_1, Y_2, Y_3 sets of residues modulo p, $|Y_i| \geq \beta p$ with $0 < \beta < 1$. The set $S = Y_1 + Y_2 + Y_3$ contains an arithmetic progression of length at least p^δ, with $\delta > 0$ depending only on γ .*

Proof. By a standard averaging argument we find $x, y \in \mathbb{Z}_p$ such that $Z = Y_1 \cap (Y_2 + x) \cap (Y_3 + y)$ has at least $\beta^3 p$ elements. We apply the previous lemma to this set with $\gamma = \beta^3$. The arithmetic progression found in this way gives our arithmetic progression by a shift, since clearly $Y_1 + Y_2 + Y_3 \supset Z + Z + Z - (x + y)$. ∎

The proof of [7] could also be modified to directly handle the case of different summands.

We now give a modular analogue to Theorems 4.2 and 4.3.

LEMMA 8.4: *Let r, s, t be non-zero integers, $\alpha \in (0, 1/2)$ and $\beta \in (0, 1)$. There is a $p_1 = p_1(r, s, t, \alpha, \beta)$ such that for every prime $p > p_1$ there is a set $X \subset \mathbb{Z}_p$ such that $|X| > \beta p$ and there is no $Y \subset \mathbb{Z}_p$ satisfying $|Y| > \beta p$, $rY + sY + tY \subset X - X$.*

Proof. If $p > \max(|r|, |s|, |t|)$, then $Y_1 = rY$, $Y_2 = sY$, $Y_3 = tY$ all have $> \beta p$ elements. Thus $S = rY + sY + tY$ contains an arithmetical progression of length p^δ with some $\delta = \delta(\beta)$. If $p > p_1(1/2 - \alpha)$, then we can find X such that the length of any arithmetic progression in $X + X = X - X$ is less than the quantity in (8.1). So this set X is good as soon as p_1 is large enough to guarantee

$$\exp(\log p)^{2/3+\varepsilon} < p^\delta$$

for $p > p_1$. ∎

Proof of Theorem 4.2. We have to construct, for given r, s, t, α, β, a set A with $\overline{d}(A) > \alpha$ such that $A - A$ does not contain any set of the form $rB + sB + tB$ with $\overline{d}(B) > \beta$.

To this end take a set X produced by the previous lemma and let A be the set of those integers whose residue modulo p lies in X. Clearly $d(A) = |X|/p > \alpha$. Take any set B such that $rB + sB + tB \subset A - A$. Let Y be the set of residues modulo p of elements of B. Then clearly $rY + sY + tY \subset X - X$, thus $|Y| < \beta p$ and hence $\overline{d}(B) \leq |Y|/p < \beta$. ∎

Proof of Theorem 4.3. We have to construct, for given r, s, t, α, β and $n > n_0$, a set $A \subset [1, N]$ with $|A| > \alpha N$ such that $A - A$ does not contain any set of the form $rB + sB + tB$ with $|B| > \beta N$.

To this end take a number $\alpha' \in (\alpha, 1/2)$, and apply the lemma above with α' in the place of α for the largest prime $p < N$. For sufficiently large N we have $\alpha' p > \alpha N$, thus the set X obtained satisfies $|X| > \alpha N$. Let A be the set of those integers in $[1, p]$ whose residue modulo p lies in X. Clearly $|A| = |X| > \alpha N$.

Take any set B such that $rB + sB + tB \subset A - A$. Let Y be the set of residues modulo p of elements of B. Then clearly $rY + sY + tY \subset X - X$, thus $|Y| < \beta p < \beta N$.

Take any two elements of B, say x, y. Since both $rx + sx + tx$ and $ry + sx + tx$ belong to A, their difference is at most $p-1$. Thus $|x-y| \le |rx-ry| \le p-1$, that is, the elements of B are pairwise incongruent modulo p, hence $|B| = |Y| < \beta N$ as desired. ∎

9. Concluding remarks and open problems

1. Density. For uniformity, we formulated our results with upper density. However, as far as difference sets are concerned, it does not matter what concept of density we use. Given a set A with upper Banach density α, one can find another set A' with asymptotic density α, which has the following property: for any finite $F \subset A'$ there is an x such that $F + x \subset A$. In particular, this implies that $A' - A' \subset A - A$. This can be found for the case of sets of positive integers (where the definition of density is modified in the natural way) in Ruzsa [10]; the case of $\mathbb{Z}$ or $\mathbb{Z}^d$ can be handled similarly. It can also be found (for $\mathbb{Z}$) in Furstenberg's book [8], Theorem 3.20 or in [1], Theorem 2.2.

