**COMBINATORICA**
Bolyai Society – Springer-Verlag

COMBINATORICA **18** (1) (1998) 13–25

# ON LINEAR COMBINATORICS II. STRUCTURE THEOREMS VIA ADDITIVE NUMBER THEORY

GYÖRGY ELEKES*

*Received November 28, 1996*

This article is the second one in a series of three. Part I [1] contained concurrency results for sets of linear mappings of $\mathbf{R}$ with few compositions and/or small image sets. Here the fine structure of such sets of mappings will be described in terms of generalized arithmetic and geometric progressions, yielding Freiman–Ruzsa type results for a non-Abelian group.

## 1. Introduction

The goal of this paper (actually the goal of the sequence of articles of which this one is the second member) is to prove structure theorems specific to Euclidean geometry — i.e., some which do not hold true within finite planes. (See the introduction of Part I for a brief historic overview of certain such results.)

Here the concurrency results of part I are utilized and strong structure theorems are deduced — generalizing some results of Freiman and Ruzsa to the non-Abelian group of linear functions with composition as the group operation.

### 1.1. Composition sets

Throughout this paper, $\mathcal{L}$ will denote the set of non-constant linear functions $x \mapsto ax+b$ $(a \neq 0)$, i.e., the non-degenerate affine mappings of $\mathbf{R}$.

Finite sets of linear mappings will be denoted by $\Phi$ or $\Psi$. As usual, $\phi \circ \psi$ will denote the composition $x \mapsto \phi(\psi(x))$. Moreover, we write

$$\Phi \circ \Psi = \{\phi \circ \psi \ ; \ \phi \in \Phi, \psi \in \Psi\}.$$

In what follows, we call $\Phi \circ \Psi$ a *composition set*.

Mathematics Subject Classification (1991): 20F12, 51A25, 52C10

* Research partially supported by HU-NSF grants OTKA T014105 T014302 and T019367

0209–9683/98/$6.00 ©1998 János Bolyai Mathematical Society

In the first part we addressed the following question: *what is the structure of* $\Phi$ *and* $\Psi$ *if* $\Phi\circ\Psi$ *is not too big as compared to* $\Phi$ *and* $\Psi$*; e.g.,* $|\Phi\circ\Psi|\le 10000n$ *while* $|\Phi|=|\Psi|=n$*?* (Or any positive constant $C$ in place of 10000 above; see the introduction of Part I for the motivations to this problem.)

Now we start with showing two types of "small" composition sets, which will play a key role in the sequel.

**Example 1.** Let $n$ be a large integer multiple of 100 and

$$\Phi=\Psi=\{x\mapsto ix+l\ ;\ i=1\dots 100,\ l=1\dots n/100\},$$

i.e., $\Phi$ and $\Psi$ both consist of 100 bundles of parallel lines of slopes $1,\dots,100$. Then

$$|\Phi\circ\Psi|\le\binom{100}{2}\frac{101}{100}n<10000n.$$ ∎

Before turning our attention to the other example, it is worth mentioning that, instead of $1\dots 100$, any fixed positive integers $s_1,s_2,\dots,s_C$ (with $n$ a multiple of $C$) would result in $|\Phi\circ\Psi|\le C^*n$ for some $C^*=C^*(C,s_1,\dots,s_C)$. (E.g., if we put $s=\max s_i$, then $C^*=2C^2s^2$ will do.) Also, $1\dots n/100$ can be substituted by an arbitrary arithmetic progression if appropriate slopes are chosen.

**Example 2.** Let $n$ be a multiple of 10; moreover, $v_1,v_2,\dots,v_{10}$, $w_1,w_2,\dots,w_{10}$ and $u$ arbitrary reals. Put

$$\Phi=\{x\mapsto 2^i(x-u)+v_t\ :\ t=1\dots 10,\ i=1\dots\frac{n}{10}\};\text{ and}$$
$$\Psi=\{x\mapsto 2^j(x-w_t)+u\ :\ t=1\dots 10,\ j=1\dots\frac{n}{10}\}.$$

Then $|\Phi\circ\Psi|\le 2\cdot 10^2(n/10)=20n$. ∎

(Note that here, again, any geometric progression will do in place of $2^i$.)

Further examples and results on more special types of small composition sets can be found in Section 3.

Using commutator graphs and other combinatorial tools, it was proven in part I [1] (see Theorem 1 there) that small composition sets must be similar to the above two examples in the sense that they always contain many functions whose graphs are parallel or concurrent lines. Here in Part II we use that result to deduce more refined structure theorems and show that all such structures must also be relatives of either arithmetic or geometric progressions or, perhaps, some generalized versions of these. The latter are presented in the next section.

