PRODUCT THEOREMS IN SL_2 AND SL_3

[1]Mei-Chu Chang

Abstract We study product theorems for matrix spaces. In particular, we prove the following theorems.

Theorem 1. For all $\varepsilon > 0$, there is $\delta > 0$ such that if $A \subset SL_3(\mathbb{Z})$ is a finite set, then either A intersects a coset of a nilpotent subgroup in a set of size at least $|A|^{1-\varepsilon}$, or $|A^3| > |A|^{1+\delta}$.

Theorem 2. Let A be a finite subset of $SL_2(\mathbb{C})$. Then either A is contained in a virtually abelian subgroup, or $|A^3| > c|A|^{1+\delta}$ for some absolute constant $\delta > 0$.

Here $A^3 = \{a_1a_2a_3 : a_i \in A, i = 1, 2, 3\}$ is the 3-fold product set of A.

§0. Introduction.

The aim of this paper is to establish product theorems for matrix spaces, in particular $SL_2(\mathbb{Z})$ and $SL_3(\mathbb{Z})$. Applications to convolution inequalities will appear in a forthcoming paper.

Recall first Tits' Alternative for linear groups G over a field of characteristic 0: Either G contains a free group on two generators or G is virtually solvable (i.e. contains solvable subgroup of finite index). For a solvable group G, one has to distinguish further the cases G not virtually nilpotent and G with a nilpotent subgroup of finite index. Also in the solvable non-virtually nilpotent case, G is of exponential growth. In particular, G contains a 'free semi-group' on two generators. (See [Ti].)

The 'growth' here refers to the size of the balls

$$B_\Gamma(n) = \{x \in G : d_\Gamma(x, e) \leq n\}$$

[1]partially supported by NSF.
2000 Mathematics Subject Classification. 05A99, 15A99, 05C25; 20G40, 20D60, 11B75 .
Keywords: product theorem, Subspace theorem, trace amplification, nilpotent group.

where d_Γ refers to the distance on the Caley graph associated to a given finite set of generators Γ of G, and e is the identity of G.

Uniform statements on the exponential growth were obtained recently in the work of Eskin-Mozes-Oh [EMO] and Breuillard [B].

Nilpotent groups are of polynomial growth. This explains the exponential versus polynomial growth dichotomy for linear groups. (See [G].)

Here we are interested in the amplification of large subsets A of G under a few product operations, thus

$$|A^n| > |A|^{1+\varepsilon},$$

where

$$A^n = A \cdots A = \{a_1 \cdots a_n : a_i \in A\}$$

is the *n-fold product set* of A. (it is known that if a bounded n suffices, then already $n = 3$ will do, cf. [T] or Proposition 1.6).

From previous growth dichotomy discussion, such "product-phenomenon" may not be expected in nilpotent groups. For instance, let A be the following subset of the Heisenberg group

$$A = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} : \ a, b, c \in \mathbb{Z}, \ a, b \in [1, N], \ c \in [1, N^2] \right\}.$$

Then

$$|A| \sim N^4 \sim |A^3|.$$

Our main result is the following:

Theorem 1. *For all $\varepsilon > 0$, there is $\delta > 0$ such that if $A \subset SL_3(\mathbb{Z})$ is a finite set, then one of the following alternatives holds.*

(i) *A intersects a coset of a nilpotent subgroup in a set of size at least $|A|^{1-\varepsilon}$.*

(ii) $|A^3| > |A|^{1+\delta}$.

The main tool involved in the proof is the Subspace Theorem by Evertse, Schlickewei, and Schmidt. (cf. [ESS])

Moreover, we also rely essentially on some techniques introduced by H. Helfgott in the study of the product phenomenon in groups $SL_2(\mathbb{Z}_p)$ and earlier work of the author [C].

Let us point out that generalizing Theorem 1 to $SL_d(\mathbb{Z})$ is quite feasible using the same type of approach. One can further replace $\mathbb{Z}$ by the integers in a given algebraic number field K, $[K : \mathbb{Q}] < \infty$, but the case $SL_3(\mathbb{R})$ would be more problematic (because of the use of the Subspace Theorem).

On the other hand, an easy adaptation of Helfgott's methods permits us to show:

Theorem 2. *Let A be a finite subset of $SL_2(\mathbb{C})$. Then one of the following alternatives holds.*

(i) *A is contained in a virtually abelian subgroup*

(ii) $|A^3| > c|A|^{1+\delta}$ *for some absolute constant $\delta > 0$.*

This last result applies in particular for finite subsets $A \subset F_2 < SL_2(\mathbb{Z})$, where F_2 is the free group on two generators.

Here, the first alternative reads now

(i') A is contained in a cyclic subgroup of F_2.

There should be a direct combinatorial proof of this, possibly providing more information on δ.

The results obtained in this paper belong to the general research area of arithmetic combinatorics. In particular, obtaining general sum-product theorems and product theorems in certain abelian or non-abelian groups has been an active research topic in recent years. Besides the scalar fields of real and complex numbers, these problems have been investigated in characteristic p (in prime fields $\mathbb{F}_p$ and their Cartesian products $\mathbb{F}_p \times \mathbb{F}_p$) and also in residue classes $\mathbb{Z}/n\mathbb{Z}$ under various assumptions on n. Among the different motivations and implications of those results, one should certainly mention the estimates of certain exponential sums (see [BGK], [BC]) over small multiplicative subgroups and the applications to pseudo-randomness problems in computer science. (see [BIW], [BKSSW]). It turns out that sum-product results in the commutative case permit one to obtain product theorems in certain non-abelian setting. In a remarkable paper [H], H. Helfgott proves that if $A \subset SL_2(\mathbb{Z}_p)$ is not contained in a proper subgroup and $|A| < p^{3-\varepsilon}$, then $|A^3| > |A|^{1+\delta}$ with $\delta = \delta(\varepsilon)$. Generalizing Helfgott's results to higher dimensions remains unsettled at this point. In this paper we consider the corresponding problem in characteristic zero. For $SL_3(\mathbb{Z})$ this question is easier. Our main result depends however on the Subspace Theorem. It is not clear how to elaborate a counterpart of this approach in characteristic p. It would be quite interesting to find a different method to prove our result.

The paper is organized as follows:

Section 1 consists of some preliminary material for n by n matrices and some elementary facts about sum-product sets. In Section 2 we give a technical proposition about sets of traces of elements in $GL_3(\mathbb{C})$. In Section 3 we state the version of the Subspace Theorem which we will use. In Sections 4 and 6 we give the proofs of Theorem 1 and Theorem 2 respectively. In Section 5 we give a different proof (using Subspace Theorem) of the version of Theorem 2 for $SL_2(\mathbb{Z})$ (though Theorem 5.1 clearly follows from Theorem 2).

Notations. When working on n-fold sum-product sets, sometimes it is more convenient to consider symmetric sets or even sets involving few products. Hence we define

$$A^{[n]} = (\{1\} \cup A \cup A^{-1})^n.$$

We use A^n for both the n-fold product set and n-fold Cartesian product when there is no ambiguity.

