On the structure of sets with many k-term arithmetic progressions

by

GYÖRGY ELEKES (Budapest)

1. Introduction. Extremal problems can be studied at three levels. First you determine the optimum (maximum or minimum) of the quantity in question. Then you characterize the structures for which the optimum is attained. Finally, you may wonder how stable the extremal structures are: not far from the optimum, do the structures remain close to the optimal one? Of course, stability questions of this type are usually much harder than the original extremum problems.

A good example is the number of 3-term arithmetic progressions in a set of n reals. The maximum itself is not difficult to find (probably folklore; see also Section 1.3, Case I) and the optimal structures are close to an n-term AP. However, it is a profound, widely applicable (and, last but not least, not at all easy) result of Additive Number Theory that near the optimum the structure cannot change much [BS94]. Also for $k \geq 4$, similar results hold.

The goal of this paper is to study the structure of n-sets which contain an "asymptotically optimal" number of k-term arithmetic progressions, for k large: $k \to \infty$ (together with n). As we shall see, the larger the size k, the more stable the (nearly) optimal structure.

1.1. *The problem.* For $\mathcal{H} \subset \mathbb{R}$ and $3 \leq k \leq |\mathcal{H}|$, we write

$$\mathcal{AP}(\mathcal{H}, k) := \#\{k\text{-term AP's in } \mathcal{H}\}, \quad \mathcal{AP}(n, k) := \max_{|\mathcal{H}|=n} \mathcal{AP}(\mathcal{H}, k).$$

A fundamental and very useful theorem of Szemerédi and Trotter [ST83] implies the following upper bound:

$$\mathcal{AP}(n, k) \leq C \cdot \frac{n^2}{k}$$

for an absolute constant $C > 0$. The proof can be found, e.g., in [Ele99b] (see

---

[145]

also [Ele02, Theorem 3.13] for a more general result). Moreover, in Corollary 2.3 we also present a simple direct proof.

Can the order of magnitude $|\mathcal{H}|^2/k$ be attained for some $|\mathcal{H}|$? The answer is "yes", as shown by, say, arithmetic progressions of n terms. Indeed, e.g., for $\mathcal{H} = \{0, 1, \ldots, n-1\}$, for any difference $\delta \in \{1, \ldots, (n-1)/(2k)\}$ and every "starting value" $0 \le a \le (n-1)/2$, the k-term AP's $a, a+\delta, \ldots, a+(k-1)\delta$ are all contained in $\mathcal{H}$, yielding a total of $(n-1)^2/(4k)$ such progressions.

It is natural to ask the following.

QUESTION 1.1. *What is the structure of $\mathcal{H}$ if $\mathcal{AP}(\mathcal{H}, k)$ attains the (optimal) order of magnitude $|\mathcal{H}|^2/k$?*

1.2. *The main result.* For technical reasons (to be clarified later) in certain formulas we shall use $k - 1$ instead of just k—which preserves the asymptotic behavior of the functions involved—just as in the denominator in inequality (1) below.

THEOREM 1.2 (Main Theorem). *Let $c_0 > 0$ and $\varepsilon_0 \in (0, 1/3)$ be arbitrary constants. Assume that $\mathcal{H} \subset \mathbb{R}$ with $n := |\mathcal{H}|$ elements and*

$$\text{(1)} \qquad\qquad \mathcal{AP}(\mathcal{H}, k) \ge c_0 \cdot \frac{|\mathcal{H}|^2}{k-1}.$$

Then there exist C_1 AP's $\mathcal{A}_1, \ldots, \mathcal{A}_{C_1}$, each of length at most $C_2 n$, such that "almost every" k-AP in $\mathcal{H}$—precisely: with the exception of $\le \varepsilon_0 \mathcal{AP}(\mathcal{H}, k)$ of them—is completely contained in one of the $\mathcal{A}_i$, provided that k is "not too small": $k \ge C_3$. Here the bounds $C_i = C_i(c_0, \varepsilon_0) > 1$ for $i = 1, 2, 3$ do not depend on $|\mathcal{H}| = n$.

For "very large" values of k we can do even better.

THEOREM 1.3. *If, for a $c_1 > 0$, we consider AP's of length $k \ge c_1 n$ then in the statement of Theorem 1.2 we can even require that each k-term AP (with NO exception!) be contained in one of the C_1 AP's, each of size at most $C_2 n$—provided that $n > n_0(c_0, c_1)$.*

REMARK 1.4. Such "total covers" cannot be expected for sublinear values of k (i.e. if $n/k \to \infty$). Indeed, if half of $\mathcal{H}$ form an arithmetic progression, plus we just "drop at random" $n/(2k)$ AP's of length k then the latter cannot be covered with a bounded number of AP's (not even with generalized AP's).

1.3. *Previous results.* The structure of sets with many 3-term or longer arithmetic progressions, as well as (more generally) of those which contain many similar copies of a fixed pattern, has been studied for a long time; see e.g. [BS94], [LR96], [EE94], [AEFM04], [ER06], and [Ele02] for an overview. These results are closely related to some "inverse theorems"

of Additive Number Theory and small sumsets; see Freiman [Fre66, Fre73], Ruzsa [Ruz92, Ruz94], Bilu [Bil99].

Certain structures called "generalized arithmetic progressions" play a central role in these investigations.

DEFINITION 1.5. Let d and $n_1, \ldots, n_d$ be positive integers and $\Delta_1, \ldots, \Delta_d$ arbitrary real or complex numbers (or vectors). A set G is a *generalized arithmetic progression* ("generalized AP" or "gAP" for short) of dimension d and size $n = n_1 \cdot \ldots \cdot n_d$ if

$$G = \Big\{ \sum_{i=1}^{d} k_i \cdot \Delta_i \, ; \, 0 \le k_i < n_i \text{ for } i = 1, \ldots, d \Big\},$$

and these elements are all distinct.

In what follows, $\mathcal{G}_n^d$ will denote the class of generalized AP's of dimension *not exceeding* d and size *at most* n. The most interesting property of these structures is that, if their dimension d is constant while their size n is large ($\to \infty$), then their *sum-set* $G + G := \{g' + g'' \, ; \, g', g'' \in G\}$ is not too much larger than G itself: $|G + G| \le 2^d |G|$.

As for the structure of sets with $\mathcal{AP}(\mathcal{H}, k)$ large—and for some generalizations—the following has been shown:

CASE I: *k fixed.* For $k = 3$, the exact value $\mathcal{AP}(n, 3) = \lfloor (n-1)^2/4 \rfloor$ is not difficult to demonstrate. On the one hand, the ith element of $\mathcal{H}$ can be the middle element of at most $i - 1$ or $n - i$ three-term AP's, depending on whether $i \le n/2$ or $i \ge n/2$, respectively. Moreover, this upper bound can really be attained, e.g., for $\mathcal{H}$ an AP or an AP with one or both of the outermost elements moved one step further, as in 13457 or 123457.

Balog and Szemerédi [BS94] considered the structure of sets $\mathcal{H}$ which contain a near-optimal number of 3-AP's. They proved that if $\mathcal{AP}(\mathcal{H}, 3) \ge c|\mathcal{H}|^2$ then a large portion of $\mathcal{H}$ is contained in a not-too-large gAP.

For similar copies of more general patterns, i.e. not just AP's but a wide class of point sets B fixed while $|\mathcal{H}| \to \infty$, it was shown in [AEFM04] that if $\mathcal{H}$ contains $\ge c|\mathcal{H}|^2$ similar (i.e. magnified/shrunk and possibly shifted) copies of B then, again, a large part of $\mathcal{H}$ is contained in a suitable gAP.