## 2. Notions and the main result

### 2.1. Arithmetic and geometric GPs

Generalized arithmetic progressions were introduced by Szemerédi in his famous paper [7].

**Definition 1.** Let $d$ and $n_1, n_2, \ldots, n_d$ be positive integers and $\Delta_1, \Delta_2, \ldots, \Delta_d$ arbitrary real numbers. We shall call the set

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

a *generalized arithmetic progression* of dimension $d$ and parameters $n_i$ and $\Delta_i$. Similarly — instead of differences $\Delta_i$ — with quotients $q_1, q_2, \ldots, q_d$ positive reals, $\mathcal{G}$ is a *generalized geometric progression* if

$$\mathcal{G} = \left\{ \pm \prod_{i=1}^{d} q_i^{k_i} \ ; \ 0 \le k_i < n_i \text{ for } i = 1 \ldots d \right\}.$$

The $\pm$ sign is placed in the definition for convenience; we just wanted to avoid having to play around with quotients of different signs.

We shall use the shorthands "arithmetic GP" and "geometric GP" for the above structures.

**Remark 2.** Note that the sums or products in the above definition may not be distinct, i.e., $|\mathcal{G}| < \prod n_i$ or $|\mathcal{G}| < 2\prod n_i$ are possible, respectively.
In what follows, $\mathcal{G}_{d,n}$ will denote an arithmetic or geometric GP of dimension *not exceeding* $d$ and size *at most* $n$. For short, we shall also use expressions like "there exists an arithmetic or geometric $\mathcal{G}_{d,n}$".

**Proposition 3.** *If $\mathcal{G}$ is an arithmetic or geometric GP of dimension $d$ then $|\mathcal{G} \pm \mathcal{G}| \le 2^d|\mathcal{G}|$ or $|\mathcal{G} \cdot \mathcal{G}|, |\mathcal{G}/\mathcal{G}| \le 2^d|\mathcal{G}|$, respectively.*

This statement is, probably, folklore. It is really obvious if the sums or products in Definition 1 are all distinct. Otherwise, the following simple proof (communicated to us by I. Z. Ruzsa) can be applied.

**Proof.** For each subset $S \subset \{1, 2, \ldots, n\}$, consider the "vertex" $V_S$ of $\mathcal{G}$ defined in case of arithmetic or geometric GPs by $V_S = \sum_{i \in S} (n_i - 1)\Delta_i$ or $V_S = \prod_{i \in S} q_i^{n_i - 1}$, respectively. (As usual, the empty sum is 0 while the empty product is 1.)

Put $V \stackrel{\text{def}}{=} \{ V_S \ ; \ S \subset \{1, 2, \ldots, n\} \}$. Then $\mathcal{G} + \mathcal{G} = \mathcal{G} + V$ and $\mathcal{G} - \mathcal{G} = \mathcal{G} - V$ or $\mathcal{G} \cdot \mathcal{G} = \mathcal{G} \cdot V$ and $\mathcal{G}/\mathcal{G} = \mathcal{G}/V$; moreover, $|V| \le 2^d$. ∎

We also define arithmetic GP-type structures as generalizations of Example 1.

**Definition 4.** Let $\mathcal{G} \subset \mathbf{R}$ be an arithmetic GP, $\Phi, \Psi \subset \mathcal{L}$ and $C$ a positive integer. We say that the pair $(\Phi, \Psi)$ is an arithmetic GP-type structure based upon $\mathcal{G}$ with $C$ slopes if there are non-zero reals $s_1, s_2, \ldots, s_C$ such that

$$\Phi^{-1} \cup \Psi \;\; = \;\; \{x \mapsto s_i x + g \; ; \;\; 1 \leq i \leq C \;\; \text{and } g \in \mathcal{G}\}.$$

**Remark 5.** At first glance it might seem awkward that $\Phi^{-1}$ occurs together with $\Psi$. However, a closer look shows that "$\Phi \circ \Psi$ is small" is equivalent to "$\Psi^{-1} \circ \Phi^{-1}$ is small"; thus $\Phi^{-1}$ and $\Psi$ must play equal roles.