2. Bohr neighbourhoods. The proof in Section 6 worked through deciding whether a sumset of a certain type is or is not a Bohr neighbourhood of 0. The proofs in Section 7 were seemingly different. However, the proof of Lemma 7.2 actually works through finding a shifted Bohr k, η-set in the sumset. A long arithmetic progression is then easily found by Dirichlet's approximation method. So there is a closer connection with the proof in section 6 than the wording shows.

The unsolved cases of the density case are closely connected with the following unsolved question. If $A \subset \mathbb{Z}$, $d(A) > 0$, must $A - A$ be a neighbourhood of some number in the Bohr topology? By Kříž' theorem we know it is not necessarily

a neigbourhood of 0, and 0 is the "most natural difference". We also know that the difference set of a large set in $\mathbb{Z}_p$ may not contain a Bohr k, η-set. These results suggest a negative answer. However, so far we could not find a way to connect the finite and infinite cases.

If the answer to this question is positive, that is, $A - A$ is always a Bohr neigbourhood, then one can easily deduce that it contains sets of the form $rB + sB + tB$ with $d(B) > 0$ for arbitrary prescribed r, s, t. If the answer is negative, we are confident that the answer to the inclusion question is negative as well. We remark that (in the case $r+s+t \neq 0$) the condition $\overline{d}(B) > 0$, or even the stronger condition $\underline{d}(B) > 0$ does not imply that $rB + sB + tB$ is a Bohr neighbourhood; indeed, it may have large gaps, while a Bohr neighbourhood always has bounded gaps. We do not know whether the assumption $d(B) > 0$ suffices.

Acknowledgement. The authors are grateful to Prof. Norbert Hegyvári for directing their attention to the paper of Varnavides and to the referee for some corrections.

References

[1] V. Bergelson, *Sets of recurrence of $\mathbb{Z}^m$-actions and properties of sets of differences in $\mathbb{Z}^m$*, Journal of the London Mathematical Society. Second Series. **31** (1985), 295–304.

[2] V. Bergelson, *Combinatorial and diophantine applications of ergodic theory*, (with appendices by A. Leibman and by A. Quas and M. Wierdl) Handbook of dynamical systems (B. Hasselblatt and A. Katok, eds.), vol. 1B, Elsevier B. V., Amsterdam, 2005, pp. 745–841.

[3] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, Journal of American Mathematical Society **9** (1996), 725–753.

[4] V. Bergelson and R. McCutcheon, *An ergodic IP polynomial Szemerédi theorem*, Memoirs of the American Mathematical Society **146**, (2000), No 695, vii + 106 pp.

[5] N. N. Bogolyubov, *Some algebraical properties of almost periods*, Zapiski Kafedry Matematicheskoy Fiziki Akademii Nauk Ukrainy **4** (1939), 185–194.

[6] J. Bourgain, *Double recurrence and almost sure convergence*, Journal für die Reine und Angewandte Mathematik **404** (1990), 140–161.

[7] G. A. Freiman, H. Halberstam, and I. Z. Ruzsa, *Integer sum sets containing long arithmetic progressions*, Journal of the London Mathematical Society **46** (1992), 193–201.

[8] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, Princeton, 1981.

[9] I. Kříž, *Large independent sets in shift-invariant graphs*, Graphs and Combinatorics **3** (1987), 145–158.

[10] I. Z. Ruzsa, *On difference sets*, Studia Scientiarum Mathematicarum Hungarica **13** (1978), 319–326.
[11] I. Z. Ruzsa, *Arithmetic progressions in sumsets*, Acta Arithmetica **60** (1991), 191–202.
[12] P. L. Varnavides, *On certain sets of positive density*, Journal of the London Mathematics Society **34** (1959), 358–360.