FACTORIZATION IN GENERALIZED ARITHMETIC PROGRESSIONS AND APPLICATION TO THE ERDŐS-SZEMERÉDI SUM-PRODUCT PROBLEMS

M. Chang

1. Introduction and Statements

Given integers $c_0, c_1, \cdots, c_d$ and $J_1, \ldots, J_d \geq 1$, a (generalized) d-dimensional arithmetic progression P is

$$P = P(c_0; c_1, \ldots, c_d; J_1, \ldots, J_d) = \{c_0 + \sum_{i=1}^{d} k_i c_i | k_i \in \mathbb{Z}, 0 \leq k_i \leq J_i\}. \tag{1.1}$$

A progression is called *proper* provided all expressions in (1.1) represent different numbers.

The importance of this notion appears for instance in the context of Freiman's theorem on sumsets.

Theorem 0. *Let A be a sufficiently large finite set of integers such that*

$$|A + A| \leq C_1 |A|, \tag{1.2}$$

for some constant C_1.

Then A is contained in a proper d-dimensional arithmetic progression P satisfying

$$d \leq [C_1 - 1] \tag{1.3}$$

$$|P| \leq C_2 |A|, \tag{1.4}$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

where $C_2 = C_2(C_1)$ is a constant depending on C_1.

This result has been the focus of research and improvements over recent years. See for instance [B-L], [Bi], [Ch1], [C-Z], [E], [E-S], [Fr], [F-H-R], [G1], [G2], [He], [K-L-T], [K-T], [Na], [N-T], [Ru1], [Ru2], [Ru3], [S-T], and [T].. It is shown in particular in [Ch1] that one may take in (1.4)

$$\log C_2 \lesssim C_1^2 \big(\log C_1\big)^3. \tag{1.5}$$

In [E-S], it is conjectured that either the sumset $A+A$ or the product set $A.A$, A being an arbitrary finite set of integers, needs to have essentially extremal size, in the following sense

$$\max(|A+A|, |A.A|) > c_\varepsilon |A|^{2-\varepsilon} \text{ for all } \varepsilon > 0, \tag{1.6}$$

and, more generally for all $h \geq 2$

$$\max(|\underbrace{A+\cdots+A}_{h}|,\ |\underbrace{A.\cdots.A}_{h}|) > c_{\varepsilon,h} |A|^{h-\varepsilon} \text{ for all } C_1 > 0. \tag{1.7}$$

These problems are still open. Concerning (1.6), best result to date is due to Elekes [E]

$$\max(|A+A|, |A.A|) > |A|^{5/4}. \tag{1.8}$$

The proof makes essential use of the geometric Szemerédi-Trotter theorem [S-T], which has shown itself a rather powerful tool in this type of questions. There are a few special instances where (1.6), (1.7) may be established. First, it is shown in [Ch2] (among other things) that if A satisfies

$$|A.A| < C|A|, \tag{1.9}$$

then for all h

$$|\underbrace{A+\cdots+A}_{h}| > c_h(C)|A|^h. \tag{1.10}$$

(the argument uses a weak form of Freiman's theorem and methods from Harmonic Analysis).

In another recent paper [E-R], the general inequality

$$|A+A|^4.|A.A|.\log|A| > |A|^6 \tag{1.11}$$

is obtained (again based on the Szemerédi-Trotter theorem). As a consequence,

$$|A+A| < C|A| \tag{1.12}$$

implies that

$$|A.A| > C^{-4}\frac{|A|^2}{\log|A|}. \tag{1.13}$$

The purpose of this paper is to prove the following facts.

Theorem 1. *Let A be a finite set of integers satisfying (1.12) for some constant C, and let $r_h(n;A)$ be the number of representations of n as a product of h elements in A. Then there is a uniform estimate for all $n \in \mathbb{Z}$*

$$r_2(n;A) < e^{C'(C)\frac{\log|A|}{\log\log|A|}}. \tag{1.14}$$

More generally, for $h \geq 2$

$$r_h(n;A) < e^{C_h(C)\frac{\log|A|}{\log\log|A|}}. \tag{1.15}$$

An immediate corollary of Theorem 1 is the following extension of the Erdős-Szemerédi theorem to multiple factors.

Theorem 2. *If A satisfies (1.12), then for all $h \geq 2$*

$$|\underbrace{A.\cdots.A}_{h\text{ fold}}| > c_{\varepsilon,h}|A|^{h-\varepsilon} \text{ for all } \varepsilon > 0. \tag{1.16}$$

The starting point is Freiman's theorem and our assumption (1.12) permits us to replace A by a proper generalized d-dimensional arithmetic progression. Thus Theorem 1 is a consequence of Theorem 0 and

Proposition 3. *Let P be as in (1.1) and $J = \max_i J_i$. Then, for all $n \in \mathbb{Z}$*

$$r_2(n;P) < e^{C_d \frac{\log J}{\log\log J}}, \tag{1.17}$$

and similarly, for arbitrary $h \geq 2$

$$r_h(n;P) < e^{C_{d.h} \frac{\log J}{\log\log J}}. \tag{1.18}$$

Thus estimates (1.17), (1.18) are uniform: they only depend on d, J, but not on the generators $c_0, c_1, \dots, c_d$ of P and n.

Remark 1. In Proposition 3, we do not use the fact that P is proper. Moreover, the statement remains true for arbitrary real or complex generators $c_0, c_1, \dots, c_d \in \mathbb{C}$ and $n \in \mathbb{C}$. In fact, the proof will be given in this generality (see section 3).

Remark 2. The proof of Proposition 3 presented below uses essentially the theory of algebraic number fields, hence methods very different from those mentioned earlier. We give the argument for $h = 2$. The general case can be gotten by induction. (Note that the progression does not need to be proper.)