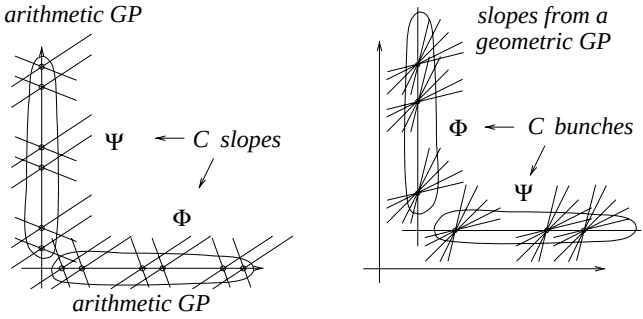


*Figure 1.* $\Phi$ (not $\Phi^{-1}$) and $\Psi$ in an **(a)** arithmetic **(b)** geometric GP-type structure.

Similarly, we define geometric GP-type structures as generalizations of Example 2.

**Definition 6.** Let $\mathcal{G} \subset \mathbf{R}$ be a geometric GP, $\Phi, \Psi \subset \mathcal{L}$ and $C$ a positive integer. We say that the pair $(\Phi, \Psi)$ is a geometric GP-type structure based upon $\mathcal{G}$ with $C$ bunches if there are real numbers $u_1, u_2, \ldots, u_C$ and $v$ such that

$$\Phi^{-1} \cup \Psi \;\; = \;\; \{x \mapsto g(x - u_i) + v \; ; \;\; 1 \leq i \leq C \;\; \text{and } g \in \mathcal{G}\}.$$

It is worth to note that both an arithmetic and a geometric GP-type structure has, indeed, small composition sets — with a constant depending only on $C$ and the dimension of $\mathcal{G}$. The trivial proof (using Proposition 3) is left to the reader.

### 2.2. The Main Theorem

**Theorem 1. (Main Theorem)** *For every $C > 0$ there are $C^* = C^*(C) > 0$, $C^{**} = C^{**}(C) > 0$ and $d^* = d^*(C) > 0$ with the following property. If $\Phi, \Psi \subset \mathcal{L}$ with $|\Phi|, |\Psi| \geq n$ and*

$$|\Phi \circ \Psi| \leq Cn$$

*then $(\Phi,\Psi)$ is contained in an arithmetic or in a geometric GP-type structure with $\le C^*$ slopes or bundles, respectively, based upon an arithmetic or geometric $\mathcal{G}_{d^*,C^{**}n}$.*

The proof can be found in Section 5.

## 3. Special composition sets

The operation of composing two mappings does not commute, i.e., $\phi\circ\psi\neq\psi\circ\phi$ in general. Thus, in a composition set $\Phi\circ\Psi$, changing the role of the operands may result in a dramatic change of size.

**Example 3.** If

$$\Phi=\{x\mapsto 2^t x+1\ :\ t=1\dots n\};\text{ and}$$
$$\Psi=\{x\mapsto 2^t x\ :\ t=1\dots n\},$$

then $|\Phi\circ\Psi|=2n-1$ while $|\Psi\circ\Phi|=n^2$. ∎

That is why, in what follows, we shall also consider more symmetric relatives of the "asymmetric" composition sets $\Phi\circ\Psi$.

### 3.1. Totally symmetric composition sets

We shall call $\Phi\circ\Phi$ a "totally symmetric" composition set and show that if such a set is small then $\Phi$ is really symmetric in some geometric sense.

**Corollary 7.** *For every $C>0$ there are $C^*=C^*(C)>0$ and $d^*=d^*(C)>0$ with the following property. Let $\Phi\subset\mathcal{L}$, $|\Phi|=n$. Assume $|\Phi\circ\Phi|\le Cn$. Then*

(i) *either there are not more than $C^*$ slopes in $\Phi$ and the points of intersection of the $\phi\in\Phi$ with the axes are contained in a horizontal and a vertical copy of an arithmetic $\mathcal{G}_{d^*,C^*n}$;*

(ii) *or the graphs of the $\phi\in\Phi$ — with the exception of $C^*$ of them — all pass through a point $(v,v)$ and the slopes come from a geometric $\mathcal{G}_{d^*,C^*n}$.*

**Proof.** Use Theorem 1. The only detail worth mentioning to Case (ii) is that if both $\Phi$ and $\Phi^{-1}$ have not more than $m$ distinct points of intersection with the horizontal line $y=v$, then at most $(m-1)^2$ of the $\phi\in\Phi$ can avoid the point $(v,v)$. ∎

### 3.2. Slopes and arithmetic GPs

There remains one more question: *how shall the arithmetic GP $\mathcal{G}$ and the slopes $s_i$ be related to each other in order to fulfill condition (i) of the above Corollary?*

One might guess that the slopes (or most of them) must be rational. However, this would be false, as shown by the following example.