Recently it was shown in [ER06] that if $\mathcal{AP}(\mathcal{H}, 3) \ge c|\mathcal{H}|^2$ then at most C gAP's, each of size $\le C|\mathcal{H}|$, contain "almost all" 3-AP's; more precisely, at most $\varepsilon|\mathcal{H}|^2$ are missing, provided that $n > n_0(c, \varepsilon)$—see also Theorem 2.7. (Here $C = C(c, \varepsilon)$ does not depend on $|\mathcal{H}|$.)

CASE II: *"medium" values of k.* For $k \to \infty$ but $k/|\mathcal{H}| \to 0$ (while $|\mathcal{H}| \to \infty$), to our best knowledge, no structure result has been found so far; our Theorem 1.2 is the first such example.

CASE III: *"large" values of k.* For $k \geq c|\mathcal{H}|$ (where $c > 0$ is fixed while $|\mathcal{H}| \to \infty$), our Theorem 1.3 will follow easily (in Section 6) from a structure result that characterizes point sets with many similar subsets of size k [Ele99a].

1.4. *The structure of the paper.* The road which leads to the proof of Main Theorem 1.2 consists of four parts, whose details will be given in Sections 2–5, respectively. At the beginning of Sections 2–4, rather technical "statements" can be found that sum up what we want to achieve during that section. On the one hand, we do not call them "theorems" since they are not really interesting in their own right while, on the other hand, they are far more complicated than just "lemmata"; actually, their proofs usually require lots of (sub?)lemmata.

The four main ideas (or rather goals to achieve) can be described as follows.

(a) first we find $\sim k$ small gAP's, each of size $\sim n/k$ such that almost every k-term AP in $\mathcal{H}$ will have $\sim k$ neighboring pairs of elements, each pair contained in one of the small gAP's (different pairs in different gAP's)—see Statement 2.1;

(b) from these gAP's we build a constant number of larger gAP's, each of size $\sim n$, such that a.e. k-term AP in $\mathcal{H}$ has $\sim k$ neighboring pairs of elements, all within one of the new gAP's—see Statement 3.1;

(c) then these gAP's are "blown up" so that a.e. k-term AP in $\mathcal{H}$ be completely "swallowed" (i.e. contained) in one of them—see Statement 4.1;

(d) finally, to prove the Main Theorem 1.2, we again blow up the gAP's even further so that a.e. k-term AP in $\mathcal{H}$ be contained in one single "fiber"—a simple (non-generalized) arithmetic sub-progression—of one of them.

2. Finding (small) gAP's. To begin with, we find some gAP's that, together, cover not just many elements of $\mathcal{H}$, but also a positive proportion of almost all k-AP's in question (the "good" ones will be called "regular"). To make these gAP's really useful in the forthcoming steps, we make sure that they are not too independent of each other, e.g., they, together, only have few distinct differences.

STATEMENT 2.1. *Let $c_0 > 0$ and $\varepsilon_1 \in (0, 1/3)$ be arbitrary constants and assume that, just as in Theorem* 1.2, $\mathcal{AP}(\mathcal{H}, k) \geq c_0 \cdot |\mathcal{H}|^2/(k-1)$ *for some* $\mathcal{H} \subset \mathbb{R}$ *with* $n := |\mathcal{H}|$ *and* $k \geq C_3$. *Then we can select a subset* $\mathcal{R} \subseteq \{k\text{-}AP\text{'s}$ *in* $\mathcal{H}\}$—*which we shall call the set of* regular *AP's—and* $t = C_5 k$ *gAP's*

$$\mathcal{G}_1, \ldots, \mathcal{G}_{C_5 k} \in \text{gAP}^d_{C_6 n/k}$$

(*which may not be distinct*) *with the following properties*:

(i) *at most* $(\varepsilon_1/2) \cdot c_0 n^2/(k-1)$ *of the original k-AP's are irregular*, *i.e.*

$$|\mathcal{R}| \geq (1 - \varepsilon_1/2) \cdot \mathcal{AP}(\mathcal{H}, k);$$

(ii) *for every regular* $A \in \mathcal{R}$ *of difference, say,* δ, *there exist at least* $c_6 k$ *consecutive pairs* $a' \in A$, $a'' = a' + \delta \in A$, *each pair contained in one of the* $\mathcal{G}_i$, *with different pairs in different* $\mathcal{G}_i$;

(iii) *consequently, the total number of* (*distinct*) *differences of the regular AP's is at most* $C_7 n/k$;

(iv) *if* $x_1, y_1 \in \mathcal{G}_j$ (*for some* $j \leq t$) *occur as consecutive elements of a regular* $A_1 \in \mathcal{R}$, *then there exist at least* $m = c_8 n/k$ *elements* $y_1, \ldots, y_m$ (*including the original* y_1) *such that* x_1 *and* y_i ($i = 1, \ldots, m$) *are consecutive elements of some* $A_i \in \mathcal{R}$,

where $d = d(c_0, \varepsilon_1)$ *and all* $c_i = c_i(c_0, \varepsilon_1) > 0$ *and* $C_i = C_i(c_0, \varepsilon_1) > 1$ *are independent of* $|\mathcal{H}| = n$.

2.1. *Preparations for the proof of Statement 2.1.* The main idea is based upon a method of J. Solymosi [unpublished]. We first define a graph (multiple edges allowed) on $\mathcal{H}$ as vertex set as follows: for each k-AP $A \subset \mathcal{H}$ we draw $k - 1$ edges that connect the consecutive pairs of A. We denote this edge set by $E = E(\mathcal{H}, k)$. Since a fixed pair $h', h'' \in \mathcal{H}$ can only be neighboring in at most $k - 1$ distinct k-AP's,

(2)   each edge in E has multiplicity $< k$.

It is also useful to imagine a simple graph (with no multiple edges) on the larger vertex set $\mathcal{H} \times \{0, 1, \ldots, k-1\}$: for each k-AP $\{a, a + \delta, \ldots, a + (k-1)\delta\} \subset \mathcal{H}$ we draw the graph of the linear function $y = (x - a)/\delta$ (which maps A onto the set $\{0, 1, \ldots, k-1\}$) and connect the neighboring pairs of points along this line—note that a similar idea already appears in [Szé97]. Our original edge set E can be thought of as the projection of all these "non-horizontal" edges to $\mathcal{H}$.

This way to each edge there corresponds a unique k-AP; moreover, it also justifies the usage of expressions like "the straight line which contains edge $e \in E$". However, the graph that we shall really use will be the one defined on $\mathcal{H}$ (see Figure 1).

Put $n := |\mathcal{H}|$. Now (1), which appears again as an assumption of Statement 2.1, is equivalent to

$$|E(\mathcal{H}, k)| \geq c_0 n^2.$$

Here comes the essence of Solymosi's method. First we pick a sufficiently small $\varepsilon \in (0, 1)$ (to be specified later) and cut $\mathcal{H}$ into εk consecutive parts $\mathcal{H}_1, \ldots, \mathcal{H}_{\varepsilon k}$, each with $n/(\varepsilon k)$ elements (see Figure 1, which also indicates
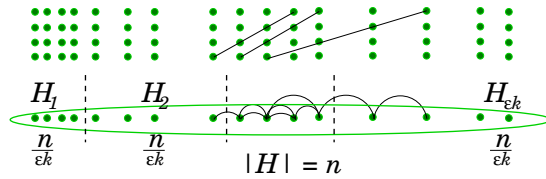
Fig. 1. Some 4-term arithmetic progressions and the corresponding edges

that the cutting lines may not be equidistant). Then we delete from E those edges which are cut (at most εk edges in each k-AP), thus defining a subgraph with edge set

$$E_\varepsilon = E_\varepsilon(\mathcal{H}, k) := \{\text{uncut edges of } E(\mathcal{H}, k)\}.$$

This graph (and some of its subgraphs below) will work as our main tool(s) for proving Statement 2.1. Technically, we shall usually count some of its edges, instead of counting AP's. Formally, too, we shall typically state our results/bounds in terms of edge sets—but sometimes we also translate the statements to the language of k-AP's or of straight lines, if doing so is worth the effort.