The *n-fold sum set* of A is $nA = A + \cdots + A = \{a_1 + \cdots + a_n : a_1, \cdots, a_n \in A\}$. The *difference set* $A - A$ and the *inverse set* A^{-1} can be defined similarly.

For a matrix g, $\mathrm{Tr}(g)$ is the trace of g. Therefore, Tr can be viewed as a function on $Mat_n(\mathbb{C})$.

Note that the properties under consideration (e.g. the size of a set of matrices or the trace of a matrix) are invariant under base change (i.e. conjugation by an invertible matrix).

We follow the trend that ε, (respectively, δ, or C) may represent various constants, even in the same setting. Also, $f(x) \sim g(x)$ means $f(x) = cg(x)$ for some constant c which may depend on some other parameters.

§1. Preliminaries.

Lemma 1.1. *Let $A \subset GL_n(\mathbb{C})$ be finite. Then there is a subset $A' \subset A$ of size $|A'| > |A|^{1-\varepsilon}$ such that for any $\tilde{g} \in Mat_n(\mathbb{C})$, one of the following alternatives holds.*

(i) $Tr\left(\tilde{g}(A' - A')\right) = \{0\}$.

(ii) $|Tr\,(\tilde{g}A')| > |A|^{\delta}$ *for any* $\delta < 1 - \varepsilon$.

Proof. Let $V < Mat_n(\mathbb{C})$ be a linear space of the smallest dimension for which there is a subset $A' \subset A$ so that

$$|A'| > |A|^{1-\varepsilon} \tag{1.1}$$

for some $\varepsilon > 0$, and

$$A' - A' \subset V. \tag{1.2}$$

It is clear that V exists, since a decreasing sequence of subspaces of $Mat_n(\mathbb{C})$ has length at most n^2+1.

Take $\tilde{g} \in Mat_n(\mathbb{C})$. Assume (ii) fails. Then there is $z \in \mathbb{C}$ such that

$$|\{g \in A' : Tr\,(\tilde{g}g) = z\}| \geq \frac{|A'|}{|Tr\,(\tilde{g}A')|} \geq \frac{|A'|}{|A|^\delta} > |A|^{1-\varepsilon-\delta}.$$

Denote

$$A'' = \{g \in A' : Tr\,(\tilde{g}g) = z\}. \tag{1.3}$$

By (1.2) and (1.3), we have

$$A'' - A'' \subset V \cap \{g \in Mat_n(\mathbb{C}) : \mathrm{Tr}\ (\tilde{g}g) = 0\} =: W.$$

From the minimality assumption on V, it follows that $V = W$ and from (1.2)

$$Tr\,\big(\tilde{g}(A' - A')\big) \subset Tr\,(\tilde{g}V) = \{0\}.$$

Hence (i) holds. $\square$

Lemma 1.2. *Let $A \subset GL_n(\mathbb{C})$ be finite. Assume*

$$\forall\, \tilde{g} \in A^{[n-1]},\ \ Tr\,\big(\tilde{g}(A - A)\big) = \{0\}. \tag{1.4}$$

Then for any $g \in A - A$, the eigenvalues of g are zero and $g^n = 0$.

Proof. Let g be an element in $A - A$. Then for $i = 1, \dots, n$, the matrix g^{i-1} is a linear combination of elements in $A^{[n-1]}$. Hence assumption (1.4) implies that

$$\mathrm{Tr}\ g^i = 0, \quad \text{for } i = 1, \dots, n. \tag{1.5}$$

There is $b \in GL_n(\mathbb{C})$ to put g in the upper triangular form.

$$\bar{g} := b^{-1}gb = \begin{pmatrix} g_{11} & g_{12} & g_{13} & \cdots \\ 0 & g_{22} & g_{23} & \\ 0 & 0 & g_{33} & \\ \vdots & & & \end{pmatrix} \tag{1.6}$$

It follows from (1.5) that $\mathrm{Tr}\ \bar{g} = \mathrm{Tr}\ \bar{g}^2 = \cdots = \mathrm{Tr}\ \bar{g}^n = 0$. Namely,

$$\sum_{i=1}^n g_{ii} = \sum_{i=1}^n g_{ii}^2 = \cdots = \sum_{i=1}^n g_{ii}^n = 0.$$

We claim that $g_{ii} = 0$ for all $1 \leq i \leq n$.

Assume not. Let $\{\lambda_i\}_{1\leq i\leq m}$ be the set of distinct elements in $\{g_{ii}\}_{1\leq i\leq n}\backslash\{0\}$ and $a_i \geq 1$ be the corresponding multiplicities. Then

$$\sum_{i=1}^{m} a_i\lambda_i = \sum_{i=1}^{m} a_i\lambda_i^2 = \cdots = \sum_{i=1}^{m} a_i\lambda_i^n = 0.$$

This means the vectors

$$\begin{pmatrix} \lambda_1 \\ \lambda_1^2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_1^n \end{pmatrix}, \cdots, \begin{pmatrix} \lambda_m \\ \lambda_m^2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_m^n \end{pmatrix}$$

are linearly dependent. A contradiction follows.

Therefore in (1.6), $\bar{g}^n = 0$, which implies $g^n = 0$. We proved that all elements of $A - A$ have zero eigenvalues, hence are nilpotent. □

Remark 1.2.1. Assumption (1.4) implies that Tr $(A - A)^{\leq n} = \{0\}$.

Remark 1.2.2. It is clear that the proof only needs condition (1.5) rather than assumption (1.4).

Next, we will study sets consisting of matrices of rank ≤ 1.

We recall that, via the identification $Mat_n(\mathbb{C}) \simeq Hom(\mathbb{C}, \mathbb{C}) \simeq \mathbb{C}^{n\vee} \otimes \mathbb{C}^n$, for a rank one matrix $g \in Mat_n(\mathbb{C})$, there exist $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \in \mathbb{C}^n$ such that

$$g = y^\vee \otimes x = \big(x_i y_j\big)_{1\leq i,j\leq n} \tag{1.7}$$

Lemma 1.3. *Let $B \subset Mat_n(\mathbb{C}) \simeq \mathbb{C}^{n\vee} \otimes \mathbb{C}^n$ be a finite set satisfying the property that for any $g \in B$, rank $g \leq 1$ and*

$$Tr(B^2) = \{0\}. \tag{1.8}$$

Then there exist $\bar{g} = \bar{y}^\vee \otimes \bar{x} \in B\backslash\{0\}$ and a subset $\bar{B} \subset B$ such that $|\bar{B}| > \frac{1}{2}|B|$ and for all $g = y^\vee \otimes x \in \bar{B}$

$$y \cdot \bar{x} = \sum_{i=1}^{n} \bar{x}_i y_i = 0$$

Proof. For any $g = y^\vee \otimes x, g' = {y'}^\vee \otimes x' \in B$, (1.7) and (1.8) imply

$$0 = \operatorname{Tr} gg' = \sum_{i,j} x_i y_j x'_j y'_i = \Big(\sum_i x_i y'_i\Big)\Big(\sum_j x'_j y_j\Big).$$