**Example 4.** Let $n=2m^2$ and $\mathcal{G}=\{k+l\sqrt{2}\ ;\ 0<k,l\leq m\}$. If

$$\Phi = \{x \mapsto x+g\ ;\ g\in\mathcal{G}\}\ \cup\ \{x\mapsto \sqrt{2}x+g\ ;\ g\in\mathcal{G}\}$$

then $|\Phi|=2m^2=n$, while $|\Phi\circ\Phi|\leq 20m^2=10n$. ∎

Now one might have another guess: the slopes can be arbitrary. This, again, is false, as shown by the following assertion.

**Theorem 2.** *For every $C>0$ and $d\geq 1$ there exists a sequence $M_n=M_n(C,d)$ $(n\in\mathbf{N})$, with*

$$\frac{M_n}{n}\to 0 \text{ (for } n\to\infty),$$

*and with the following property.*
*Let $\Phi\subset\mathcal{L}$ with $|\Phi|=n$. Assume that there are not more than $C$ slopes in $\Phi$ and the points of intersection of the $\phi\in\Phi$ with the axes are contained in a horizontal and a vertical copy of an arithmetic $\mathcal{G}_{d,Cn}$. Then, for any arbitrary transcendental number $\alpha$,*

$$|\{\phi\in\Phi\ ;\ (\text{slope of } \phi)=\alpha\}|\leq M_n.$$

*Moreover, if $d=d(C)$ is a function of $C$, then, of course, $M_n=M_n(C)$ will only depend on $C$.*

The proof is given in Section 5.2.

**Remark 8.** Both Corollary 7 and Theorem 2 can be generalized to symmetric but not totally symmetric composition sets $(\Phi\circ\Psi)\cup(\Psi\circ\Phi)$. The only modification is that in Corollary 7 we will have an arithmetic GP $\mathcal{G}$ on one axis and $s\mathcal{G}$ on the other one, where $s$ (or $-s$) is the most frequent slope while Theorem 2 holds true for slopes which are transcendental multiples of the above $s$.

The reader has seen lots of results of type *"the number of ... is at most $C^*$"*. Such experiences may generate the feeling that even in Theorem 2, such a stronger statement might hold true (i.e., the number of $\phi\in\Phi$ of slope $\alpha$ could be bounded by some $C^*$). This, again, is false. Actually, the above quantity can be very close to $n$, as shown by the following example.

**Example 5.** Let $m$, $d$ be positive integers, $n=m^d$. Define the arithmetic GPs

$$\mathcal{G} = \left\{ \sum_{i=1}^{d} k_i \cdot \pi^i \ ; \ 0 \le k_i < m \text{ for } i = 1 \dots d \right\}; \text{ and}$$

$$\mathcal{G}_2 = \left\{ \sum_{i=2}^{d} k_i \cdot \pi^i \ ; \ 0 \le k_i < m \text{ for } i = 2 \dots d \right\}.$$

Here $|\mathcal{G}|=m^d=n$ and $|\mathcal{G}_2|=m^{d-1}=n^{1-1/d}$. Moreover, note that

$$\mathcal{G}_2 \subset \mathcal{G} \cap \pi\mathcal{G}. \tag{1}$$

Now if

$$\Phi = \{x \mapsto x+g \ ; \ g \in \mathcal{G}\} \ \cup \ \{x \mapsto \pi x + g_2 \ ; \ g_2 \in \mathcal{G}_2\},$$

then, by (1), $|\Phi \circ \Phi| \le 4|\mathcal{G}+\mathcal{G}| \le 2^{d+2}|\mathcal{G}| = 2^{d+2}n$, while the number of $\phi \in \Phi$ of slope $\pi$ is $n^{1-1/d}$. ∎

It may be the subject of further investigations whether or not the number of $\phi \in \Phi$ of slope $\alpha$ can be bigger than $n^{1-1/d}$ if the dimension of $\mathcal{G}$ is a fixed integer $d$ while its size (and, together with it, the size of $\Phi$) goes to infinity.

## 4. Image sets

**Definition 9.** For $H \subset \mathbf{R}$ and $\Phi \subset \mathcal{L}$, we put

$$\Phi(H) = \{\phi(h) \ ; \ \phi \in \Phi, h \in H\}$$

and call it an *image set*.

Image sets, as seen in part I [1], are closely related to composition sets. Thus, it is not surprising that also the structure of small image sets can be described in terms of arithmetic and geometric GPs.