Actually, we shall (slightly) modify Solymosi's method: instead of just considering edges, we shall mostly be interested in *pairs of consecutive edges* (which represent three-term AP's). That is why, in each k-AP (on each line), we divide the edges of the original E into $(k-1)/2$ consecutive pairs and even delete those whose *pair* is cut by one of the εk cutting lines, thus defining yet another subgraph $E_\varepsilon^*(\mathcal{H}, k)$.

LEMMA 2.2.

(i) $|E_\varepsilon(\mathcal{H}, k)| \geq (1 - \varepsilon) \cdot (k - 1) \cdot \mathcal{AP}(\mathcal{H}, k)$;
(ii) $|E_\varepsilon^*(\mathcal{H}, k)| \geq (1 - 2\varepsilon) \cdot (k - 1) \cdot \mathcal{AP}(\mathcal{H}, k)$;
(iii) *each of the εk $\mathcal{H}_i$ defined by the cutting, and having $m = n/(\varepsilon k)$ points, contains at most*

$$k \cdot \binom{m}{2} < \frac{1}{2\varepsilon^2} \cdot \frac{n^2}{k}$$

edges of E_ε (counted with multiplicity $< k$), yielding a total of

$$|E_\varepsilon(\mathcal{H}, k)| \leq \frac{1}{2\varepsilon} \cdot n^2. \ \blacksquare$$

As a consequence, the upper bound mentioned in the introduction is not difficult to directly establish (with multiplicative constant $C = 2$—and $k-1$ in place of k).

COROLLARY 2.3.

$$\mathcal{AP}(\mathcal{H}, k) \leq 2 \cdot \frac{|\mathcal{H}|^2}{k - 1}.$$

Proof. By (i) and (iii) of the foregoing lemma, applied to $\varepsilon = 1/2$, we have

$$\frac{1}{2} \cdot (k-1) \cdot \mathcal{AP}(\mathcal{H}, k) \leq |E_{1/2}| \leq |\mathcal{H}|^2,$$

whence the required inequality. ∎

REMARK 2.4. If we apply a theorem of Szemerédi and Trotter to the straight lines and the points of the Cartesian product in Figure 1, then we get another proof of the foregoing upper bound, via incidence geometry.

2.2. *"Good edges" and small subsets.* In an average k-AP, are the edges concentrated, or rather spread nearly evenly among the $\mathcal{H}_i$? As we shall see immediately, concentrated edges are quite rare.

DEFINITION 2.5. For a "badness" $B > 0$ (to be specified later), we call an edge $e \in E_\varepsilon^*$ *B-bad* if the k-AP that defines it (the line which passes through it) contains at least B edges within the $\mathcal{H}_i$ that contains e.

LEMMA 2.6. *At most*

$$\frac{1}{B} \cdot \frac{2n^2}{\varepsilon}$$

edges of E_ε^* are B-bad, which does not exceed $\varepsilon c_0 n^2$ if $B \geq 2/(\varepsilon^2 c_0)$.

Proof. So far we have made (and, in later steps, will again make) good use of the εk-cut. However, just for the purposes of the present proof, we slightly alter it, defining a refinement by cutting the εk parts $\mathcal{H}_i$ of $\mathcal{H}$ (of size $n/(\varepsilon k)$ each) into $B/2$ smaller parts of size $2n/(B\varepsilon k)$. These define a total of $B\varepsilon k/2$ smaller pieces. Some simple observations follow:

— on each line, in each $\mathcal{H}_i$, we have #{newly cut edges} $\leq B/2$, whence

(3) $$\frac{1}{2} \cdot \#\{B\text{-bad edges}\} \leq \#\{\text{still uncut edges}\};$$

— the inequality (3) also applies to *all* edges of any fixed line;
— the same inequality applies even to all edges of E_ε^*.

Thus, to prove Lemma 2.6, it suffices to bound (from above) the quantity on the right hand side of (3):

$$\#\{\text{still uncut edges}\} \leq k \cdot \frac{B\varepsilon k}{2} \cdot \binom{\frac{2n}{B\varepsilon k}}{2} < k \cdot \frac{B\varepsilon k^2}{2} \cdot \frac{1}{2} \cdot \left(\frac{2n}{B\varepsilon k}\right)^2$$
$$= \frac{1}{B} \cdot \frac{1}{\varepsilon} \cdot n^2,$$

which, together with (3), yields the required inequality. ∎

Now we delete all $B = 2/(\varepsilon^2 c_0)$-bad edges (they appear in pairs) and call the resulting edge set $E_\varepsilon^{\mathrm{good}}(\mathcal{H}, k)$. Thus we have

$$(4) \qquad |E_\varepsilon^{\mathrm{good}}(\mathcal{H}, k)| \geq (1 - 3\varepsilon) \cdot (k - 1) \cdot \mathcal{AP}(\mathcal{H}, k),$$

by Lemmata 2.2(ii) and 2.6.

2.3. *Finding the gAP's.* In order to make use of the many three-term AP's determined by pairs of consecutive edges of $E_\varepsilon^{\mathrm{good}}$, we define a new graph, again on $\mathcal{H}$ as vertex set, by connecting two-step neighbors. More precisely, for each pair of consecutive edges

$$(a_\gamma, a_{\gamma+1}) \in E_\varepsilon^{\mathrm{good}} \quad \text{and} \quad (a_{\gamma+1}, a_{\gamma+2}) \in E_\varepsilon^{\mathrm{good}}$$

within a k-AP, we connect a_γ to $a_{\gamma+2}$, and denote the resulting edge set by T (short for "two-step edges"). Moreover, we also write $T_i := T|_{\mathcal{H}_i}$ for the subgraphs spanned by the $\mathcal{H}_i$. Note that we consider T and the T_i as simple graphs with no multiple edges.

Of course, for each two-step edge, its two endpoints must lie in a common $\mathcal{H}_i$, and so must their midpoint.

We quote a fact from Additive Number Theory, concerning small sumsets along graphs.

THEOREM 2.7. *Let (V, E) be an arbitrary graph, $|V| = m$. Assume that $V \subset \mathbb{R}$ and, for each edge $(x, y) \in E$, we have*

$$\frac{x + y}{2} \in V.$$

Further, let an $\alpha \in (0, 1)$ be given. Then we can find disjoint subsets of the vertices, say $V_1, \ldots, V_r$, such that $|V_i| \geq \alpha m$ (hence $r \leq 1/\alpha$), together they contain all but αm^2 edges of E, and also some gAP's

$$\mathcal{G}_1, \ldots, \mathcal{G}_r \in \mathrm{gAP}_{Cm}^d$$

such that $V_i \subset \mathcal{G}_i$ for all $i \leq r$. (Here $d = d(\alpha)$ and $C = C(\alpha)$ do not depend on m.) Furthermore, the degree of each vertex in a spanned subgraph $E|_{V_i}$ is at least $\alpha(|V_i| - 1) \geq \alpha(\alpha m - 1) \geq \alpha^2 m/2$ (if $m \geq 2/\alpha$).

Proof. See [ER06, Theorem 5.4] (with $\lambda = 1$ there).