Hence either $x \cdot y' = \sum x_i y'_i = 0$, or $x' \cdot y = \sum x'_i y_i = 0$. The lemma follows from the following fact. □

Fact 1.4. *Let B be a set with $|B| = N$ and let $\sim$ be a relation on B. Assume that for any $b, b' \in B$, either $b \sim b'$ or $b' \sim b$. Then there exist $\bar{b} \in B$ and $\bar{B} \subset B$ such that $|\bar{B}| \geq \frac{1}{2}|B|$ and for all $b \in \bar{B}$, we have $b \sim \bar{b}$.*

Proof. For $b \in B$, denote

$$B_b^+ = |\{b' : b' \sim b\}|, \text{ and } B_b^- = |\{b' : b \sim b'\}|.$$

Then

$$\sum_b {B_b}^+ + \sum_b {B_b}^- = N^2 \text{ and } \sum_b {B_b}^+ = \sum_b {B_b}^-.$$

Hence $\sum_b {B_b}^+ = \sum_b {B_b}^- = \frac{N^2}{2}$ and there exists b such that $B_b^+ \geq \frac{N}{2}$.

For the rest of the section we will recall some general facts for sum-product sets. Specifically, we are interested in the quantitative growth of n-fold product sets or sum-product sets.

Fact 1.5. (Ruzsa's triangle inequality)

Let A, B, C be finite subsets of an abelian group $\langle G, \cdot \rangle$. Then

$$|AB| \leq \frac{|AC|\ |C^{-1}B|}{|C|}$$

Proposition 1.6. *Let A be a finite subset of an abelian group $\langle G, \cdot \rangle$ and let*

$$S = A \cup A^{-1}. \tag{1.9}$$

Assume

$$|A^3| < K|A|. \tag{1.10}$$

Then

$$|S^n| = K^{c(n)}|S|, \tag{1.11}$$

with $c(n) \leq 3(n-2)$.

Proof. Ruzsa's triangle inequality implies

$$|A^2A^{-1}| \leq \frac{|A^2A|\,|A^{-1}A^{-1}|}{|A|} < K^2|A| \tag{1.12}$$

$$|A^{-1}A^2| \leq \frac{|A^{-1}A^{-1}|\,|AA^2|}{|A|} < K^2|A| \tag{1.13}$$

$$|AA^{-1}A| \leq \frac{|AA^{-1}A^{-1}|\,|AA|}{|A|} < K^3|A| \tag{1.14}$$

(We also use (1.12) for the second inequality in (1.14).)

Therefore,

$$|S^3| < K^3|S|. \tag{1.15}$$

Assume $|S^n| = K^{c(n)}|S|$ with $c(n) \leq 3(n-2)$. Then by Ruzsa's triangle inequality, induction and (1.15)

$$|S^{n+1}| \leq \frac{|S^{n-1}S|\,|SS^2|}{|S|} \leq K^{c(n)}K^3|S|.$$

Hence

$$c(n+1) \leq c(n) + 3.$$

On the other hand, $c(3) \leq 3$. $\square$

Lemma 1.7. *Let A be a finite subset of a ring $\langle R; +, \cdot\rangle$. Then for $n = 2^k$ we have*

$$|2^nS^n| = |S\ |^{c(n)}$$

with $c(n) > n^{\log_2(\frac{5}{4})}$.

Proof. We will prove by induction on k. Assume $|2^nS^n| = |S\ |^{c(n)}$ with $c(n) > n^{\log_2(\frac{5}{4})}$. By the sum-product theorem in $\mathbb{C}$, either

$$|2^nS^n + 2^nS^n| > |2^nS^n|^{\frac{5}{4}}$$

or

$$|2^nS^n \cdot 2^nS^n| > |2^nS^n|^{\frac{5}{4}}.$$

Therefore, we have

$$|S\ |^{c(2n)} = |2^{2n}S^{2n}| > |2^nS^n|^{\frac{5}{4}} > (|S\ |^{c(n)})^{\frac{5}{4}}$$

and

$$c(2n) > c(n)\ 2^{\log_2(\frac{5}{4})} > n^{\log_2(\frac{5}{4})}\ 2^{\log_2(\frac{5}{4})} = (2n)^{\log_2(\frac{5}{4})}. \qquad \square$$

Proposition 1.8. *Let S be a finite subset of a ring $\langle R; +, \cdot\rangle$. Then for any $a_1, \ldots, a_{2^k}$ in R,*

$$|a_1 S^k + \ldots + a_{2^k} S^k| > |S|^{b(k)},$$

where $b(k) \to \infty$ as $k \to \infty$. In fact, $\log b(k) \sim \log k$.

Proof. First, we note that for sets A, B, by Ruzsa's triangle inequality (on addition), we have

$$|A - A| \leq \frac{|A+B|^2}{|B|}.$$

Hence

$$|A + B| \geq |A - A|^{1/2} |B|^{1/2}. \tag{1.16}$$

Claim. Let $A_1, \ldots, A_{2^k}$ be subsets of R. We take $s \sim \log_2 k$ and $\ell \sim \frac{k}{s} \sim \frac{k}{\log k}$. Then

$$|A_1 + \cdots + A_{2^k}| > \min_j |2^\ell (A_j - A_j)|^{\frac{1}{2}}.$$

Applying (1.16), we have

$$\begin{aligned} |A_1 + \cdots + A_{2^k}| \geq & |A_1 + \cdots + A_{2^{k-1}}|^{1/2} \\ & |(A_{2^{k-1}+1} - A_{2^{k-1}+1}) + \cdots + (A_{2^k} - A_{2^k})|^{1/2}. \end{aligned} \tag{1.17}$$

Repeating s times, we see that the right-hand side of (1.17) is bounded below by

$$\begin{aligned} & |A_1 + \cdots + A_{2^{k-s}}|^{1/2^s} \, |(A_{2^{k-s}+1} - A_{2^{k-s}+1}) + \cdots + (A_{2^{k-s+1}} - A_{2^{k-s+1}})|^{1/2^s} \\ & |(A_{2^{k-s+1}+1} - A_{2^{k-s+1}+1}) + \cdots + (A_{2^{k-s+2}} - A_{2^{k-s+2}})|^{1/2^{s-1}} \\ & \ldots \\ & |(A_{2^{k-2}+1} - A_{2^{k-2}+1}) + \cdots + (A_{2^{k-1}} - A_{2^{k-1}})|^{1/2^2} \\ & |(A_{2^{k-1}+1} - A_{2^{k-1}+1}) + \cdots + (A_{2^k} - A_{2^k})|^{1/2} \end{aligned} \tag{1.18}$$

which is bounded further by

$$\min_{j_1 \leq 2^{k-1}} \left| (A_{j_1+1} - A_{j_1+1}) + \cdots + (A_{j_1+2^{k-s}} - A_{j_1+2^{k-s}}) \right|^{(1-\frac{1}{2^s})}. \tag{1.19}$$

(This estimate is very rough. We omit the first absolute value completely. As for the other absolute values we take only the first 2^{k-s} differences. Therefore, $j_1 \in \{2^{k-s}, 2^{k-s+1}, \ldots, 2^{k-1}\}$.)