**Definition 10.** Let $\mathcal{G} \subset \mathbf{R}$ be an arithmetic GP and $C$ a positive integer. We say that $\Phi \subset \mathcal{L}$ and $H \subset \mathbf{R}$ is an arithmetic GP-type structure based upon $\mathcal{G}$ with $C$ slopes if there are non-zero reals $s_1, s_2, \dots, s_C$ such that

$$\begin{aligned} H &= \mathcal{G}; \\ \Phi &= \{x \mapsto s_i(x+g) \ ; \ \ 1 \le i \le C \ , \ g \in \mathcal{G}\}; \ \text{ and so} \\ \Phi(H) &= \bigcup_{i=1}^{C} s_i(\mathcal{G}+\mathcal{G}). \end{aligned}$$

**Definition 11.** Similarly, if $\mathcal{G} \subset \mathbf{R}$ is a geometric GP then $\Phi \subset \mathcal{L}$ and $H \subset \mathbf{R}$ is a geometric GP-type structure based upon $\mathcal{G}$ with $C$ bunches if there are reals $u, v_1, v_2, \dots, v_C$ such that

$$H = (\mathcal{G} \cup \{0\}) + u;$$
$$\Phi = \{x \mapsto g(x-u) + v_i \ ; \ \ 1 \le i \le C \ , \ g \in \mathcal{G}\}; \ \text{ and so}$$
$$\Phi(H) = \bigcup_{i=1}^{C} \big((\mathcal{G} \cdot \mathcal{G} \cup \{0\}) + v_i\big).$$

**Remark 12.** These structures have, indeed, small image sets — as shown by Proposition 3.

**Theorem 3.** *For every $C > 0$ there are $C^* = C^*(C) > 0$ and $d^* = d^*(C) > 0$ with the following property. If $\Phi \subset \mathcal{L}$ and $H \subset \mathbf{R}$ with $|\Phi|, |H| \ge n$ and*

$$|\Phi(H)| \le Cn$$

*then $(\Phi, H)$ is contained in an arithmetic or in a geometric GP-type structure with $\le C^*$ slopes or with $\le C^*$ bundles, respectively, based upon an arithmetic or geometric $\mathcal{G}_{d^*, C^* n}$.*

The proof can be found in Section 5.

### 4.1. Uniqueness

**Remark 13.** In the above Theorem $\mathcal{G}$ is not uniquely determined by $H$ and $\Phi$; e.g., it can simply be extended arbitrarily, provided that its actual dimension and size are sufficiently small, as compared to the bounds we imposed on them.
However, for geometric GP-type structures and large values of $n$, we have no such freedom for the value of $u$ which occurs in Definition 11. This is demonstrated by the following observation.

**Theorem 4.** *In case (ii) of Theorem 3, $H$ uniquely determines parameter $u$ of the geometric GP-type structure (by which $\mathcal{G} \cup \{0\}$ is shifted in Definition 11), for $|H| > n_0 = n_0(C)$.*

## 5. Proof of the Theorems

Our first main tool is the following result (proven as Theorem 1 in part I [1]).

**Proposition 14.** *For every $C > 0$ there is a $c^* = c^*(C) > 0$ with the following property. Assume that*

$$|\Phi \circ \Psi| \le Cn$$

*for some $\Phi, \Psi \subset \mathcal{L}$ with $|\Phi|, |\Psi| \geq n$. Then $\Phi$ and $\Psi$ contain some $\Phi^* \subset \Phi$ and $\Psi^* \subset \Psi$ with $|\Phi^*|, |\Psi^*| \geq c^* n$, for which*

(i) *either both $\Phi^*$ and $\Psi^*$ consist of parallel lines;*

(ii) *or both $\Phi^*$ and $\Psi^*$ consist of concurrent lines.*

Also, in the sequel we shall heavily rely upon the following result of Freiman and Ruzsa [2, 5, 6].

**Proposition 15. (Freiman–Ruzsa Theorem)** *For every $C > 0$ there is a $d^* = d^*(C) > 0$ and a $C^* = C^*(C) > 0$ with the following property.*
*If*

$$|A + B| \leq Cn \quad \text{or} \quad |A - B| \leq Cn$$

*for some $A, B \subset \mathbf{R}$ with $|A|, |B| \geq n$, then $A \cup B$ is contained in an arithmetic $\mathcal{G}_{d^*, C^* n}$. Similarly, if $|A \cdot B| \leq Cn$ or $|A/B| \leq Cn$ (and, in either case, $0 \notin A, B$), then $A \cup B$ is contained in a geometric $\mathcal{G}_{d^*, C^* n}$.*

As mentioned before, this statement holds true for any torsion-free Abelian group.

### 5.1. Proof of Theorem 1

To prove the bound on the number of slopes or on that of the concurrent bunches is straightforward.