REMARK 2.8. Of course, the statement is only interesting if E contains at least αm^2 edges—since otherwise no V_i is necessary to reduce the number of "leftover" edges below this bound.

For each $i = 1, \ldots, \varepsilon k$, we apply Theorem 2.7 to $V = \mathcal{H}_i$, $E = T_i$, $\alpha = \varepsilon^2 c_0$ and $m = |\mathcal{H}_i| = n/(\varepsilon k)$. We get $r \leq 1/\alpha = 1/(\varepsilon^2 c_0)$ gAP's

$$(5) \qquad \mathcal{G}_{ij} \in \mathrm{gAP}_{Cn/k}^d \quad (j = 1, \ldots, r)$$

which, together, cover all but

$$\alpha|\mathcal{H}_i|^2 = \varepsilon^2 \cdot c_0 \cdot \left(\frac{n}{\varepsilon k}\right)^2 = c_0 \cdot \frac{n^2}{k^2}$$

edges of T_i within $\mathcal{H}_i$. (Recall that there are no multiple edges in T.) Summing for all $i = 1, \ldots, \varepsilon k$, we conclude that the $\mathcal{G}_{ij}$ cover all but $\varepsilon \cdot c_0 n^2 / k$ edges of T. For the edges of $E_\varepsilon^{\text{good}}$ (and taking multiplicity $< k$ into account) we have

(6) $\#\{\text{edges of } E_\varepsilon^{\text{good}} \text{ not covered}\} < 2\varepsilon c_0 n^2 \le 2\varepsilon (k-1)\mathcal{AP}(\mathcal{H}, k).$

Furthermore, each $h \in \mathcal{H}_i$ is connected to

(7) $$\ge \frac{\alpha^2 m}{2} = \frac{\varepsilon^4 c_0^2}{2} \cdot \frac{n}{k}$$

points within the same $\mathcal{G}_{ij}$ by an edge of T_i—or, perhaps, to none.

We sum up what we have achieved in the following lemma.

LEMMA 2.9. *For simplicity of notation, let $\hat{\mathcal{G}}_1, \ldots, \hat{\mathcal{G}}_t$ denote the gAP's $\mathcal{G}_{ij}$ just found; and $\mathcal{E}_\varepsilon^{\mathcal{G}}$ the set of those edges of $E_\varepsilon^{\text{good}}$ which, together with their twin pairs, are covered by one of the $\hat{\mathcal{G}}_s$. Then $\mathcal{E}_\varepsilon^{\mathcal{G}}$ and the $\hat{\mathcal{G}}_s$ $(s \le t)$ have the following properties*:

(i) *the number t of the $\hat{\mathcal{G}}_s$ satisfies*

$$t \le (\varepsilon k) \cdot r = (\varepsilon k) \cdot \frac{1}{\varepsilon^2 c_0} = \frac{1}{\varepsilon c_0} \cdot k;$$

(ii) *their dimension and size are bounded respectively by*

$$d = d(\varepsilon^2 c_0) \quad \text{and} \quad Cn/k \quad (\text{for } C = C(\varepsilon^2 c_0)),$$

by (5);

(iii) *together they cover all but $\varepsilon \cdot c_0 n^2 / k$ edges of T*;

(iv) *consequently, all but*

$$\le 2(k-1) \cdot \varepsilon \cdot \frac{c_0 n^2}{k} < 2\varepsilon \cdot c_0 n^2 \le 2\varepsilon \cdot (k-1)\mathcal{AP}(\mathcal{H}, k)$$

edges of $E_\varepsilon^{\text{good}}$ are preserved in $E_\varepsilon^{\mathcal{G}}$;

(v) *if an $h \in \mathcal{H}$ is incident to at least one edge of T then it is connected to at least $(\varepsilon^4 c_0^2 / 2) \cdot n/k$ other points of the same gAP, by (7).* ∎

2.4. *Regular lines and AP's: the proof of Statement 2.1.* To start with, we fix

$$\varepsilon := \frac{\varepsilon_1}{20}.$$

For simplicity, we just denote by $\mathcal{G}_i$ the gAP's $\hat{\mathcal{G}}_i$ found in the foregoing Lemma 2.9. According to (i) there, their number does not exceed

$$\frac{1}{\varepsilon c_0} \cdot k = \frac{20}{\varepsilon_1 c_0} \cdot k = C_5 k,$$

if we pick $C_5 = 20/(\varepsilon_1 c_0)$. Also their dimension and size will be bounded by $d := d(\varepsilon^2 c_0)$ and $C_6 \cdot n/k = C(\varepsilon^2 c_0) \cdot n/k$, respectively, by Lemma 2.9(ii).

We define the set $\mathcal{R}$ of regular k-AP's (with a slight abuse of notation) in terms of their corresponding straight lines:

$$\mathcal{R} := \{\text{lines with } \geq (k-1)/2 \text{ edges of } E_\varepsilon^{\mathcal{G}}\},$$

where we use the constant factor $1/2$ just for convenience; anything in $(0,1)$ would do. We are left to show that they—together with the foregoing gAP's $\mathcal{G}_i$—satisfy the requirements (i)–(iv) of Statement 2.1.

(i) Originally, every line contained $k-1$ edges of E, as defined in Section 2.1. Our $E_\varepsilon^{\mathcal{G}}$ still contains all but

$$(2\varepsilon + \varepsilon + 2\varepsilon) \cdot c_0 n^2 = 5\varepsilon \cdot c_0 n^2$$

of these edges, by Lemmata 2.2(ii), 2.6 and 2.9(iv). Thus the number of irregular lines/k-AP's (i.e. from which more than $(k-1)/2$ edges are missing) is

$$< \frac{5\varepsilon \cdot c_0 n^2}{(k-1)/2} = 10\varepsilon \cdot c_0 \frac{n^2}{k-1} = \frac{\varepsilon_1}{2} \cdot c_0 \frac{n^2}{k-1},$$

as required.

(ii) follows from the definition of $\mathcal{R}$ and the fact that any such line contains at most $B = 2/(\varepsilon^3 c_0)$ edges from any $\mathcal{H}_i$ and, therefore, from any $\mathcal{G}_i$. Thus we can choose—by $k - 1 \geq k/2$—

$$c_6 = c_6(c_0, \varepsilon_1) := \frac{1}{4B} = \frac{\varepsilon^3 c_0}{8} = \frac{\varepsilon_1^3 c_0}{8 \cdot 10^3}.$$

(iii) is the only part worth describing in detail. We want to show that the total number of distinct differences δ of the regular AP's $A \in \mathcal{R}$ of type $a + i\delta$ $(i = 0, 1, \ldots, k-1)$ is at most $C_7 n/k$.

(a) On the one hand, each such δ occurs in at least $c_6 k$ of the difference-set $\Delta(\mathcal{G}_i)$ by (iii), whence

$$\sum_i |\Delta(\mathcal{G}_i)| \geq (c_6 k) \cdot (\# \text{ of distinct } \delta).$$

(b) On the other hand, $|\Delta(\mathcal{G}_i)| \leq 2^d |\mathcal{G}_i| \leq 2^d C_6 \cdot n/k$ for $i = 1, \ldots, C_5 k$. Hence

$$\sum_{i=1}^{C_5 k} |\Delta(\mathcal{G}_i)| \leq (C_5 k) \cdot \left(2^d C_6 \cdot \frac{n}{k}\right) = (C_5 \cdot 2^d C_6) \cdot n.$$

Putting (a) and (b) together, the number of distinct differences cannot exceed $C_7 \cdot n/k$ (for $C_7 := C_5 \cdot 2^d C_6/c_6$), as required.