In the next section, we present some algebraic number theory facts that are then used in Section 3 to prove Proposition 3.

In Section 4, we give another application of Theorem 1 to the Erdős -Szemerédi sum-product problem along graphs (see Proposition 4.4).

Section 5 contains further applications of Proposition 3 (together with Remark 1) to problems considered in [E-S], [E-N-R] and [E2] (see Section 2.1.1).

We show in particular the following facts.

Proposition 5. *Let A be an arbitrary finite set of complex numbers, $|A| = N$, such that*

$$|A + A| < C|A| \tag{1.19}$$

for some constant C. Then

$$\left|\frac{1}{A}+\frac{1}{A}\right| > e^{-C'\frac{\log N}{\log\log N}}N^2. \tag{1.20}$$

Proposition 6. *Let $A \subset \mathbb{C}$ be as in Proposition 5 satisfying (1.19). Let $p(X) \in \mathbb{C}[X]$ be a polynomial of degree $r \geq 2$. Then*

$$|p(A)+p(A)| > e^{-C'\frac{\log N}{\log\log N}}N^2, \tag{1.21}$$

where $C' = C'(C, r)$.

We denoted here

$$\frac{1}{A} = \{\frac{1}{x}|x \in A\} \qquad \text{and}$$
$$p(A) = \{p(x)|x \in A\}.$$

Remark 3. It was shown in [E-N-R] that if $A \subset \mathbb{R}$ is a finite set with $|A| = N$ and f a strictly convex (or concave) function defined on an interval containing A, then

$$|A \pm A|.|f(A) \pm f(A)| > cN^{5/2}, \tag{1.22}$$

where c is an absolute constant.

Their proof uses extensions of the Szemerédi-Trotter result to so-called ‘pseudo-line systems’, but seem so far only established in the real case.

An immediate consequence of (1.22) ([E-N-R]) is

$$|A+A|.\left|\frac{1}{A}+\frac{1}{A}\right| > cN^{5/2} \tag{1.23}$$

for $A \subset \mathbb{R}, |A| = N$.

The ‘natural’ conjecture is again the validity of (1.22) and (1.23) with lower bound $c_\varepsilon N^{3-\varepsilon}$, for all $\varepsilon > 0$. Thus Proposition 5 establishes this fact for (1.23), in the extremal case when $|A+A| < CN$ for some constant C. Moreover both Proposition 5 and 6 remain valid in the complex case.

Acknowledgement. The author would like to thanks J. Bourgain and P. Sarnak for helpful communication.

2. Some Algebraic Preliminaries

We first specify certain terminology, in order to ease the exposition.

Let us fix a large integer J.

Definition. A polynomial $p(X) \in \mathbb{Z}[X_1, \dots, X_r]$ is a *good* polynomial, provided its degree is bounded and all its coefficients are integers bounded by some power J^C of J, where C stands for an unspecified constant, understood to remain bounded when $J \to \infty$.

In the sequel, the number of variables r and the degree of $p(X)$ will always remain bounded.

Let us also agree to use in the above definition the letter "C" for possibly different constants. Thus with this convention the class of good polynomials is clearly closed under addition and multiplication. Furthermore, if p, q are good, so will be the resultant $\operatorname{Res}(p, q; X_i)$ of p, q with respect to one of the variables X_i (it is given by the Sylvester determinant, which is a fixed polynomial expression in the coefficients).

Definition. An algebraic number is *good*, if it is a root of a good polynomial.

Thus, the set of good algebraic numbers is closed under addition, multiplication and division. Moreover, the root of a polynomial of bounded degree whose coefficients are good algebraic numbers is a good algebraic number.

A *good algebraic integer* is an algebraic integer, which is a good algebraic number. Clearly, good algebraic integers form a ring.

Definition. Let R be an integral domain, and let $p(X)$ and $q(X) \in R[X]$ be polynomials. We call $r(X)$ a *good remainder* of $p(X)$ divided by $q(X)$, if $r(X)$ is the remainder (in the usual sense) multiplied by common denominators of all its coefficients, such that $r(X) \in R[X]$.

Clearly, we have

$$p(X) = q(X)h(X) + \frac{r(X)}{(q_d)^M},$$

where q_d is the leading coefficient of $q(X)$, and $M < deg(q(X))$.

Also, if p and q are good polynomials, then so is the good remainder of p divided by q.

Lemma 2.1. *Let $\alpha \in \mathbb{C}$ be a good algebraic number. Then*

$$J^{-C} < |\alpha| < J^C.$$

Proof. It suffices to show the upper bound.

Let $\sum_{j=0}^{d} a_j X^j \in \mathbb{Z}[X]$ be a good polynomial of degree d satisfied by α. Therefore, we have

$$\sum_{j=0}^{d} a_j \alpha^j = 0 \text{ with } d < C, \text{ and } |a_j| < J^C, \text{ for all j.}$$

We may assume that $|\alpha| > 1$.

Then clearly

$$|\alpha|^d \leq |a_d| \ |\alpha^d| \leq \sum_{j \leq d-1} |a_j| \ |\alpha|^j \leq J^C |\alpha|^{d-1}.$$

Lemma 2.2. *Let $p \in \mathbb{Z}[X]$ be a good polynomial and $q \in \mathbb{Z}[X]$ divide p in $\mathbb{Z}[X]$. Then q is a good polynomial.*

Proof. Let $p = \sum_{j=0}^{d} a_j X^j, q = \sum_{j=0}^{d'} b_j X^j$. Thus $d' \leq d$ is bounded and $b_{d'} | a_d$, hence $|b_{d'}| < J^C$. If we factor q in $\mathbb{C}[X]$ as $q(X) = b_{d'} \prod_{i=1}^{d'} (X - \alpha_i)$, then for all $i = 1, \ldots, d'$, $p(\alpha_i) = 0$, hence $|\alpha_i| < J^C$ by (2.1). Expressing the coefficients b_j in $b_{d'}$ and the $\{\alpha_i\}$, obviously we have $|b_j| < J^C$.