Repeating the process on the sets $A_{j_1+1} - A_{j_1+1}, \dots, A_{j_1+2^{k-s}} - A_{j_1+2^{k-s}}$, we have

$$\begin{aligned}&\left|(A_{j_1+1} - A_{j_1+1}) + \cdots + (A_{j_1+2^{k-s}} - A_{j_1+2^{k-s}})\right| \\ &> \min_{j_2 \le 2^{k-s-1}} \left|2(A_{j_2+1} - A_{j_2+1}) + \cdots + 2(A_{j_2+2^{k-2s}} - A_{j_2+2^{k-2s}})\right|^{(1-\frac{1}{2^s})}.\end{aligned}$$

Iterating $\ell + 1$ times, we have

$$\begin{aligned}|A_1 + \cdots + A_{2^k}| &> \min_j |2^\ell (A_j - A_j)|^{(1-\frac{1}{2^s})^{\ell+1}} \\ &> \min_j |2^\ell (A_j - A_j)|^{(1-\frac{1}{k})^{\frac{k}{\log k}}} \\ &> \min_j |2^\ell (A_j - A_j)|^{1/2}.\end{aligned}$$

Taking $A_j = a_j S^k$ in the Claim, by our choice of ℓ and Lemma 1.7, we have

$$\begin{aligned}|a_1 S^k + \ldots + a_{2^k} S^k| >& |(2^\ell (S^k - S^k)| \\ >& |2^\ell S^\ell| \\ =& |S|^{c(\ell)}.\end{aligned}$$

Let $b(k) = c(\ell)$. Then $\log b(k) = \log c(\ell) \sim \log(\ell) \sim \log k$. $\square$

§2. The set of traces.

This is a variant of Helfgott's result.

Proposition 2.1. *Let $A \subset GL_3(\mathbb{C})$ be a finite set. Then one of the following alternatives holds.*

(i) *There is a subset A' of A, $|A'| > |A|^{1-\varepsilon}$ which is contained in a coset of a nilpotent subgroup.*

(ii) *There is some $\tilde{g} \in A^{[3]}$ such that*

$$|Tr\,(\tilde{g}A)| > |A|^\delta.$$

Proof. Assume (ii) fails. Namely, we assume that for all $\tilde{g} \in A^{[3]}$, $|Tr\,(\tilde{g}A)| \le |A|^\delta$. Lemma 1.1 implies that there exists $A' \subset A$,

$$|A'| > |A|^{1-\varepsilon} \tag{2.1}$$

such that

$$\forall \tilde{g} \in A^{[3]},\ Tr\left(\tilde{g}(A' - A')\right) = \{0\}. \tag{2.2}$$

Fix some element $\xi \in A'$ and let

$$B = A' - \xi \subset A' - A'.$$

Then (2.2) and Remark 1.2.1 imply

$$\text{Tr } (B^2) = \{0\}. \tag{2.3}$$

We consider two cases.

Case 1. $g^2 = 0$ for all $g \in B$.

Claim 1. rank $g \leq 1$.

Indeed, Lemma 1.2 and (2.2) imply that g has the following upper triangular form.

$$\bar{g} = b^{-1}gb = \begin{pmatrix} 0 & g_{12} & g_{13} \\ 0 & 0 & g_{23} \\ 0 & 0 & 0 \end{pmatrix} \tag{2.4}$$

for some $b = b(g) \in GL_3(\mathbb{C})$. The assumption $g^2 = 0$ implies that

$$\bar{g}^2 = \begin{pmatrix} 0 & 0 & g_{12}g_{23} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0.$$

Thus either $g_{12} = 0$ or $g_{23} = 0$ and g is of rank at most 1.

Claim 2. After suitable changes of bases, there is a subset $\bar{\bar{B}}$ of B, $|\bar{\bar{B}}| \geq \frac{1}{4}|B|$, consisting of matrices of the form

$$\begin{pmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ * & 0 & 0 \end{pmatrix}. \tag{2.5}$$

Proof of Claim 2.

We apply Lemma 1.3 to find some $\bar{y}^\vee \otimes \bar{x} \in B \backslash \{0\}$ and a subset $\bar{B} \subset B$ such that $|\bar{B}| > \frac{1}{2}|B|$ and for all $y^\vee \otimes x \in \bar{B}$

$$\bar{x} \cdot y = \sum_{i=1}^{3} \bar{x}_i y_i = 0. \tag{2.6}$$

An appropriate base charge (e.g. change $\bar{x}$ to the standard base $\vec{e}_3$) permits us then to ensure that

$$y_3 = 0$$

for all $y^\vee \otimes x \in \bar{B}$ with $y = (y_1, y_2, y_3)$.

Repeating the preceding, we have for any $g = y^\vee \otimes x, g' = y'^\vee \otimes x' \in \bar{B}$

$$(x \cdot y')(x' \cdot y) = \Big(\sum_{i=1,2} x_i y_i'\Big)\Big(\sum_{j=1,2} x_j' y_j\Big) = 0.$$

Hence we may apply Lemma 1.3 again on $\bar{B}$ to find $\bar{\bar{g}} = \bar{\bar{y}}^\vee \otimes \bar{\bar{x}} \in \bar{B}\backslash\{0\}$ and a subset $\bar{\bar{B}} \subset \bar{B}, |\bar{\bar{B}}| > \frac{1}{2}|\bar{B}|$, such that for all $y^\vee \otimes x \in \bar{\bar{B}}$

$$\bar{\bar{x}} \cdot y = \sum_{i=1}^{2} \bar{\bar{x}}_i y_i = 0. \tag{2.7}$$

A further base change permits us to ensure that also

$$y_2 = 0$$

for any $g = y^\vee \otimes x \in \bar{\bar{B}}$, which therefore has the form

$$g = \vec{e}_1^\vee \otimes x.$$

Again, (2.2) implies

$$x_1 = \text{ Tr } g = 0.$$

and g has the form in (2.5) and Claim 2 is proved.

Write

$$\xi + \bar{\bar{B}} = \xi(1 + \xi^{-1}\bar{\bar{B}}) \subset A'. \tag{2.8}$$

The elements of $\xi^{-1}\bar{\bar{B}}$ are still of the form (2.5) since they are of zero-trace by (2.2).

Hence $1 + \xi^{-1}\bar{\bar{B}} \subset N$, where N is the nilpotent group

$$\begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{pmatrix}.$$

Recalling (2.1), $|\bar{\bar{B}}| > \frac{1}{4}|A|^{1-\varepsilon}$ and we therefore showed that A intersects a coset of a nilpotent subgroup in a set of size at least $|A|^{1-\varepsilon}$.