(iv) is obvious by Lemma 2.9(v), for

$$c_8 = c_8(c_0, \varepsilon_1) := \frac{\varepsilon^4 c_0^2}{2} = \frac{\varepsilon_1^4 c_0^2}{2 \cdot 10^4}.$$

This finishes the proof of Statement 2.1. ∎

3. Creating few large gAP's from small ones

STATEMENT 3.1. *Let every constant, also $\mathcal{H} \subset \mathbb{R}$ and $n := |\mathcal{H}|$, and even the set $\mathcal{R}$ of regular k-AP's be as in Statement 2.1. Then we can find few larger, but not too large gAP's (in place of the many small ones): there exist C_{10} gAP's*

$$\mathcal{G}_1', \ldots, \mathcal{G}_{C_{10}}' \in \mathrm{gAP}_{C_{11}n}^{d+1},$$

such that, for every regular $A \in \mathcal{R}$, at least one $\mathcal{G}_i'$ contains $c_7 k$ consecutive pairs from A. Here, again, $c_7 = c_7(c_0, \varepsilon_1) > 0$ and all $C_i = C_i(c_0, \varepsilon_1) > 1$ are independent of $|\mathcal{H}| = n$.

Proof. Of course, we first use Statement 2.1 to find some $\mathcal{G}_1, \ldots, \mathcal{G}_{C_5 k} \in \mathrm{gAP}_{C_6 n/k}^d$. Then we define certain "expansions" and "kernels" of them, which are the subject of Section 3.1. The proof continues after these definitions, in Section 3.2.

3.1. *Expansions and kernels of gAP's*

DEFINITION 3.2. *Let $\mathcal{G}$ be a gAP of dimension d, differences $\delta_1, \ldots, \delta_d$ and size $n_1 \times \cdots \times n_d$. For a positive integer r we call*

$$\left\{ \sum_{j=1}^d i_j \cdot (r \cdot \delta_j) \, ; \, \forall j = 1, \ldots, d \, |i_j| \le n_j/r \right\}$$

the r-expansion of $\mathcal{G}$.

WARNING: this expansion is not larger than $\mathcal{G}$ itself; rather, it is just a subset of it (see Figure 2).
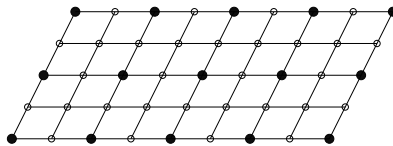

Fig. 2. A gAP and its 2-expansion (solid bullets)

PROPOSITION 3.3. *$(2r-1)^d$ shifted copies of the r-expansion cover $\mathcal{G}$.* ∎

DEFINITION 3.4. For any gAP $\mathcal{G}$ of dimension d, we also define the
c-kernel $\mathcal{G}^{(c)}$ of $\mathcal{G}$ to be its $(1/c)!$-expansion.

For any set $S \subset \mathbb{R}$, we put $\Delta(S) := \{s' - s''\,; \; s', s'' \in S\}$ and, inductively,

$$\Delta^{(j)} := \Delta(\Delta^{(j-1)}), \quad \text{where} \quad \Delta^{(1)} = \Delta(S).$$

PROPOSITION 3.5. *Let $C_1, C_2 \geq 1$ be arbitrary constants while d, t and m positive integers. Assume we are given t sets $Y_0, Y_1, \ldots, Y_{t-1} \subset \mathbb{R}$ with $|Y_i| \geq m$ for $0 \leq i < t$ while*

$$\left| \bigcup_{i=0}^{t-1} Y_i \right| \leq C_1 m.$$

Assume, moreover, that each Y_i is contained in a not-too-large gAP, i.e. there exist $\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{t-1}$ such that

$$Y_i \subset \mathcal{G}_i \in \mathrm{gAP}_{C_2 m}^d \quad \text{for } i = 1, \ldots, t-1.$$

Then there exist $1 + t' \geq 1 + t/(8C_1^2)$ of the Y_i—for simplicity of notation we may assume without loss of generality that after an appropriate permutation of the indices they are $Y_0, Y_1, \ldots, Y_{t'}$—for which

$$(8) \qquad \bigcup_{i=0}^{t'} \mathcal{G}_i^{(c^*)} \subseteq \Delta^{(C^*)}(Y_0).$$

Consequently, if Y_0 intersects one of the selected $\mathcal{G}_i$ $(i = 1, \ldots, t')$, say $a \in Y_0 \cap \mathcal{G}_i$, then for the coset $a + \mathcal{G}_i^{(c^*)}$ we have

$$a + \mathcal{G}_i^{(c^*)} \subseteq \Delta^{(C^*+2)}(Y_0)$$

for, e.g., $C^* = C^*(C_1, C_2, d) := 40C_1 C_2 + d$ and $c^* = c^*(C_1, C_2) := 1/(4C_1 C_2)$ —both independent of t and m. (Of course, the size of the $\mathcal{G}_i^{(c^*)}$ will depend on d too.)

Proof. (8) is Theorem 3.7 in [EHR09]; the rest is an obvious consequence. ∎

3.2. *Promising systems of cosets and the rest of the proof of Statement 3.1.* Put $\hat{C} := C_7/c_5$ and let $C^* = C^*(\hat{C}, d)$ be as in Proposition 3.5.

DEFINITION 3.6. $\mathcal{S} = \{\mathcal{C}_1, \ldots, \mathcal{C}_s\}$ is a *promising system of cosets* if

(i) the $\mathcal{C}_j$ are $(3\hat{C})!$-cosets of distinct $\mathcal{G}_1, \ldots, \mathcal{G}_s$, respectively;
(ii)

$$s \geq \frac{c_5^2 c_6}{12 C_7^2} \cdot k;$$

(iii) there exists a regular AP $A = \{a, a + \delta, \ldots, a + (k-1)\delta\} \in \mathcal{R}$ and indices $0 \le i_1 < \cdots < i_s \le k-1$ such that, for $1 \le j \le s$,

$$a + i_j \delta \in \mathcal{C}_j,$$
$$a + (i_j + 1)\delta \in \mathcal{C}_j,$$
$$(a + i_j \delta, a + (i_j + 1)\delta) \in \mathcal{E}_\varepsilon^{\mathcal{G}};$$

(iv) there is a $0 \le j \le s$ such that

$$(9) \qquad \bigcup_{i=1}^{s} \mathcal{C}_i \subseteq \mathcal{G}'_j := \Delta^{(C^*+2)}(\mathcal{G}_j) + \{-(k-1)\delta, \ldots, -\delta, 0, \delta, (k-1)\delta\}.$$

LEMMA 3.7. *Consider a maximal number of disjoint promising systems* $\mathcal{S}_1, \ldots, \mathcal{S}_r$ *of cosets—where "disjoint" does NOT mean that the COSETS* $\mathcal{C}_j$ *in the* S_i *are disjoint, rather, the SET–SYSTEMS* $\mathcal{S}_i$ *are required to be disjoint, i.e. no two share a common coset. Then*

(i) $r \le C_{10} = C_{10}(c, \varepsilon_1)$;
(ii) *the cosets which occur in the* $\mathcal{S}_i$, *together, cover at least* $(c_6/2)k$ *points of each regular* $A \in \mathcal{R}$;
(iii) *consequently, so do the* $\mathcal{G}'$ *defined in* (9).