Lemma 2.3. *For an algebraic integer α of bounded degree, the following properties are equivalent*

(1) α is a good algebraic integer.

(2) The minimum polynomial of α is a good polynomial.

(3) All conjugates $\sigma_i(\alpha)$ satisfy

$$J^{-C} < |\sigma_i(\alpha)| < J^C. \tag{2.4}$$

We denote here by $\sigma_1, \dots, \sigma_n$ the n Q-isomorphisms of $\mathbb{Q}(\alpha)$ into the field $\mathbb{C}$ of complex numbers, where $n = [Q(\alpha) : \mathbb{Q}]$ is assumed bounded.

Proof.

(1) implies (2): The minimum polynomial of α divides in $\mathbb{Z}[X]$ a good polynomial and hence is good by Lemma 2.2.

(2) implies (3): By Lemma 2.1.

(3) implies (1): α is root of the minimal polynomial $\prod_{i=1}^n \big(X - \sigma_i(\alpha)\big)$ whose coefficients are integers bounded by J^C.

Lemma 2.4. *Let α be a good algebraic integer and let $\mathbb{Q}(\alpha) \subset K$ be a finite extension with $[K : \mathbb{Q}]$ bounded. Then the norm $N_{K/\mathbb{Q}}(\alpha)$ is a rational integer bounded by J^C. (the constant C depends on α and $[K : \mathbb{Q}]$).*

Proof. Let $d = [K : \mathbb{Q}]$ and $\nu_1, \dots, \nu_d$ be the d $\mathbb{Q}$-isomorphisms of K into $\mathbb{C}$. Then the characteristic polynomial $\prod_{j=1}^d \big(X - \nu_j(\alpha)\big)$ of α relative to the extension $K/\mathbb{Q}$ is a power of the minimum polynomial of α. Lemma 2.3 implies

$$N_{K/\mathbb{Q}}(\alpha) = \prod \nu_j(\alpha) \text{ is bounded by } J^C.$$

The following fact is the key result for what follows

Proposition 2.5. *Let K be an extension of $\mathbb{Q}$ with $[K : Q]$ bounded. Let $\alpha \in K\backslash\{0\}$ be a good algebraic integer. Then the number of factorizations in K*

$$\alpha = \alpha_1.\alpha_2 \tag{2.6}$$

with α_1, α_2 good algebraic integers, is at most

$$e^{C' \frac{\log J}{\log\log J}}. \tag{2.7}$$

Proof. Denote $d = [K : \mathbb{Q}]$ and $\nu_1, \dots, \nu_d$ the d $\mathbb{Q}$-isomorphisms of K into $\mathbb{C}$. According to Lemma 2.4,

$$0 \neq |N(\alpha)| = |N_{K/\mathbb{Q}}(\alpha)| < J^C. \tag{2.8}$$

If (2.6) holds, then $N(\alpha_1)|N(\alpha)$ and hence we are bounding the following

$$\sum_{\substack{a\in\mathbb{Z}\\ a|N(\alpha)}} \{\alpha_1|\alpha_1 \text{ is a good algebraic integer and } N(\alpha_1) = a\}. \tag{2.9}$$

First, fixing a, we consider the number $\omega(a)$ of non-associate elements β of K for which $N(\beta) = a$. Recall that elements $\beta, \gamma \in K$ are associate if $\beta = \xi\gamma$, where ξ is a unit. This problem is discussed in [B-S] (see Chapter 3, Section 7) and one has the estimate

$$\omega(a) \leq \tau(a)^d, \tag{2.10}$$

where $\tau(a)$ is the divisor function of a (see [B-S], p. 220).

In order to solve our problem, it remains to determine the number of units ξ in K which are good. Indeed, we only need to consider units ξ that appear as $\xi = \frac{\alpha_1'}{\alpha_1''}$, where α_1', α_1'' are associated good algebraic integers in K.

Let U be the group of units of K and consider the logarithmic homomorphism

$$\ell : U \to \mathbb{R}^d : \zeta \to (\log|\nu_1(\zeta)|, \dots, \log|\nu_d(\zeta)|) \tag{2.11}$$

It is well known that $\ell(U)$ is a lattice $\mathcal{L}$ of dimension $\leq d-1$. (In what follows, we will not treat the roots of unity separately).

Let ξ be a good unit. From Lemma 2.3

$$J^{-C} < |\nu_j(\xi)| < J^C, \qquad (1 \leq j \leq d)$$

hence

$$\big|\log|\nu_j(\xi)|\,\big| < C\log J. \qquad (1 \leq j \leq d). \tag{2.12}$$

Hence, for such units ξ, $\ell(\xi)$ lies within a cube in $\mathbb{R}^d$, centered at 0 and of size $(2C\log J)^d$.

We claim that

$$|\{\xi \in U|\max_{1\leq j\leq d}\big|\log|\nu_j(\zeta)|\big| \leq 1\}| < C_d, \tag{2.13}$$

where C_d is a constant depending on d.

Indeed, an element in the set (2.13) has characteristic polynomial $\prod_{j=1}^d \big(X-\nu_j(\zeta)\big)$ with (integer) coefficients bounded by $\prod_{j=1}^d(1+|\nu_j(\zeta)|) < (1+e)^d$. Therefore, the boundedness of the number of the roots of such polynomials implies the boundedness of the number of all such ζ in (2.13).

Dividing the cube into unit cubes and using translation, we see that there are at most $C_d(2C\log J)^d$ good units ξ in U.