Case 2: There is some $h \in B$ with $h^2 \neq 0$.

We do a base change so that h has the upper triangular form

$$h = \begin{pmatrix} 0 & h_{12} & h_{13} \\ 0 & 0 & h_{23} \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.9}$$

Hence,

$$h^2 = \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ where } a = h_{12}h_{23} \neq 0.$$

For $g = (g_{ij}) \in B$

$$0 = \text{ Tr } (h^2 g) = a g_{31},$$

hence

$$g_{31} = 0.$$

Also

$$0 = \text{ Tr } \left(h^2 g^2\right) = a(g^2)_{31},$$

hence

$$g_{32} g_{21} = 0.$$

Therefore, either $g_{32} = 0$, or $g_{21} = 0$. Assume that more elements $g \in B$ have $g_{32} = 0$. (The other case is similar.) Let $\bar{B} \subset B, |\bar{B}| \geq \frac{1}{2}|B|$ be a subset such that

$$g_{31} = g_{32} = 0 \text{ for } g \in \bar{B}.$$

Next, recalling (2.9), write for $g \in \bar{B}$,

$$0 = \text{Tr } (hg) = h_{12} g_{21} = 0,$$

hence also $g_{21} = 0$.

Thus the elements $g \in \bar{B}$ satisfy

$$g_{21} = g_{31} = g_{32} = 0$$

and recalling (2.2) and Lemma 1.2,

$$g_{11} = g_{22} = g_{33} = 0.$$

Hence the elements of $\bar{B}$ are strictly upper triangular

$$g = \begin{pmatrix} 0 & g_{12} & g_{13} \\ 0 & 0 & g_{23} \\ 0 & 0 & 0 \end{pmatrix}$$

Denote

$$\zeta = \xi^{-1}.$$

By (2.2) again

$$0 = \text{ Tr } (\zeta h^2) = \zeta_{31}$$

and for $g \in \bar{B}$

$$\text{Tr}\,(\zeta g) = 0 = \text{ Tr } \big((\zeta g)^2\big). \tag{2.10}$$

Since

$$\zeta g = \begin{pmatrix} 0 & \zeta_{11} g_{12} & \zeta_{11} g_{13} + \zeta_{12} g_{23} \\ 0 & \zeta_{21} g_{12} & \zeta_{21} g_{13} + \zeta_{22} g_{23} \\ 0 & 0 & \zeta_{32} g_{23} \end{pmatrix}$$

(2.10) implies

$$\begin{aligned} &\zeta_{21} g_{12} + \zeta_{32} g_{23} = 0 \\ &(\zeta_{21} g_{12})^2 + (\zeta_{32} g_{23})^2 = 0 \end{aligned}$$

and

$$\zeta_{21} g_{12} = \zeta_{32} g_{23} = 0.$$

Therefore $\zeta \bar{B}$ are strictly upper triangular and

$$\xi + \bar{B} = \xi(1 + \zeta \bar{B}) \subset A \cap \xi N.$$

The conclusion is the same as in Case 1. $\square$

Remark 2.1.1. More generally, previous argument shows that if $A \subset GL_3(\mathbb{C})$ is a finite set and M large, then one of the following holds.

(1) There is $\tilde{g} \in A^{[3]}$ such that $|Tr\,(\tilde{g} A)| > M$,

(2) There is a subset A' of $A, |A'| > M^{-C}|A|$ (C an absolute constant) such that A' is contained in a coset of a nilpotent subgroup.

Remark 2.1.2. The preceding remains valid for $\mathbb{C}$ replaced by a finite field.

§3. Some applications of the Subspace Theorem.

Our main tool to prove Theorem 1 is the finiteness theorem of Evertse, Schlickewei, and Schmidt which we state here in a form convenient for later purpose.

Theorem 3.1. [ESS] *Let $G < \langle \mathbb{C}^*, \cdot \rangle$ be a multiplicative group of rank r, and let $a_1, a_2, \dots, a_t \in \mathbb{C}$. One may then associate to each subset $S \subset \{1, \dots, t\}$ with $|S| \geq 2$, a subset $\mathcal{C}_S \subset \mathbb{C}^{|S|} = \mathbb{C} \times \cdots \times \mathbb{C}$ of size*

$$|\mathcal{C}_S| < C(r, t) \tag{3.1}$$

such that the following holds.

Let $x = (x_1, \dots, x_t) \in G^t = G \times \cdots \times G$ be a solution of the equation

$$a_1 x_1 + \cdots + a_t x_t = 0.$$

Then there is a partition $\pi = \{\pi_\alpha\}$ of $\{1, \dots, t\}$ such that $|\pi_\alpha| \geq 2$ and for each α there is an element $y \in \mathcal{C}_{\pi_\alpha}$ such that $(x_j)_{j \in \pi_\alpha}$ is a scalar multiple of y.

There is the following corollary.

Lemma 3.2. *Let G be as in Theorem 3.1 and fix an integer $t \geq 2$. Let $a_1, \dots, a_{2t} \in \mathbb{C} \backslash \{0\}$. There is a set $E \subset \mathbb{C}$ depending on $a_1, \dots, a_{2t}$,*

$$|E| < C(r, t) \tag{3.2}$$

such that the following holds.

Let $\mathcal{A}$ be a finite subset of $G^t = G \times \cdots \times G$ and such that

$$\frac{x_i}{x_j} \notin E \text{ for all } x \in \mathcal{A} \text{ and } 1 \leq i \neq j \leq t. \tag{3.3}$$

Then

$$|\{(x, x') \in \mathcal{A} \times \mathcal{A} : a_1 x_1 + \cdots + a_t x_t = a_{t+1} x'_1 + \cdots + a_{2t} x'_t\}| < C(r, t) |\mathcal{A}|. \tag{3.4}$$

Proof. Apply Theorem [ESS] to the equation

$$a_1 x_1 + \cdots + a_t x_t - a_{t+1} x_{t+1} - \cdots - a_{2t} x_{2t} = 0, \tag{3.5}$$

where we denoted $x' = (x_{t+1}, \dots, x_{2t})$.

Let $\mathcal{C}_S, S \subset \{1, \dots, 2t\}$ with $|S| \geq 2$ be the corresponding systems. Define

$$E = \bigcup_{S \subset \{1, \dots, 2t\}} \Big\{ \frac{z_i}{z_j} : z \in \mathcal{C}_S,\ 1 \leq i \neq j \leq t, \text{ or } \ t+1 \leq i \neq j \leq 2t \Big\}. \tag{3.6}$$

If (3.5) holds, there is a partition $\{\pi_\alpha\}$ of $\{1,\dots,2t\}$ such that for each α there is an element $y\in\mathcal{C}_{\pi_\alpha}$ with

$$\frac{x_i}{x_j}=\frac{y_i}{y_j} \text{ for } i,j\in\pi_\alpha. \tag{3.7}$$

If we assume (3.3), then $|\pi_\alpha\cap\{1,\dots,t\}|\le 1$ and $|\pi_\alpha\cap\{t+1,\dots,2t\}|\le 1$. Hence $|\pi_\alpha|=2$ and π_α intersects both $\{1,\dots,t\}$ and $\{t+1,\dots,2t\}$ in one element. Since $\{\pi_\alpha\}$ is a partition of $\{1,\dots,2t\}$, it follows from (3.7) that given $x=(x_1,\dots,x_t)$, the element $x'=(x_{t+1},\dots,x_{2t})$ will be determined up to $t!|E|^t<C(r,t)$ possibilities. This proves Lemma 3.2.