1. First we apply Proposition 14 and get a regular substructure $\Phi^*$ (we shall not need $\Psi^*$ here) with $|\Phi^*| \stackrel{\text{def}}{=} n^* \geq c^* n$.
2. According to its nature, we distinguish two cases.
   (i) If $\Phi^*$ consists of parallel lines, let $\phi_i \in \Phi^*$ be the graph of $x \mapsto sx + a_i$ $(i = 1 \dots n^*)$ while $\psi_j \in \Psi$ be that of $x \mapsto t_j x + b_j$. Then $\phi_i \circ \psi_j$ is given by

   $$x \mapsto st_j x + (sb_j + a_i).$$

   Here, for distinct $t_j$, also the slopes $st_j$ are distinct. Moreover, for $j$ fixed, there are at least $n^* \geq c^* n$ distinct constant terms $sb_j + a_i$. Therefore, the number of distinct slopes $t_j$ cannot exceed $(Cn)/(c^* n) = C/c^*$.

   (ii) If the lines in $\Phi^*$ are concurrent, let $\phi_i \in \Phi^*$ be the graph of $x \mapsto a_i(x - v) + u$ $(i = 1 \dots n^*)$ while, without loss of generality, $\psi_j \in \Psi$ can be written in the form $x \mapsto b_j(x - u_j) + v$. Then $\phi_i \circ \psi_j$ is given by

   $$x \mapsto a_i b_j (x - u_j) + u.$$

   Here, for distinct $u_j$, the points of intersection with the line $y = u$ are distinct. Moreover, for $j$ fixed, there are at least $n^* \geq c^* n$ distinct slopes

$a_ib_j$. Therefore, the number of distinct intersections $u_j$ cannot exceed $(Cn)/(c^*n)=C/c^*$.

3. We do the same for the largest parallel/concurrent bundle $\hat{\Psi}$ and the set of inverses $\hat{\Psi}^{-1}\circ\Phi^{-1}$ and get similar structure for $\Phi^{-1}$.

Let us sum up what we have achieved.

There is a $C^*=C^*(C)$; moreover

(i) either there are some $s_i$ $(i=1\dots C^*)$ such that

$$\Phi^{-1}=\{x\mapsto s_i\cdot x+a_{ir}\ ;\ i\le C^*, r\le l_i\};$$
$$\Psi=\{x\mapsto s_i\cdot x+b_{it}\ ;\ i\le C^*, t\le m_i\},$$

where $l_i\ge 0$ (resp. $m_i\ge 0$) is the number of the $\phi\in\Phi$ (resp. $\psi\in\Psi$) of slope $s_i$ while the $a_{ir}$ (resp. $b_{it}$) denote the points of intersection of these lines with the $y$-axis;

(ii) or there are $u_i$ $(i=1\dots C^*)$ and $v$ such that

$$\Phi^{-1}=\{x\mapsto a_{ir}(x-u_i)+v\ ;\ i\le C^*, r\le l_i\};$$
$$\Psi=\{x\mapsto b_{it}(x-u_i)+v\ ;\ i\le C^*, t\le m_i\},$$

where $l_i\ge 0$ (resp. $m_i\ge 0$) is the number of the $\phi\in\Phi$ (resp. $\psi\in\Psi$) through $(u_i,v)$ while the $a_{ir}$ (resp. $b_{it}$) denote the slopes of these lines.

We are left to show the arithmetic or geometric GP property of the $a_{ir}$ and $b_{it}$.

(i) In the "$C^*$ slopes" case, consider any pair

$$\phi:\ x\mapsto\frac{1}{s_i}(x-a_{ir});$$
$$\psi:\ x\mapsto s_jx+b_{jt}.$$

Then $\phi\circ\psi: x\mapsto\frac{s_j}{s_i}x+\frac{b_{jt}-a_{ir}}{s_i}$. Here the constant term can only take $Cn$ or fewer distinct values while the $s_i$ at most $C^*$. Thus, for the values of the differences

$$\{b_{jt}-a_{ir}\ ;\ 1\le i,j\le C^*, r\le l_i, t\le m_j\}$$

there are not more than $C^*Cn$ distinct possibilities. By Proposition 15, all the $a_{ir}$ and the $b_{jt}$ must be contained in an appropriate arithmetic GP.

(ii) In the "$C^*$ concurrent bunches" case, let

$$\phi:\ x\mapsto\frac{x-v}{a_{ir}}+u_i;$$
$$\psi:\ x\mapsto b_{jt}(x-u_j)+v.$$

Use Proposition 15 again, for the leading coefficients $b_{jt}/a_{ir}$ of the $\phi\circ\psi$, yielding an appropriate geometric GP.