Proof. (i) Altogether there are $\le (2 \cdot (3\hat{C})!)^d \cdot C_5 k$ cosets, while each promising system consists of $(c_5^2 c_6)/(12 C_7) \cdot k$ of them, whence

$$r \le \frac{(2 \cdot (3\hat{C})!)^d \cdot C_5 k}{\frac{c_5^2 c_6}{12 C_7} \cdot k} = \frac{12 C_5 C_7 \cdot (2 \cdot (3\hat{C})!)^d}{c_5^2 c_6} =: C_{10}.$$

(ii) Assume for a contradiction that fewer than $(c_6/2)k$ points of a regular AP $A = \{a, a + \delta, \ldots, a + (k-1)\delta\} \in \mathcal{R}$ are covered by the cosets in the $\mathcal{S}_j$. By Statement 2.1(iii), at least $c_6 k$ of the gAP's $\mathcal{G}_j$ contain two consecutive elements (distinct pairs!) of A. At least half of these pairs are NOT intersected by any of the cosets in the $\mathcal{S}_j$, say—without loss of generality, after an appropriate permutation of the indices—

$$x_1 = a + i_1 \delta \in \mathcal{G}_1 \quad \text{and} \quad y_1 = a + (i_1 + 1)\delta \in \mathcal{G}_1,$$
$$\vdots$$
$$x_t = a + i_t \delta \in \mathcal{G}_t \quad \text{and} \quad y_t = a + (i_t + 1)\delta \in \mathcal{G}_t,$$

where $t \ge (c_6/2)k$. Consequently, by Statement 2.1(ii) and for each $j \le t$, there exist

$$m := c_5 \frac{n}{k}$$

elements $y_1, \ldots, y_m \in \mathcal{G}_j$ such that, for all $i = 1, \ldots, m$, the difference $y_i - x_j$ occurs as a difference of some regular $A \in \mathcal{R}$.

Put

$$Y_j := \{y_1 - x_j, y_2 - x_j, \ldots, y_m - x_j\} \quad \text{for } j = 1, \ldots, t,$$

where the y_i, of course, depend on j too. According to Statement 2.1(iv),

$$\left| \bigcup_{j=1}^{t} Y_j \right| \le C_7 \cdot \frac{n}{k} = \frac{C_7}{c_5} \cdot c_5 \frac{n}{k} = \frac{C_7}{c_5} \cdot m = \hat{C} \cdot m.$$

Thus we can use Proposition 3.5 and get a $\mathcal{G}_{j_0}$ and a $C^* = C^*(\hat{C}, d)$ such that $\Delta^{(C^*)}(\mathcal{G}_{j_0})$ covers the $(3\hat{C})!$-kernel of at least

$$t' \ge \frac{t}{6\hat{C}^2} = \frac{c_6}{2} \cdot k \cdot \frac{c_5^2}{6C_7^2} = \frac{c_5^2 c_6}{12 C_7^2} \cdot k$$

other $\mathcal{G}_j$. Hence $\Delta^{(C^*+2)}(\mathcal{G}_{j_0})$ covers the cosets which contain the corresponding x_j. These cosets will, therefore, give yet another promising system of cosets, disjoint from the $\mathcal{S}_j$—which gives the required contradiction.

Finally, (iii) is obvious from (ii). $\blacksquare$

Now the rest of the proof of Statement 3.1 is easy. Consider a maximal promising system of cosets. The corresponding $\mathcal{G}'_1, \ldots, \mathcal{G}'_{C_{10}}$—as in (iv) of the definition of such a system—have all the required properties, by the foregoing lemma. $\blacksquare$

4. Covering with gAP's

STATEMENT 4.1. *In Statement 3.1 we can even find (still at most C_{10}) slightly larger gAP's*

$$\mathcal{G}''_1, \ldots, \mathcal{G}''_{C_{10}} \in \mathrm{gAP}^{d+1}_{C_{12}n},$$

and require that every regular $A \in \mathcal{R}$ be completely contained in one of them.

We need the following.

PROPOSITION 4.2. *Let $c \in (0, 1)$ and m be a positive integer. Assume that $S \subset \{0, 1, \ldots, m\}$ with*

$$(10) \qquad\qquad |S| \ge c(m + 1) \ge 2.$$

If, moreover, we also assume that the original S contains two consecutive elements, then $\Delta^{(\hat{C})}(S) \supseteq \{0, 1, \ldots, m\}$ for, e.g., $\hat{C} = \hat{C}(c) := 10 \log(1/c)$.

Proof. See [EHR09, Lemma 3.1(i)].

Now the proof of Statement 4.1 is easy: apply Proposition 4.2 to $m = k - 1$, an arbitrary regular AP, say $\mathcal{A}$, in place of $\{0, 1, \ldots, m\}$, and $c = c_7$ and $S := \mathcal{A} \cap \mathcal{G}'_i$, where $\mathcal{G}'_i$ was found in Statement 3.1. $\blacksquare$

5. Transforming gAP's into AP's

5.1. *Bijective ("unique") representations of gAP's*

LEMMA 5.1. *Let $\mathcal{G} \in \text{gAP}^d$ be represented by the d-dimensional parallelogram lattice $\mathcal{L}_\mathcal{G}$. If $\Delta^{(2)}(\mathcal{L}_\mathcal{G})$ is a bijective representation of $\Delta^{(2)}(\mathcal{G})$ (and consequently $\mathcal{L}_\mathcal{G}$ is a bijective representation of $\mathcal{G}$) then each k-AP in $\mathcal{G}$ is represented by a k-AP (of d-dimensional vectors) in $\mathcal{L}_\mathcal{G}$.*

Proof. Consider the points $a_1, \ldots, a_k \in \mathcal{G}$ of a k-AP, represented by vectors $\underline{v}_1, \ldots, \underline{v}_k \in \mathcal{L}_\mathcal{G}$, respectively. We want to show that they, too, form a k-AP.

Put $\underline{\delta} := \underline{v}_2 - \underline{v}_1$ and consider (also in $\mathbb{R}^d$) the k-AP

$$
\begin{aligned}
\hat{\underline{v}}_1 &:= \underline{v}_1, \\
\hat{\underline{v}}_2 &:= \hat{\underline{v}}_1 + \underline{\delta} = \underline{v}_2, \\
\hat{\underline{v}}_3 &:= \hat{\underline{v}}_2 + \underline{\delta} \quad \text{(which might not be } \underline{v}_3 \text{ itself)}, \\
\hat{\underline{v}}_4 &:= \hat{\underline{v}}_3 + \underline{\delta}, \\
&\vdots \\
\hat{\underline{v}}_k &:= \hat{\underline{v}}_{k-1} + \underline{\delta}.
\end{aligned}
$$

(11)

This k-AP, again, represents $a_1, \ldots, a_k$—though it might not be contained in $\mathcal{L}_\mathcal{G}$. We show (by induction) that this is not the case: $\hat{\underline{v}}_i = \underline{v}_i$ for all $i \leq k$. Indeed, this holds for $i = 1, 2$ by definition. For the inductive step, if $\hat{\underline{v}}_i = \underline{v}_i$ then

$$
\hat{\underline{v}}_{i+1} = \hat{\underline{v}}_i + \underline{\delta} = \underline{v}_i + (\underline{v}_2 - \underline{v}_1) = \underline{v}_i - (\underline{v}_1 - \underline{v}_2) \in \Delta^{(2)}(\mathcal{L}_\mathcal{G}),
$$

and also

$$
\underline{v}_{i+1} \in \mathcal{L}_\mathcal{G} \subseteq \Delta^{(2)}(\mathcal{L}_\mathcal{G}),
$$

thus $\hat{\underline{v}}_{i+1} = \underline{v}_{i+1}$ since $\Delta^{(2)}(\mathcal{L}_\mathcal{G})$ is a bijective representation of $\Delta^{(2)}(\mathcal{G})$. We conclude that the k-AP defined in (11) does, indeed, work. ∎

LEMMA 5.2. *Let d and k be positive integers, $\alpha > 0$, and $\mathcal{L} \subset \mathbb{R}^d$ an $m_1 \times \cdots \times m_d$ parallelogram lattice. Assume that $\mathcal{L}$ contains at least*

$$
\alpha \cdot \frac{m^2}{k-1}
$$

k-term AP's, for $m = \prod m_i$. Then $\mathcal{L}$ is the union of at most $1/\alpha$ AP's—provided that $k \geq k_0 = k_0(\alpha) := 1 + 4/\alpha$.