Hence, we also have:

Lemma 3.3. *Let G be as in Theorem 3.1. Given $a_1,\dots,a_t\in\mathbb{C}\backslash\{0\}$, there is a subset $E\subset\mathbb{C}$ with $|E|<C(r,t)$, such that if $\mathcal{A}$ is a finite subset of $G^t=G\times\cdots\times G$ and*

$$\frac{x_i}{x_j}\notin E \text{ for all } x=(x_s)_s\in\mathcal{A} \text{ and } 1\le i\ne j\le t \tag{3.8}$$

then

$$\Big|\Big\{\sum_{s=1}^t a_s x_s : x\in\mathcal{A}\Big\}\Big|>\frac{1}{C(r,t)}|\mathcal{A}|. \tag{3.9}$$

Proof.

Denote $R=\{\sum a_s x_s : x\in\mathcal{A}\}$ and let for $z\in\mathbb{C}$

$$n(z)=|\{x\in\mathcal{A}:\sum a_s x_s=z\}|.$$

Then

$$|\mathcal{A}|=\sum_{z\in R}n(z)\le|R|^{1/2}\Big[\sum_{z\in R}n(z)^2\Big]^{1/2}<C(r,t)|R|^{1/2}|\mathcal{A}|^{1/2}$$

by (3.4). Therefore

$$|R|>\frac{1}{C(r,t)}|\mathcal{A}|$$

and (3.9) holds.

§4. The proof of Theorem 1.

We specialize further $A\subset SL_3(\mathbb{Z})$ not satisfying alternative (i) of Theorem 1.

Hence by Proposition 2.1,

$$|Tr\,(\tilde{g}A)|>|A|^\theta \tag{4.1}$$

for some $\theta > 0$ and $\tilde{g} \in A^{[3]}$.

We assume

$$|A^3| < |A|^{1+\delta} \tag{4.2}$$

with the aim to reach a contradiction for δ small. (In any case $\delta < \theta/21$. cf. (4.6) and (4.19))

Assumption (4.2) implies that for any given $s \in \mathbb{Z}_+$

$$|A^{[s]}| < |A|^{1+\delta_s} \tag{4.3}$$

where $\delta_s \leq 3(s-2)\delta$. (See [T] or Proposition 1.6.)

We will now repeat an argument due to H. Helfgott [H].

First, we will find a large subset of $A^{-1}A$ consisting of simultaneously diagonalizable matrices.

Denote $T = Tr\,(\tilde{g}A)$ and let for each $\tau \in T$ an element $g_\tau \in A$ be specified such that

$$\mathrm{Tr}\,(\tilde{g}g_\tau) = \tau. \tag{4.4}$$

Claim 1. There are $g_1, g_\tau \in A$ and $A_1 \subset A$ with $|A_1| > |A|^\theta$ such that $g_1^{-1}A_1$ is contained in the centralizer of $\tilde{g}g_\tau$.

Proof. Since the conjugacy classes

$$C_\tau = \{g\tilde{g}g_\tau g^{-1} : g \in A\} \subset A^{[6]}$$

are disjoint and in view of (4.1) and (4.3) we may specify $\tau \in T\backslash\{3, -1\}$ such that

$$|C_\tau| < \frac{|A|^{1+\delta_6}}{|T|} < |A|^{1+\delta_6-\theta}. \tag{4.5}$$

Therefore there exists some $g_1 \in A$ such that

$$|\{g \in A : g\tilde{g}g_\tau g^{-1} = g_1\tilde{g}g_\tau g_1^{-1}\}| \geq \frac{|A|}{|C_\tau|} > |A|^{\theta-\delta_6}. \tag{4.6}$$

(Here δ_6 is negligible, since we can take δ as small as we like.)

Let $A_1 = \{g \in A : g\tilde{g}g_\tau g^{-1} = g_1\tilde{g}g_\tau g_1^{-1}\}$. Thus for $g \in A_1$

$$(g_1^{-1}g)(\tilde{g}g_\tau) = (\tilde{g}g_\tau)(g_1^{-1}g), \tag{4.7}$$

which means that the elements of $g_1^{-1}A_1 \subset A^{-1}A$ commute with $\tilde{g}g_\tau$. $\square$

We will need the following elementary fact from algebra.

Fact 4.1. Let $f(x) \in \mathbb{Z}[x]$ be a monic cubic polynomial over $\mathbb{Z}$. Then either $f(x)$ is irreducible over $\mathbb{Q}$ and has three distinct roots, or one of the roots is in $\mathbb{Q}$ and the other two roots are quadratic conjugates, or $f(x)$ has three roots in $\mathbb{Q}$. Hence if the constant term of $f(x)$ is -1, the only possible multiple roots are $1, 1, 1$ or $1, -1, -1$.

Let K be the splitting field of the characteristic polynomial $\det(\tilde{g}g_\tau - \lambda)$ of $\tilde{g}g_\tau$. Since $\det(\tilde{g}g_\tau - \lambda)$ has degree 3, we have $[K : \mathbb{Q}] \leq 6$. The eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of $\tilde{g}g_\tau$ are distinct, because by (4.4), $\lambda_1 + \lambda_2 + \lambda_3 = \tau \notin \{3, -1\}$. Therefore $\tilde{g}g_\tau$ is diagonalizable over the extension field K of $\mathbb{Q}$. With this basis, the commutativity property

$$h\tilde{g}g_\tau = \tilde{g}g_\tau h \text{ for } h \in g_1^{-1}A_1$$

implies $h_{ij}\lambda_j = \lambda_i h_{ij}$, hence $h_{ij} = 0$ for $i \neq j$.

We have obtained a subset

$$D = g_1^{-1}A_1 \subset A^{-1}A$$

of simultaneously diagonalizable elements, where by Claim 1

$$|D| > |A|^\theta. \tag{4.8}$$

We use the basis introduced above with which the elements of D are diagonal.

According to Fact 4.1, for the elements $g \in D$, there are two possibilities. Either the eigenvalues $\lambda_i(g), 1 \leq i \leq 3$ form a system of conjugate algebraic units, or $\{1, -1\} \cap \{\lambda_i(g) : i = 1, 2, 3\} \neq \phi$ and the other two eigenvalues are conjugate quadratic units. We assume the first alternative (the second may be handled similarly and is in fact easier).