This finishes the proof of Theorem 1. ∎

### 5.2. Proof of Theorem 2

Assume for a contradiction that for some $C$ and infinitely many values of $n$, there exist $\Phi=\Phi_n$ for which at least $cn$ of the $\phi\in\Phi$ have slope $\alpha=\alpha_n$, for some fixed $c>0$.

1. For the arithmetic GP $\mathcal{G}$,
$$\frac{n}{C}\le|\mathcal{G}|\le Cn.$$
(Here the first inequality comes from the fact that there are at most $C$ slopes.) Put $\mathcal{H}=(-\mathcal{G})\cap(\alpha G)$. Then, by the indirect assumption, $|\mathcal{H}|\ge cn$.
2. $|\mathcal{H}+(-\mathcal{G}\cup\alpha G)|\le|-\mathcal{G}-\mathcal{G}|+|\alpha\mathcal{G}+\alpha\mathcal{G}|\le 2|\mathcal{G}+\mathcal{G}|\le 2^{d+1}Cn$ by Proposition 3. Thus Proposition 15 implies the existence of an arithmetic GP $\mathcal{G}^*$ of dimension not exceeding $d^*$ and size $|\mathcal{G}^*|\le C^*n$ for which
$$-\mathcal{G}\cup\alpha G\subset\mathcal{G}^*. \tag{2}$$
3. For every $a,b\in-\mathcal{G}$, $(a\ne b)$, the quadruple
$$Q_{a,b}\stackrel{\text{def}}{=}\{2a,a+b,2b,2a+\alpha(b-a)\}\subset(-\mathcal{G})+(-\mathcal{G})+(\alpha G)-(\alpha G)\subset$$
$$\subset\mathcal{G}^*+\mathcal{G}^*+\mathcal{G}^*-\mathcal{G}^*,$$
by (2). Here the four-term sum set on the right has not more than $N=2^{4d^*}C^*n$ elements by Proposition 3. Moreover, $Q_{a,b}$ is the homothetic image of $\{0,1,2,\alpha\}$ under the mapping $x\mapsto(b-a)x+2a$.
4. Hence an $N$-element set contains at least
$$\binom{|\mathcal{G}|}{2}\ge\binom{n/C}{2}>\frac{n^2}{3C^2}=\frac{N^2}{3C^2 2^{8d^*}C^{*2}}=\hat{c}N^2$$
homothetic copies of $\{0,1,2,\alpha\}$ for a transcendental $\alpha$. However, a result of Laczkovich and Ruzsa ([4]) says that the number of such copies cannot exceed a certain sequence $S_N$ where $S_N/N^2\to 0$ (if $N\to\infty$); a contradiction if $N$ is large enough. ∎

### 5.3. Proof of Theorem 3

First we show that if an image set $\Phi(H)$ is small then also $\Phi\circ\Phi^{-1}$ must be small.

**Lemma 16.** *For every $C>0$ there is a $C^*=C^*(C)>0$ such that if $\Phi\subset\mathcal{L}$, $H\subset\mathbf{R}$ with $|\Phi|,|H|\ge n$ and $|\Phi(H)|\le Cn$ then $|\Phi\circ\Phi^{-1}|\le C^*n$.*

**Proof.** Put $K:=\Phi(H)$. Now if $\phi_1,\phi_2\in\Phi$ then the graph of $\phi_1\circ\phi_2^{-1}$ contains $n$ or more points of $K\times K$, e.g., those of type $(\phi_1(h),\phi_2(h))$ for $h\in H$. According to Lemma 20 in part I, there cannot exist more than $C^*n$ such $\phi_1\circ\phi_2^{-1}$. ∎

Now we can apply Theorem 1 to $\Phi\circ\Phi^{-1}$ and deduce that there are $C^*$ or fewer slopes or concurrent bunches in $\Phi$. (We also know an arithmetic or geometric GP-type structure for $\Phi$; however, this will be of little use for us since we know no such thing about $H$.)

(i) In the "$C^*$ slopes" case, use the notation

$$\Phi = \{x \mapsto s_i(x - a_{ir}) \ ; \ i \le C^*, r \le l_i\},$$

where $l_i \ge 0$ is the number of the $\phi \in \Phi$ of slope $s_i$ while the $a_{ir}$ denote the points of intersection of these lines with the $x$-axis.
Then $\Phi(H) = \{s_i(h - a_{ir}) \ ; \ i \le C^*, r \le l_i, h \in H\}$. Since $|\Phi(H)| \le Cn$ and there are at most $C^*$ of the $s_i$, the differences $h - a_{ir}$ can only take $C^*Cn$ or fewer distinct values. Hence Proposition 15 implies the existence of an appropriate arithmetic GP.