Proof. Without loss of generality assume that $m_1 \geq \cdots \geq m_d$. Let m_r be the last $m_i \geq k$, i.e.

$$
m_1 \geq \cdots \geq m_r \geq k > m_{r+1} \geq \cdots \geq m_d.
$$

Note that $r \geq 1$, since otherwise $\mathcal{L}$ contains no k-term AP at all.

The differences $\underline{\delta}$ of the possible k-term AP's must all be of the form $\underline{\delta} = (\delta_1, \ldots, \delta_r, 0, \ldots, 0)$ with $|\delta_i| < (m_i - 1)/(k - 1)$ for $i = 1, \ldots, r$. Moreover, for each $\underline{\delta}$, every element of $\mathcal{L}$ can be the starting value of at most one k-AP (in $\mathcal{L}$) of difference $\underline{\delta}$, which implies the second inequality in the formula below. Assume $r \geq 2$, i.e. $2(r - 1) \geq r$. Then

$$\alpha \cdot \frac{m^2}{k-1} \leq \mathcal{H}(\mathcal{L}, k) \leq m \cdot \#\{\underline{\delta}\} < m \cdot 2^r \cdot \frac{m_1 \cdot \ldots \cdot m_r}{(k-1)^r}$$

$$< \frac{m^2}{k-1} \cdot \left(\frac{4}{k-1}\right)^{r-1} \cdot \frac{1}{m_{r+1} \cdot \ldots \cdot m_d},$$

whence

$$(12) \qquad \left(\frac{k-1}{4}\right)^{r-1} \cdot m_{r+1} \cdot \ldots \cdot m_d < \frac{1}{\alpha}.$$

Now if $k \geq k_0(\alpha) := 1 + 4/\alpha$ then $r - 1 > 0$ is impossible, i.e. $r = 1$. In this case, on the one hand, all k-term AP's in $\mathcal{L}$ must be parallel to the first coordinate direction—since the other dimensions are too short: $m_i < k$. On the other hand, again by (12), $m_2 \cdot \ldots \cdot m_d < 1/\alpha$. We conclude that $\mathcal{L}$ can, indeed, be covered by this few fibers, each an m_1-term AP in the first coordinate direction. ∎

5.2. *Proof of the Main Theorem 1.2.* Assume that, just as in Statement 4.1, every regular AP $A \in \mathcal{R}$ is contained in one of C_{10} gAP's

$$\mathcal{G}_1'', \ldots, \mathcal{G}_{C_{10}}'' \in \text{gAP}_{C_{12}n}^{d+1}.$$

Of course, we would like to use the foregoing two lemmata; however, the representation of the $\mathcal{G}_i''$—let alone that of the $\Delta^{(2)}(\mathcal{G}_i'')$—may not be bijective. The following result will help overcome this difficulty.

THEOREM 5.3. *If $\mathcal{G} \in \text{gAP}^d$ and t is a positive integer then $\mathcal{G}$ is contained in a $\mathcal{G}^+ \in \text{gAP}^d$ of size $|\mathcal{G}^+| \leq C_{13}|\mathcal{G}|$ such that the representation of $\Delta^{(t)}(\mathcal{G}^+)$ is unique (hence so are those of the $\Delta^{(i)}(\mathcal{G}^+)$ for $i \leq t$), where $C_{13} = C_{13}(d, t)$ does not depend on $|\mathcal{G}|$.*

Proof. See Green [Gre05, Theorem 2.1]. The statement is also contained implicitly in Bilu's "strong Freiman theorem" [Bil99] and the details of the proof on p. 82 of Bilu's paper.

Now we apply Theorem 5.3 to $t = 2$ and the $\mathcal{G}_i''$ to get some $\mathcal{G}_i''' := (\mathcal{G}_i'')^+$ for $i \leq C_{10}$ such that the representation of each $\Delta^{(2)}(\mathcal{G}_i''')$ is bijective. Thus we can apply Lemmata 5.1 and 5.2 (the latter for $\alpha = (\varepsilon_1/2)c_0/C_{10}$) to show that at most $(\varepsilon_1/2)c_0 n^2/(k-1)$ of the k-AP's can be covered by those $\mathcal{G}_i'''$ whose dimension is larger than one.

The rest, i.e. $\geq (1 - \varepsilon_1)$-proportion of the k-AP's, are contained in $d = 1$ dimensional, proper arithmetic progressions, provided that k is sufficiently large: $k > 1 + 4/\alpha = 1 + 8C_{10}/(\varepsilon_1 c_0)$. ∎

6. Proof of Theorem 1.3. Let
$$A_1 = \{a_1, a_1 + \delta_1, \ldots, a_1 + (k-1)\delta_1\},$$
$$\vdots$$
$$A_t = \{a_t, a_t + \delta_t, \ldots, a_t + (k-1)\delta_t\}$$
be the $k = c_1 n$-term AP's in $\mathcal{H}$, where $|\mathcal{H}| = n$ and
$$t = \alpha \cdot \frac{|\mathcal{H}|^2}{k} = \frac{\alpha}{c_1} \cdot n.$$
We consider the set of linear functions
$$\Phi := \{\varphi_i : x \mapsto \delta_i x + a_i \,;\, i = 1, \ldots, t\},$$
which map $X := \{0, 1, \ldots, k-1\}$ onto the A_i. The following fact involves gGP's which can be defined quite similarly to gAP's.

PROPOSITION 6.1. *Let $C \geq 1$ be an arbitrary constant, n a positive integer, moreover $X \subset \mathbb{R}$ and Φ a set of non-constant linear functions as above. Assume that $|X|, |\Phi| \geq n$ while*
$$\left| \bigcup_{\varphi \in \Phi} \varphi(X) \right| \leq Cn,$$
where $\varphi(X)$ denotes $\{\varphi(x)\,;\, x \in X\}$. Then there exist $\leq C^*$ reals $s_1, \ldots, s_{C^*}$ such that one of the following two possibilities must hold:

(i) *either one can find a gAP $\mathcal{G} \in \mathrm{gAP}_{C^{**}n}^{d^*}$ for which*
$$X \subseteq \mathcal{G} \quad and \quad \Phi \subseteq \{\varphi : x \mapsto s_j(x + g)\,;\, j = 1, \ldots, t, g \in \mathcal{G}\};$$

(ii) *or one can find an $u \in \mathbb{R}$ and a gGP $\mathcal{G} \in \mathrm{gGP}_{C^{**}n}^{d^*}$ for which*
$$X \subseteq \mathcal{G} + u \quad and \quad \Phi \subseteq \{\varphi : x \mapsto g \cdot (x - u) + s_j\,;\, j = 1, \ldots, t, g \in \mathcal{G}\};$$
where the constants d^*, C^* and C^{**} only depend on C but not on $n = |X|$.