Summarizing, taking (2.8), (2.9), (2.10) into account, it follows that the number of factorizations in (2.6) is at most

$$\begin{aligned}&\sum_{a|N(\alpha)} C_d(2C\log J)^d\tau(a)^d\\ &\leq J^\epsilon C_d(2C\log J)^d \quad \left(e^{C_d\frac{\log J}{\log\log J}}\right)^d\\ &< e^{C'_d\frac{\log J}{\log\log J}}.\end{aligned}$$

(the constant C_d depends only on $d = [K : \mathbb{Q}]$ and the constant C involved in the assumption on α_1, α_2 in factorization (2.6).)

The main idea, explained in the next section, is to replace the factorization problem in an arithmetic progression by one in an algebraic number field to which Proposition 2.5 may be applied.

We will need one more algebraic fact.

Lemma 2.14. *Let V be a nonempty affine variety in $\mathbb{C}^r$ defined as*

$$V = \bigcap_{j=1,\dots s} [p_j = 0]$$

where $p_j(X_1, \dots, X_r) \in \mathbb{Z}[X_1, \dots, X_r]$ are good polynomials in the sense defined in the beginning of this section (r, and s are assumed bounded).

Then V contains an element $\beta = (\beta_1, \dots, \beta_r)$ whose coordinates are good algebraic numbers.

Proof. We will prove the following stronger statement by induction on the number r of variables.

Let $p_j (1 \le j \le s)$ and p be good polynomials in $\mathbb{Z}[X_1, \dots, X_r]$ such that the set

$$A = \bigcap_{j=1,\dots,s} [p_j = 0] \cap [p \ne 0] \ne \phi. \tag{2.15}$$

Then A contains an element $\beta = (\beta_1, \dots, \beta_r)$ whose coordinates are good algebraic numbers.

Of course, the case $r = 1$ is trivial. If not all p_j vanish identically, then all elements of A are good. If $p_j \equiv 0$ for all $j = 1, \dots, s$, let $0 \le \beta \le \deg p$ be an integer such that $p(\beta) \ne 0$.

Next, we show how to reduce the number of variables.

Assume $b = (b_1, \dots, b_r) \in A$.

Expand

$$p_j(X) = \sum_{0 \le k \le d_j} p_{j,k}(X_1, \dots, X_{r-1}) X_r^k$$

and denote

$$d'_j = \begin{cases} -1, \text{ if } p_{jk}(b_1, \dots, b_{r-1}) = 0 \text{ for all } 0 \le k \le d_j \\ \max\{k | p_{jk}(b_1, \dots, b_{r-1}) \neq 0\}, \text{ otherwise.} \end{cases} \tag{2.16}$$

Thus $d'_j = -1$ or $d'_j \ge 1$. If $d'_j \ge 1$, define

$$p'_j(X) = \sum_{0 \le k \le d'_j} p_{jk}(X_1, \dots, X_{r-1}) X_r^k.$$

Let further

$$A' = \bigcap_{\substack{j=1,\dots,s \\ d'_j \ge 1}} [p'_j = 0] \cap [p \neq 0] \cap B'$$

$$B' = \bigcap_{j=1,\dots,s} \Big(\bigcap_{d'_j < k \le d_j} [p_{jk} = 0] \cap [p_{j,d'_j} \neq 0] \Big)$$

where B' depends only on $X_1, \dots, X_{r-1}$.

Clearly

$$b \in A' \subset A.$$

If $d'_j = -1$ for all j, then

$$A' = [p \neq 0] \bigcap B'.$$

If $p(X) = \sum_{0 \le k \le d} p_k(X_1, \dots, X_{r-1}) X_r^k$, the assumption $p(b) \neq 0$ implies $p_k(b_1, \dots, b_{r-1}) \neq 0$ for some k. Thus $B' \cap [p_k \neq 0] \neq \phi$ and, by the induction hypothesis, contains a good point $(\beta_1, \dots, \beta_{r-1})$. Since $p(\beta_1, \dots, \beta_{r-1}, X_r) \neq 0$, there is a bounded rational integer β_r such that $p(\beta_1, \dots, \beta_{r-1}, \beta_r) \neq 0$. Hence $\beta = (\beta_1, \dots, \beta_r) \in A' \subset A$.

Next, we assume $d'_j \geq 1$ for some j. We may then define $\tilde{d}$ as the smallest integer ≥ 1 for which there is a *good* polynomial

$$\tilde{p}(X) = \sum_{0 \leq k \leq \tilde{d}} \tilde{p}_k(X_1, \dots, X_{r-1}) X_r^b \in \mathbb{Z}[X_1, \dots, X_r]$$

satisfying

$$\tilde{p}(b) = 0 \text{ and } \tilde{p}_{\tilde{d}}(b_1, \dots, b_{r-1}) \neq 0. \tag{2.17}$$

Thus

$$b \in A_1 \equiv A' \cap [\tilde{p} = 0] \cap [\tilde{p}_{\tilde{d}} \neq 0]. \tag{2.18}$$

Using the Euclidean division algorithm with respect to X_r, we may clearly replace each polynomial p'_j with $d'_j \geq \tilde{d} \geq 1$ by a good polynomial, namely, the good remainder of p'_j divided by $\tilde{p}$,

$$p''_j(X) = \sum_{0 \leq k \leq d''_j} p''_{j,k}(X_1, \dots, X_{r-1}) X_r^k$$

of degree $0 \leq d''_j \leq \tilde{d} - 1$ such that

$$[\tilde{p} = 0] \cap \tilde{p}_{\tilde{d}} \neq 0] \cap [p'_j = 0] = [\tilde{p} = 0] \cap [\tilde{p}_{\tilde{d}} \neq 0] \cap [p''_j = 0].$$