(ii) In the "$C^*$ concurrent bunches" case, let

$$\Phi = \{x \mapsto a_{ir}(x - u) + v_i \ ; \ i \le C^*, r \le l_i\},$$

where $l_i \ge 0$ is the number of the $\phi \in \Phi$ through $(u, v_i)$ while the $a_{ir}$ denote the slopes of these lines.
Then $\Phi(H) = \{a_{ir}\overline{h} + v_i \ ; \ i \le C^*, r \le l_i, \overline{h} \in (H - u)\}$. Here the products $a_{ir}\overline{h}$ cannot take more than $C^*Cn$ distinct values. Excluding $\overline{h} = 0$, a geometric GP is found by Proposition 15 again.

This finishes the proof of Theorem 3. ∎

### 5.4. Proof of Theorem 4

**Lemma 17.** *For every $C > 0$, $d > 0$ there is an $n_0 = n_0(C, d) > 0$ with the following property.*
*Let $\mathcal{G}_1$, $\mathcal{G}_2$ be geometric GPs of dimension at most $d$ each, and sizes $|\mathcal{G}_1|, |\mathcal{G}_2| \le Cn$. If $u_1$, $u_2$ are arbitrary reals such that*

$$|(\mathcal{G}_1 + u_1) \cap (\mathcal{G}_2 + u_2)| \ge n,$$

*then $u_1 = u_2$, provided that $n > n_0$.*

**Proof.** Let $H = (\mathcal{G}_1 + u_1) \cap (\mathcal{G}_2 + u_2)$. Define $\Phi = \Phi_1 \cup \Phi_2$ to be the union of two bunches of lines through $(u_1, 0)$ and $(u_2, 0)$, respectively:

$$\Phi_1 = \{x \mapsto g(x - u_1) \ ; \ g \in \mathcal{G}_1\};$$
$$\Phi_2 = \{x \mapsto g(x - u_2) \ ; \ g \in \mathcal{G}_2\}.$$

Then $|\Phi(H)| \le |\Phi_1(H)| + |\Phi_2(H)| \le |\mathcal{G}_1\mathcal{G}_1| + |\mathcal{G}_2\mathcal{G}_2| \le 2^{d+1}Cn$, by Proposition 3. Thus, according to Lemma 16, $|\Phi \circ \Phi^{-1}| \le C^*n$, for some $C^* = C^*(2^{d+1}C)$. However, it is easy to see that, unless $u_1 = u_2$,

$$|\Phi \circ \Phi^{-1}| \ge |\Phi_1 \circ \Phi_2^{-1}| = |\Phi_1||\Phi_2^{-1}| \ge n^2,$$

a contradiction for $n > n_0 = C^*$. ∎

This lemma immediately implies the statement of Theorem 4, as well. ∎

**Acknowledgments.** The author is deeply indebted to Péter Hajnal. The above results started growing from one of his questions. Also, we thank I. Z. Ruzsa for reading an earlier version of the manuscript very carefully and suggesting several corrections and simplifications.

## References

[1] György Elekes: On linear combinatorics I, *Combinatorica*, **17**(4) (1997), 447–458.

[2] Gregory A Freiman: *Foundations of a Structural Theory of Set Addition, Translation of Mathematical Monographs vol. 37*, Amer. Math. Soc., Providence, R.I., USA, 1973.

[3] Ron Graham and Jaroslav Nešetřil, editors. *The Mathematics of Paul Erdős*, Springer, 1996.

[4] Miklós Laczkovich and Imre Z Ruzsa: *The Number of Homothetic Subsets*, In Graham and Nešetřil [3], 1996.

[5] Imre Z Ruzsa: Arithmetical progressions and the number of sums, *Periodica Math. Hung.*, **25** (1992), 105–111.

[6] Imre Z Ruzsa: Generalized arithmetic progressions and sum sets. *Acta Math. Sci. Hung.*, **65** (1994), 379–388.

[7] Endre Szemerédi: On sets of integers containing no $k$ elements in arithmetic progression. *Acta Arithmetica*, **27** (1975).

György Elekes

*Department of Computer Sciences*
*Eötvös University, Budapest*
*Múzeum krt. 6–8.*
*H-1088, Budapest, Hungary*
elekes@cs.elte.hu