For the proof see [Ele98, Theorem 3]. ∎

To prove Theorem 1.3, we first show that the second (gGP) possibility in Proposition 6.1 leads to a contradiction for $X = \{0, 1, \ldots, k-1\}$. To start with, since $X \subset \mathcal{G} + u$, it is natural to consider $X_0 := X - u \subseteq \mathcal{G}$. Then we have
$$|X_0 \cdot X_0| \leq |\mathcal{G} \cdot \mathcal{G}| \leq 2^{d^*}|\mathcal{G}| \leq 2^{d^*} C^{**} n \leq \frac{2^{d^*} C^{**}}{c_1} |X_0|,$$
while
$$|X_0 + X_0| = 2k - 1 < 2|X_0|$$

is obvious. However, it was shown in [Ele97, Theorem 1] that $c|X_0|^{5/2} \leq |X_0 + X_0| \cdot |X_0 \cdot X_0|$ for an absolute constant $c > 0$. This implies

$$c|X_0|^{5/2} < 2|X_0| \cdot \frac{2^{d^*}C^{**}}{c_1}|X_0| = \frac{2^{d^*+1}C^{**}}{c_1}|X_0|^2,$$

whence

$$c_1^{1/2}n^{1/2} \leq |X_0|^{1/2} \leq \frac{2^{d^*+1}C^{**}}{c_1 c},$$

a contradiction if $n > n_0(c_0, c_1) = 2^{2d^*+2}C^{*2}/(c_1^2 c^2)$.

We are left with the first (gAP) possibility in Proposition 6.1. In this case, for each s_j $(j = 1, \ldots, C^*)$ we consider the gAP

$$\mathcal{G}_j := s_j \cdot (\mathcal{G} + \mathcal{G})$$

whose dimension is at most one more than that of $\mathcal{G}$ and whose size is at most a constant factor larger.

It would be natural to attempt to refer to the proof of the Main Theorem 1.2 in Section 5, but it allows for (not too much but) several exceptional k-term gAP's, which we want to avoid here.

Fortunately, it suffices to just repeatedly use Theorem 5.3 for the $\mathcal{G}_j$, as long as the representation of one of the $\Delta^{(2)}(\mathcal{G}_j)$ is not bijective. As soon as all such representations are unique, i.e. each $\mathcal{G}_j$ is a (bijective) projected image of a lattice $\mathcal{L}_j$, we observe that each k-AP A_i $(i = 1, \ldots, t)$ is completely covered by one of the (bounded number of) "fibers" of one of the $\mathcal{G}_j$, each an AP of size at most linear in n. (There is no "wrap-around-the-fibers" by the unique representation property.) Thus the longest dimension of a lattice, if it contains an A_i, is at least $k = c_1 n$—otherwise no k-AP could fit in it. Since

$$\frac{|\mathcal{L}_j|}{k} \leq \frac{Cn}{c_1 n} = \frac{C}{c_1},$$

we conclude that a bounded number of arithmetic progressions (the set of all fibers of the $\mathcal{G}_j$) cover *all* the A_i. ∎

Concluding remarks. From a number-theoretic point of view we have seen that, in certain situations, usual (i.e. non-generalized) arithmetic progressions can be far superior to generalized ones.

From the point of view of incidence geometry, our results can also be considered as a (small) step towards the notoriously difficult problem of characterizing point-line configurations with (asymptotically) maximal number of incidences; cf. Figure 1 and Remark 2.4.

The following questions remain open:

PROBLEM 1.

(a) *Determine the exact value of $\mathcal{AP}(n, k)$ for $k \geq 4$.*
(b) *Is it true that $\mathcal{AP}(n, 4) = \lceil n(n-3)/6 \rceil$?*

In place of AP's one can also consider arbitrary "patterns" $\mathcal{P} \subset \mathbb{R}$ and denote the number of similar copies of $\mathcal{P}$ in a set $\mathcal{H}$ by $\mathcal{S}(\mathcal{H}, \mathcal{P})$. A result of Szemerédi and Trotter mentioned in Remark 2.4 implies the existence of an absolute constant $C > 0$ such that $\mathcal{S}(\mathcal{H}, \mathcal{P}) \leq C|\mathcal{H}|^2/(|\mathcal{P}| - 1)$. This order of magnitude can be attained e.g. if $\mathcal{P}$ is the positive proportion of an AP (selected e.g. at random or by following an arbitrary "rule").

PROBLEM 2. *Is it true that if* $\mathcal{S}(\mathcal{H}, \mathcal{P}) \geq c|\mathcal{H}|^2/(|\mathcal{P}| - 1)$ *then* $\mathcal{P}$ *is contained in an AP of length* $\leq C|\mathcal{P}|$, *for a* $C = C(c)$—*provided that* $|\mathcal{P}| > k_0 = k_0(c)$?

References

[AEFM04] B. Ábrego, G. Elekes, and S. Fernández-Merchant, *Structural results for planar sets with many similar subsets*, Combinatorica 24 (2004), 541–554.

[BS94] A. Balog and E. Szemerédi, *A statistical theorem of set addition*, ibid. 14 (1994), 263–268.

[Bil99] Y. Bilu, *Structure of sets with small sumset*, Astérisque 258 (1999), 77–108.

[Ele97] G. Elekes, *On the number of sums and products*, Acta Arith. 81 (1997), 365–367.

[Ele98] —, *On linear combinatorics II*, Combinatorica 18 (1998), 13–25.

[Ele99a] —, *On linear combinatorics III*, ibid. 19 (1999), 43–53.

[Ele99b] —, *On the structure of large homothetic subsets*, in: Contemporary Trends in Discrete Mathematics (Štiřín Castle, 1997), DIMACS Ser. Discrete Math. 49, Amer. Math. Soc., 1999, 101–111.

[Ele02] —, *Sums versus products in number theory, algebra and Erdős geometry—a survey*, in: Paul Erdős and his Mathematics II, Bolyai Soc. Math. Stud. 11, 2002, 241–290.

[EE94] G. Elekes and P. Erdős, *Similar configurations and pseudo grids*, in: Intuitive Geometry (Szeged, 1991), Coll. Math. Soc. János Bolyai 63, North-Holland, Amsterdam, 1994, 85–104.

[EHR09] G. Elekes, N. Hegyvári, and I. Z. Ruzsa, *On difference sets in groups and in generalized arithmetic progressions*, in preparation, 2009.

[ER06] G. Elekes and I. Z. Ruzsa, *The structure of sets with few sums along a graph*, J. Combin. Theory Ser. A 113 (2006), 1476–1500.

[Fre66] G. A. Freiman, *Foundations of a Structural Theory of Set Addition*, Kazan Gos. Ped. Inst., Kazan, 1966 (in Russian).

[Fre73] —, *Foundations of a Structural Theory of Set Addition*, Transl. Math. Monogr. 37, Amer. Math. Soc., Providence, RI, 1973.

[Gre05] B. Green, *Notes on progressions and convex geometry*, http://www-math.mit.edu/˜green/convexnotes.pdf, 2005.

[LR96] M. Laczkovich and I. Z. Ruzsa, *The Number of Homothetic Subsets*, Springer, 1996.

[Ruz92] I. Z. Ruzsa, *Arithmetical progressions and the number of sums*, Period. Math. Hungar. 25 (1992), 105–111.

[Ruz94] —, *Generalized arithmetic progressions and sum sets*, Acta Math. Sci. Hungar. 65 (1994), 379–388.

[Szé97]   L. A. Székely, *Crossing numbers and hard Erdős problems in discrete geom-etry*, Combin. Probab. Comput. 6 (1997), 353–358.

[ST83]   E. Szemerédi and W. T. Trotter Jr., *Extremal problems in discrete geometry*, Combinatorica 3 (1983), 381–392.