In particular $p''_j(b) = 0$. By definition of $\tilde{d}$, it follows that $p''_{j,k}(b_1, \dots, b_{r-1}) = 0$ for all $k = 0, \dots, d''_j$. Therefore

$$b \in A_2 \equiv [\tilde{p} = 0] \cap [p \neq 0] \cap B_2 \subset A_1 \tag{2.19}$$

with

$$B_2 = B' \cap [\tilde{p}_{\tilde{d}} \neq 0] \cap \bigcap_{j,k} [p''_{jk} = 0]. \tag{2.20}$$

Also, $p(X)$ may be replaced by a good polynomial $\bar{p}(X) = \sum_{0 \leq k \leq \bar{d}} \bar{p}_k(X_1, \dots, X_{r-1}) X_r^k$ where $\bar{d} < \tilde{d}$ and

$$[\tilde{p} = 0] \cap [\tilde{p}_{\tilde{d}} \neq 0] \cap [p \neq 0] = [\tilde{p} = 0] \cap [\tilde{p}_{\tilde{d}} \neq 0] \cap [\bar{p} \neq 0].$$

Thus $\bar{p}(b) \neq 0$. Denote

$$\text{đ}_1 = \max\{k \leq \text{đ} | \bar{p}_k(b_1, \ldots, b_{r-1}) \neq 0\}$$

and replace B_2 by

$$b \in B_3 = B_2 \cap \bigcap_{\text{đ}_1 < k \leq \text{đ}} [\bar{p}_k = 0] \cap [\bar{p}_{\text{đ}_1} \neq 0].$$

If $\text{đ}_1 = 0$, then

$$b \in A_3 \equiv [\tilde{p} = 0] \cap B_3 \subset A_2. \tag{2.21}$$

Thus $B_3 \neq \phi$ contains a good point $(\beta_1, \ldots, \beta_{r-1})$ and solving $\tilde{p}(\beta_1, \ldots, \beta_{r-1}, X_r) = 0$ provides a good point $\beta = (\beta_1, \ldots, \beta_{r-1}, \beta_r) \in A_3 \subset A$.

Assume now $\text{đ}_1 \geq 1$.

Denoting

$$\bar{\bar{p}}\,(X) = \sum_{0 \leq k \leq \text{đ}_1} \bar{p}_k(X_1, \ldots, X_{r-1}) X_r^k$$

we have

$$b \in A_3 \equiv [\tilde{p} = 0] \cap [\bar{\bar{p}} \neq 0] \cap B_3 \subset A_2. \tag{2.22}$$

Let $R(X_1, \ldots, X_{r-1}) = \operatorname{Res}(\tilde{p}, \bar{\bar{p}}, X_r)$ be the resultant of $\tilde{p}$, $\bar{\bar{p}}$ with respect to the variable X_r. Assume $R(b_1, \ldots, b_{r-1}) = 0$.

We may then write (see [C-S-L])

$$P_0(X_r)\tilde{p}(b_1, \ldots, b_{r-1}, X_r) + P_1(X_r)\bar{\bar{p}}\,(b_1, \ldots, b_{r-1}, X_r) = 0 \tag{2.23}$$

where

$$\deg P_0 < \text{đ}_1, \deg P_1 < \text{đ}\ , \text{ and}$$

$$P_0(X_r), P_1(X_r) \text{ are not both identically zero.}$$

The coefficients of P_0, P_1 are expressed by (universal) integer polynomials in the coefficients $\tilde{p}_k(b_1, \dots, b_{r-1})$ and $\bar{p}_k(b_1, \dots, b_{r-1})$ of $\tilde{p}$, $\bar{\bar{p}}$. Hence, there are good polynomials $Q_0, Q_1 \in \mathbb{Z}[X_1, \dots, X_r]$ s.t.

$$P_0(X_r) = Q_0(b_1, \dots, b_{r-1}, X_r) \text{ and } P_1(X_r) = Q_1(b_1, \dots, b_{r-1}, X_r). \tag{2.24}$$

Substituting $X_r = b_r$ in (2.23), we get

$$Q_1(b) = P_1(b_r) = 0.$$

Since $\deg_{X_r} Q_1 < \bar{d}$, it follows again from definition of $\bar{d}$ that if $Q_1 = \sum_k Q_{1,k}(X_1, \dots, X_{r-1}) X_r^k$, necessarily $Q_{1,k}(b_1, \dots, b_{r-1}) = 0$ for all k, hence $P_1 \equiv 0$. Thus $P_0 \not\equiv 0$ and (2.23) implies $\tilde{p}(b_1, \dots, b_{r-1}, X_r) \equiv 0$, a contradiction.

Consequently, $R(b_1, \dots, b_{r-1}) \neq 0$ and $B_3 \cap [R \neq 0] \neq \phi$. Let $(\beta_1, \dots, \beta_{r-1})$ be a good point in $B_3 \cap [R \neq 0]$ and β_r satisfy $\tilde{p}(\beta_1, \dots, \beta_{r-1}, \beta_r) = 0$. Since $R(\beta_1, \dots, \beta_{r-1}) \neq 0$, $\tilde{p}(\beta_1, \dots, \beta_{r-1}, X_r)$ and $\bar{\bar{p}}(\beta_1, \dots, \beta_{r-1}, X_r)$ do not have a common root. In particular, $\bar{\bar{p}}(\beta) \neq 0$ and $\beta \in A_3 \subset A$. This completes the proof.

Remark. The above argument is elementary and self contained. We may alternatively proceed by describing the elimination ideals of $I = I(p_j; 1 \leq j \leq s)$ using Groebner basis theory (cf. [C-L-S]). Construction of a Groebner basis starting from the polynomials $p_j (1 \leq j \leq s)$ may be performed following Buchberger's algorithm and the resulting polynomials remain good polynomials (since this property is obviously preserved by the operation of taking S-polynomials).