For $g \in D$, denote

$$\Lambda(g) = \{\lambda_1(g), \lambda_2(g), \lambda_3(g)\} \subset K. \tag{4.9}$$

Let O_K be the ring of integers of K. Thus $\Lambda(g)$ is contained in the unit group of O_K which is of rank ≤ 5. This will allow us to exploit Theorem ESS (see §3) to reach a contradiction to (4.2). Also, $\Lambda(g) \cap \Lambda(g') = \phi$, if $g \neq g'$.

We claim that there is an element $h \in A$ for which there are two nonzero entries in the same row.

Indeed, otherwise for any $h \in A$ there is exactly one nonzero entry in each row and in each column. Therefore A is contained in the union of the six cosets of the diagonal subgroup. This would violate our assumption that A fails alternative (i) in Theorem 1.

Fix such an element h. Assume for instance

$$h_{12} \neq 0, h_{13} \neq o$$

(the other cases are similar).

Fix $\ell \in \mathbb{Z}_+$ and consider the following set

$$D(hD)^{\ell-1} \subset A^{-1}A(AA^{-1}A)^{\ell-1} \tag{4.10}$$

consisting of elements

$$g = g^{(1)}hg^{(2)}h \cdots hg^{(\ell)}, \text{ where } g^{(1)}, \cdots, g^{(\ell)} \in D. \tag{4.11}$$

Recall that each $g^{(s)} \in D$ is diagonal with diagonal elements $\Lambda(g^{(s)}) = \{\lambda_1(g^{(s)}), \lambda_2(g^{(s)}), \lambda_3(g^{(s)})\}$ forming a system of conjugate units in O_K. By (4.11)

$$\sum_{i,j} g_{ij} = \sum_{i_1,\dots,i_\ell} h_{i_1i_2}h_{i_2i_3} \cdots h_{i_{\ell-1}i_\ell}\lambda_{i_1}(g^{(1)})\lambda_{i_2}(g^{(2)}) \cdots \lambda_{i_\ell}(g^{(\ell)}), \tag{4.12}$$

which we view as a polynomial in $\lambda_i(g^{(s)}) \in G$, where $g^{(s)} \in D$, with $1 \leq s \leq \ell$ and $i = 1, 2, 3$.

Denote $\{a_1, \dots, a_t\}$ the non-vanishing coefficients

$$a_s = h_{i_1i_2} \cdots h_{i_{\ell-1}i_\ell} \neq 0 \tag{4.13}$$

in (4.12). We note that

$$t \leq 3^\ell. \tag{4.14}$$

We will apply Lemma 3.3 to the linear form $\sum_{1\leq s\leq t} a_s x_s, x_s \in G$. The set $\mathcal{A} \subset G^t = G \times \cdots \times G$ will consist of elements of the form

$$x = (x_s)_{1\leq s\leq t} \text{ where } x_s = \lambda_{i_1}(g^{(1)}) \cdots \lambda_{i_\ell}(g^{(\ell)}).$$

Here the index s corresponds to the multi-index $(i_1, \dots, i_\ell)$ such that (4.13) holds and $g^{(1)}, \dots, g^{(\ell)}$ range in D. (Note that $g^{(1)}, \dots, g^{(\ell)}$ stay the same for the same x.) The elements of $\mathcal{A}$ also satisfy condition (3.8) of Lemma 3.3.

Claim 2.

$$|D(hD)^{\ell-1}| > c(\ell)|\mathcal{A}|.$$

Proof. First, we observe

Fact 4.2. Let $D \subset GL_3(\mathbb{C})$ be a set of diagonal matrices obtained from a subset of $SL_3(\mathbb{Z})$ after base change. Then given any $z \in \mathbb{C}$, for $i \neq j$, there are at most four elements $g \in D$ for which

$$\frac{\lambda_i(g)}{\lambda_j(g)} = z, \tag{4.15}$$

where $\lambda_i(g)$ and $\lambda_j(g)$ are the eigenvalues of g.

In fact, if (4.15) holds for elements $g, g' \in D$, then since g, g' are diagonal, we have

$$\lambda_i(g^{-1}g') = \lambda_j(g^{-1}g').$$

Fact 4.1 implies that the eigenvalues of $g^{-1}g'$ are either $1, 1, 1$, or $1, -1, -1$. Hence $g^{-1}g'$ can only be one of the following matrices.

$$1, \begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix}, \begin{pmatrix} -1 & & \\ & 1 & \\ & & -1 \end{pmatrix}, \begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix}.$$

This shows that for given $z \in \mathbb{C}$, (4.15) may only hold for at most four elements of D.

Nest we examine condition (3.8).

Let s, s' be different multi-indices $(i_1, \dots, i_\ell)$ and $(i'_1, \dots, i'_\ell)$. Thus

$$\frac{x_s}{x_{s'}} = \frac{\lambda_{i_1}(g^{(1)}) \cdots \lambda_{i_m}(g^{(m)})}{\lambda_{i'_1}(g^{(1)}) \dots \lambda_{i'_m}(g^{(m)})} \notin E \tag{4.16}$$

where $i_m \neq i'_m$ and $i_{m+1} = i'_{m+1}, \dots, i_\ell = i'_\ell$.

Given $g^{(1)}, \dots, g^{(m-1)}$, we view (4.16) as a condition on $g^{(m)}$. The issue amounts to considering for some $z \in \mathbb{C}$ the elements $g \in D$ for which

$$\frac{\lambda_i(g)}{\lambda_j(g)} = z, \text{ where } i \neq j.$$

By Fact 4.2, it is now clear that condition (4.16) will be satisfied if we remove from D^ℓ a subset $\mathcal{D} \subset D^\ell$ where, by (4.14)

$$|\mathcal{D}| \leq 4\ell |D|^{\ell-1}|E| < C(\ell)|D|^{\ell-1}.$$

Together with Claim 3 below, we can then conclude from (3.9) that

$$|D(hD)^{\ell-1}| > \frac{1}{C(\ell)}|\mathcal{A}|. \qquad \square$$

Claim 3.

$$|\mathcal{A}| > C(\ell)\ |D|^{\ell}$$

Proof. We will show that the size of a fiber of the map

$$D^{\ell}\backslash\mathcal{D} = D \times \cdots \times D\backslash\mathcal{D} \to \mathcal{A} \text{ given by } (g^{(1)}, \ldots, g^{(\ell)}) \mapsto (x_s)_s$$

is bounded by 4^{ℓ}.

Recall that h satisfies $h_{12} \neq 0, h_{13} \neq 0$.

We proceed as follows.