3. Proof of Proposition 3

We prove (1.17) (the argument for (1.18) is similar).

Thus we need to bound the number of factorizations

$$n = (c_0 + \sum_{i=1}^{d} k_i c_i)(c_0 + \sum_{i=1}^{d} k_i' c_i) \text{ with } 0 \leq k_i, k_i' \leq J_i \leq J \tag{3.1}$$

by an expression

$$e^{C\frac{\log J}{\log\log J}}. \tag{3.2}$$

This bound is uniform in n and the generators c_i (only depends on d and J).

We prove in fact a stronger statement.

Let n and $c_i(0 \leq i \leq d)$, $c'_i(0 \leq i \leq d')$ be *arbitrary complex numbers.* Then the number of factorizations

$$n = (c_0 + \sum_{i=1}^{d} k_i c_i)(c'_0 + \sum_{i=1}^{d'} k'_i c'_i) \text{ with } k_i, k'_i \in \mathbb{Z}; |k|, |k'_i| \leq J \tag{3.3}$$

is bounded by (3.2).

Let us emphasize that only the number of factorizations is estimated, not the number of (k, k')-solutions of equation (3.3). We don't assume the progression proper.

This statement is proven by induction on $d + d'$. Thus we may assume

$$c_1 = c'_1 = 1, \tag{3.4}$$

by working on $\frac{n}{c_1 c_2}$.

Let $\mathcal{S}$ denote the set of all solutions $\bar{k} = (k, k') = (k_1, \ldots, k_d, k'_1, \ldots, k'_{d'})$ of (3.3) with $|k_i|, |k'_i| \leq J$. We fix a solution $\bar{l} = (l, l') = (l_1, \ldots, l_d, l'_1, \ldots, l'_{d'})$ in S. Let

$$V = \bigcap_{\bar{k},\in\mathcal{S}} [(X_0{+}k_1{+}\sum_{i=2}^{d} k_i X_i)(Y_0{+}k'_1{+}\sum_{i=2}^{d'} k'_i Y_i) = (X_0{+}\ell_1{+}\sum_{i=2}^{d} \ell_i X_i)(Y_0{+}\ell'_1{+}\sum_{i=2}^{d'} l'_i Y_i)]. \tag{3.5}$$

be the variety defined as common zero set of quadratic polynomials in $(X_0, X_2, \ldots, X_d, Y_0, Y_2, \ldots, Y_{d'})$.

Obviously, the number of defining polynomials may be reduced to $(d{+}1)(d'{+}1)$ and their coefficients are integers bounded by $2J^2$. From (3.3), $(c_0, c_2, \ldots, c_d, c'_0, c'_2, \ldots c'_{d'}) \in V \neq \phi$. Applying Lemma 2.14, we obtain a point $(\beta_0, \beta_2, \ldots, \beta_d, \beta'_0, \beta'_2, \ldots, \beta'_{d'}) \in V$

whose coordinates are good algebraic numbers. Thus, from definition of V, there is some λ such that

$$(\beta_0 + k_1 + \sum_{i=2}^{d} k_i\beta_i)(\beta_0' + k_1' + \sum_{i=2}^{d'} k_i'\beta_i') = \lambda \text{ for all } \bar{k} \in \mathcal{S}. \tag{3.6}$$

Multiplying an appropriate rational integer (bounded by J^C) to (3.6) permits us to obtain good algebraic integers $(\gamma_0, \gamma_1, \dots, \gamma_d, \gamma_0', \gamma_1', \dots, \gamma_{d'}')$ with $\gamma_1, \gamma_1' \neq 0$ and μ such that

$$(\gamma_0 + \sum_{i=1}^{d} k_i\gamma_i)(\gamma_0' + \sum_{i=1}^{d'} k_i'\gamma_i') = \mu \text{ for all } \bar{k} \in \mathcal{S}. \tag{3.7}$$

Assume $\mu \neq 0$.

Consider the finite extension $K = \mathbb{Q}(\gamma_i (0 \le i \le d), \gamma_i' (0 \le i \le d'))$, thus $[K : \mathbb{Q}] \le d + d' + 2$. According to Proposition 2.5, the number of factorizations of μ in (3.7) is at most $e^{C \frac{\log J}{\log \log J}}$. Thus we may specify $\bar{k} \in \mathcal{S}$ by imposing

$$\begin{cases} \gamma_0 + \sum_{i=1}^{d} k_i\gamma_i = \alpha \in K \\ \gamma_0' + \sum_{i=1}^{d'} k_i'\gamma_i' = \alpha' \in K \end{cases} \tag{3.8}$$

for some fixed α, α' (taken in a set of size at most $e^{C \frac{\log J}{\log \log J}}$).

Thus, since $\gamma_1, \gamma_1' \neq 0$, we obtain the relations

$$\begin{cases} k_1 = \frac{\alpha - \gamma_0}{\gamma_1} - \sum_{i=2}^{d} \frac{\gamma_i}{\gamma_1} k_i \\ k_1' = \frac{\alpha' - \gamma_0'}{\gamma_1'} - \sum_{i=2}^{d'} \frac{\gamma_i'}{\gamma_1'} k_i'. \end{cases} \tag{3.9}$$

Substituting (3.9) in (3.3) implies the following representation of the factors

$$\begin{cases} c_0 + \sum_{i=1}^{d} k_i c_i = c_0 + c_1 \frac{\alpha - \gamma_0}{\gamma_1} + \sum_{i=2}^{d} k_i (c_i - c_1 \frac{\gamma_i}{\gamma_1}) \\ c_0' + \sum_{i=1}^{d'} k_i' c_i' = c_0' + c_1' \frac{\alpha' - \gamma_0'}{\gamma_1'} + \sum_{i=2}^{d'} k_i' (c_i' - c_1' \frac{\gamma_i'}{\gamma_1'}) \end{cases} \tag{3.10}$$

reducing the dimensions d, d' of the respective progressions to $d - 1, d' - 1$.

If in (3.7) $\mu = 0$, then either $\gamma_0 + \sum_{i=1}^d k_i\gamma_i = 0$ or $\gamma'_0 + \sum_{i=1}^{d'} k'_i\gamma'_i = 0$. If $\gamma_0 + \sum_{i=1}^d k_i\gamma_i = 0$. Then

$$k_1 = -\frac{\gamma_0}{\gamma_1} - \sum_{i=2}^{d} \frac{\gamma_i}{\gamma_1} k_i \tag{3.11}$$

and the first factor in (3.3) becomes

$$c_0 + \sum_{i=1}^{d} k_i c_i = c_o - c_1\frac{\gamma_0}{\gamma_1} + \sum_{i=2}^{d} k_i\big(c_i - c_1\frac{\gamma_i}{\gamma_1}\big) \tag{3.12}$$

which reduces the dimension d of the first progression to $d-1$.

Thus, writing $B(d+d', J)$ for a uniform bound on the number of factorizations in (3.3), we proved that

$$B(d+d', J) \le e^{C_{d+d'} \frac{\log J}{\log\log J}} B(d+d'-1, J). \tag{3.13}$$

Consequently

$$B(d+d', J) \le e^{C'_{d+d'} \frac{\log J}{\log\log J}} \tag{3.14}$$

proving in particular Proposition 3.

4. Sums and Products along Graphs

In [E-S], the following generalization of the sum-product problem is considered. Let A be a finite set of integers, $|A| = N$ and $G \subset A \times A$ an undirected graph. We write $a \sim a'$ provided $(a, a') \in G$ and define the following restricted sum and product set

$$A \overset{G}{+} A = \{a + a' | (a, a') \in G\} \tag{4.1}$$

$$A \overset{G}{\times} A = \{aa' | (a, a') \in G\}. \tag{4.2}$$

We consider then again the question how large $|A \overset{G}{+} A| + |A \overset{G}{\times} A|$ has to be. In particular, the question was raised wether for all $\varepsilon > 0$ and $0 < \delta < 1$

$$|A \overset{G}{+} A| + |A \overset{G}{\times} A| > c_{\varepsilon,\delta} |G|^{1-\varepsilon} \text{ , if } |G| > N^{1+\delta}. \tag{4.3}$$

One may prove the following particular case

Proposition 4.4. *(4.3) holds if we assume* $|G| > \delta N^2$ *and* $|A \overset{G}{+} A| < CN$*, for arbitrary (fixed) constants* $\delta > 0, C < \infty$*.*

The proof uses the following result of Laczkovich and Rusza [L-R] (related to the Balog-Szemerédi theorem and Gower's improvement [G1]).

Proposition 4.5. *[L-R]: Assume* $G \subset A \times A$ *satisfiers* $|G| > \delta N^2$, $|A \overset{G}{+} A| < CN$*. Then there is a subset* $A' \subset A$ *such that*

$$|A' + A'| < C'(\delta, C)N \tag{4.6}$$

$$|(A' \times A') \cap G| > \delta'(\delta, C)N^2 \tag{4.7}$$

Proof of Proposition 4.4. Take A' as in Proposition 4.5. Then, applying Theorem 1

$$\delta'(\delta, C)N^2 < |(A' \times A') \cap G| \le \sum_{n \in A' \overset{G}{\times} A'} r_2(n; A') < e^{C'' \frac{\log N}{\log\log N}} |A' \overset{G}{\times} A'|$$

and hence (4.3). (Note that $|G| < N^2$).)

Remark. Using the result from [Ch1], it can also be shown that if $|G| > \delta N^2$ and $|A \overset{G}{\times} A| < CN$, then $|A \overset{G}{+} A| > \delta'(\delta, C)N^2$.

5. Proof of Propositions 5 and 6

First, since Freiman's theorem remains valid in the generality of torsion-free Abelian groups (cf. [Na]), the assumption (1.19) permit us to assume $A \subset P$ where P is a generalized d-dimensional arithmetic progression with $c_0, c_1, \dots, c_d \in \mathbb{C}$ and satisfying

$$\prod_{i=1}^{d} J_i \le C'N. \tag{5.1}$$

Using the following (for Proposition 6)

Fact. Let $d(m) = |\{n_1, n_2) \in S \times S \big| n_1 - n_2 = m\}|$. If $d(m) < |S|^\epsilon$ for all $m \in S \subset \mathbb{C}$, then $|S + S| > |S|^{2-\epsilon}$.

To obtain (1.20), (1.21), it therefore suffices to prove the following facts

$$|\{(x, y) \in P \times P \big| \frac{1}{x} + \frac{1}{y} = \xi\}| < e^{C_d \frac{\log J}{\log \log J}} \tag{5.2}$$

and

$$|\{x, y) \in P \times P \big| p(x) - p(y) = \xi\}| < e^{C_{d,r} \frac{\log J}{\log \log J}} \tag{5.3}$$

where $J = \max J_i$ and $\xi \in \mathbb{C} \backslash \{0\}$ is arbitrary (the constants in (5.2), (5.3) depend only on d and the degree $r \geq 2$ of the polynomial p).

Proof of (5.2).

Writing $x_1 = x - \frac{1}{\xi}, y_1 = y - \frac{1}{\xi}$, the equation $\frac{1}{x} + \frac{1}{y} = \xi$ becomes

$$x_1 y_1 = \frac{1}{\xi^2} \tag{5.4}$$

where

$$x_1, y_1 \in \{c_0 - \frac{1}{\xi} + \sum_{i=1}^d k_i c_i | k_i \in \mathbb{Z}, |k_i| \leq J_i\} = P - \frac{1}{\xi}. \tag{5.5}$$

Thus, according to Proposition 3 and Remark 1, (5.4) has at most $e^{C \frac{\log J}{\log \log J}}$ solutions.

Proof of (5.3).

Write

$$p(X) = \sum_{s=1}^r a_s X^s \qquad (a_s \in \mathbb{C}, a_r \neq 0, r \geq 2)$$

(we may clearly assume $a_0 = 0$).

Then

$$\xi = p(x) - p(y) = (x - y) \left[\sum_{s=1}^r a_s (x^{s-1} + x^{s-2} y + \cdots + y^{s-1}) \right] = (x - y) z. \tag{5.6}$$

Thus

$$x - y \in P - P \subset \left\{ \sum_{i=1}^{d} k_i c_i | k_i \in \mathbb{Z} \text{ and } |k_i| \leq 2J_i \right\} \tag{5.7}$$

and $z \in Q$, where Q is the generalized arithmetic progression defined as follows

$$Q = \left\{ \sum k_{s,t_0,\dots,t_d} g_{s,t_0,\dots,t_d} \middle| k_{s,\bar{t}} \in \mathbb{Z}, \text{ and } |k_{s,\bar{t}}| < (r+1)^{r+1} J^r \right\}$$

where $s, t_0, \dots, t_d$ are integers s.t.

$$1 \leq s \leq r$$

$$t_0, \dots, t_d \geq 0, \sum t_i < r$$

and generators

$$g_{s,t_0,\dots,t_d} = a_s \prod_{i=0}^{d} c_i^{t_i}.$$

Thus Q has at most r^{d+2} generators. Again from Proposition 3 and Remark 1, the number of factorizations in (5.6) is at most

$$e^{C_{d,r} \frac{\log J}{\log \log J}}.$$

If we specify

$$x - y = w \tag{5.8}$$

where thus w is taken in a set of size $< e^{C \frac{\log J}{\log \log J}}$, we obtain the equation in X

$$p(X) - p(X - w) = \xi. \tag{5.9}$$

Since $\deg p \geq 2$, $p(X) - p(X - w)$ is not constant and there are at most $r - 1$ solutions in x.

This proves (5.3).

References

[B-L]. V. Bergelson, A. Liebman, *Polynomial extensions of van der Waerden's and Szemerédi's theorem*, J. AMS, Vol 9, N3 (1996), 725-753.

[Bi]. Y. Bilu, *Structure of sets with small sumset*, in 'Structure Theory of Set Addition', Astérisque 258 (1999), 77-108.

[B-S]. L.I. Borevich, I.R. Shafarevich, *Number Theory*, Academic Press, 1966.

[Ch1]. M. Chang, *Polynomial bounds in Freiman's Theorem*, to appear in Duke Math.

[Ch2]. M. Chang, *Erdös-Szeremedi sum-product problem.*

[C-Z]. G. Cohen, G. Zemor, *Subset sums and coding theory*, in 'Structure Theory of Set Addition', Astérisque 258 (1999).

[C-S-L]. D. Cox, J. Little, D. O'Shea, *Ideals, Varieties and Algorithms*, Springer-Verlag, 1992.

[E]. G. Elekes, *On the number of sums and products*, Acta Arithmetica, 81, 4: 365–367 (1997).

[E2]. G. Elekes, *Sums versus products in Number Theory*, Algebra and Erdös Geometry.

[E-N-R]. G. Elekes, M. Nathanson, I. Rusza, *Convexity and sumsets*, J. Number Theory, to appear.

[E-R]. G. Elekes, J. Rusza, *Few sums, many products*, preprint.

[E-S]. P. Erdős, E. Szemerédi, *On sums and products of integers*, In P. Erdös, L. Alpár, G. Halász (editors), Studies in Pure Mathematics; to the memory of P. Turán, p. 213–218.

[Fr]. G. Freiman, *'Foundations of a structural theory of set addition'*, Translations of Math. Monographs, 37, AMS, 1973.

[F-H-R]. G. Freiman, H. Halberstam, I. Ruzsa, *Integer sumsets containing long arithmetic progressions*, JLMS (2), 46 (1992), no 2, 193-201.

[G1]. W.T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length 4*, GAFA 8 (1998), 529-551.

[G2]. W.T. Gowers, *A new proof of Szemerédi's theorem*, preprint.

[He]. M. Herzog, *New results on subset multiplication in groups*, in 'Structure Theory of Set Addition', Astérisque 258 (1999).

[K-L-T]. N. Katz, I. Laba, T. Tao, *An improved bound on the Minkowski dimension of Besicovitch sets in R^3*, Annals of Math..

[K-T]. N. Katz, T. Tao, *Some connections between the Falconer and Furstenburg conjectures*, New York J. Math..

[L-R]. Laczkovich, I. Ruzsa.

[Na]. M.B. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets, Springer (1996)..*

[N-T]. M. Nathanson, G. Tenenbaum, *Inverse theorems and the number of sums and products*, in 'Structure Theory of Set Addition', Astérisque 258 (1999).

[Ru1]. I. Ruzsa, *Arithmetic progressions in sumsets*, Acta Arithmetica 60 (1991), no 2, 191-202.

[Ru2]. I. Ruzsa, *Generalized arithmetic progressions and sumsets*, Acta Math. Hungar. 65 (1994), no 4, 379-388.

[Ru3]. I. Ruzsa, *An analog of Freiman's theorem in groups*, in 'Structure Theory of Set Addition', Astérisque 258 (1999).

[S-T]. F. Szemerédi, W. Trotter, *Extremal problems in Discrete Geometry*, Combinatorica, 3 (3-4): 387–392 (1983).

University of California, Department of Mathematics, Riverside, CA 92521