First we note that there exist $i_1, \ldots, i_{\ell-2}$ such that $h_{i_1 i_2} \cdots h_{i_{\ell-2}1} \neq 0$. Indeed, since h is an invertible matrix, at least one of h_{11}, h_{21}, h_{31} is nonzero. For instance, if $h_{21} \neq 0$, we can take $i_{\ell-2} = 2, i_{\ell-3} = 1, i_{\ell-4} = 2$ etc. Let $i_1, \ldots, i_{\ell-2}$ be such indices, and let $s = (i_1, \ldots, i_{\ell-2}, 1, 2)$, and $s' = (i_1, \ldots, i_{\ell-2}, 1, 3)$. Then $h_s, h_{s'} \neq 0$ and (4.13) holds for $a_s, a_{s'}$. Hence for given $x = (x_s)_s \in \mathcal{A}$,

$$\frac{x_s}{x_{s'}} = \frac{\lambda_2(g^{(\ell)})}{\lambda_3(g^{(\ell)})}$$

determines the ratio $\frac{\lambda_2(g^{(\ell)})}{\lambda_3(g^{(\ell)})}$. By Fact 4.2, this essentially specifies $g^{(\ell)}$ (up to multiplicity 4).

Next, take $i_1, \ldots, i_{\ell-3}$ and i_ℓ, i'_ℓ such that

$$h_{i_1 i_2} \cdots h_{i_{\ell-3}1} \neq 0 \text{ and } h_{2i_\ell} \neq 0, h_{3i'_\ell} \neq 0.$$

Let $s = (i_1, \ldots, i_{\ell-3}, 1, 2, i_\ell)$ and $s' = (i_1, \ldots, i_{\ell-3}, 1, 3, i'_\ell)$. Then

$$\frac{x_s}{x_{s'}} = \frac{\lambda_2(g^{(\ell-1)})}{\lambda_3(g^{(\ell-1)})}\ \frac{\lambda_{i_\ell}(g^{(\ell)})}{\lambda_{i'_\ell}(g^{(\ell)})}. \tag{4.17}$$

Since $g^{(\ell)}$ has already been specified, (4.17) allows us to determine also $g^{(\ell-1)}$ (up to multiplicity 4). Continuing, we see that (x_s) indeed determines $(g^{(1)}, \ldots, g^{(\ell)})$ up to multiplicity 4^{ℓ}. Therefore

$$|\mathcal{A}| > 4^{-\ell}|D^{\ell}\backslash\mathcal{D}| > \frac{4^{-\ell}}{2}|D|^{\ell} \qquad \square$$

Putting Claim 2 and Claim3 together, we proved that

$$|D(hD)^{\ell-1}| \gtrsim |D|^{\ell}. \tag{4.18}$$

From (4.3), (4.10), (4.18) and (4.8), this implies

$$|A|^{1+\delta_{3\ell-1}} > |A|^{\ell\theta} \tag{4.19}$$

leading to a contradiction for ℓ large enough (since δ is very small).

This concludes the argument.

Remark 4.3. To see that our result is almost the optimum, we consider the following example.

Fix large integers M and N.

Consider the set $\mathcal{M} = \{\sigma \in SL_2(\mathbb{Z}) : \ \sigma_{i,j} \leq M\}$, hence

$$|\mathcal{M}| \sim M^2.$$

Let $A \subset SL_3(\mathbb{Z})$ consisting of elements of the form

$$g = \left(\begin{array}{c|c} \sigma & \begin{matrix} x \\ y \end{matrix} \\ \hline 0\ 0 & 1 \end{array}\right), \tag{4.20}$$

where $\sigma \in \mathcal{M}$ and $x, y \in \mathbb{Z}, |x|, |y| \leq N$. Thus

$$|A| \sim M^2 N^2.$$

Clearly for g, g' of the form (4.20), we have

$$gg' = \left(\begin{array}{c|c} \sigma\sigma' & \sigma \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix} \\ \hline 0\ 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} \tilde{\sigma} & \begin{matrix} z \\ w \end{matrix} \\ \hline 0\ 0 & 1 \end{array}\right)$$

where $\tilde{\sigma} \in \mathcal{M}^2$ and $|z|, |w| \lesssim MN$.

Therefore

$$A^3 \subset \left(\begin{array}{c|c} \mathcal{M}^3 & \mathcal{M}^2\mathcal{N} + \mathcal{M}\mathcal{N} + \mathcal{N} \\ \hline 0\ 0 & 1 \end{array}\right),$$

where

$$\mathcal{N} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{Z}, |x|, |y| \le N \right\}.$$

Hence

$$|A^3| \lesssim M^6 M^4 N^2 < M^8 \, |A|.$$

By construction, the intersection of A and the coset of a nilpotent group is at most of size $\sim \frac{|A|}{M}$. Given $\varepsilon > 0$, choose N large enough to ensure $M \sim |A|^\varepsilon$. Hence $|A^3| < |A|^{1+8\varepsilon}$, proving that $\delta \le 8\varepsilon$ in Theorem 1.

Remark 4.4. It is likely that the result and proof of Theorem 1 admits a generalization to $A \subset SL_n(\mathbb{Z})$ for arbitrary n.

§5. Product theorem for $SL_2(\mathbb{Z})$.

We may carry out the preceding argument in the 2-dimensional case for finite subsets $A \subset SL_2(\mathbb{Z})$. We show the following dichotomy (compare with Helfgott's theorem for $A \subset SL_2(\mathbb{Z}_p)$).

Theorem 5.1. *Let A be a finite subset of $SL_2(\mathbb{Z})$. Then one of the following alternatives holds.*

(i) A is contained in a virtually abelian subgroup.

(ii) $|A^3| > c|A|^{1+\delta}$, *for some absolute constant $\delta > 0$.*

We outline the argument.

First, since $\det(g) = 1$ for $g \in SL_2(\mathbb{Z})$, we note that $\mathrm{Tr}(g) = \pm 2$ if and only if the characteristic polynomial $\det(g - \lambda)$ has multiple roots and the two eigenvalues of g are $1, 1$ or $-1, -1$.

Assume neither (i) nor (ii) holds.

Claim 1. There is an element $\xi \in A^{[2]}$ for which

$$\mathrm{Tr}\, \xi \ne 2, -2. \tag{5.1}$$

Proof. Assume there is none.

Take $\tilde{g} \in A \backslash \{1, -1\}$. In appropriate basis, $\tilde{g}$ has the Jordan form

$$\tilde{g} = \begin{pmatrix} \varepsilon & 1 \\ 0 & \varepsilon \end{pmatrix} \text{ with } \varepsilon = \pm 1.$$

Let

$$h = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in A,$$

hence

$$h^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} \in A^{-1}$$

and

$$\text{Tr } \tilde{g}h = \varepsilon\alpha + \gamma + \varepsilon\delta = \varepsilon \text{ Tr } h + \gamma$$
$$\text{Tr } \tilde{g}h^{-1} = \varepsilon\delta - \gamma + \varepsilon\alpha = \varepsilon \text{ Tr } h - \gamma.$$

Therefore,

$$\text{Tr } \tilde{g}h + \text{Tr } \tilde{g}h^{-1} = 2\varepsilon \text{ Tr } h.$$

From our assumption that Tr h, Tr $\tilde{g}h$, Tr $\tilde{g}h^{-1} \in \{2, -2\}$, we have Tr $\tilde{g}h$=Tr $\tilde{g}h^{-1}$. Hence $\gamma = 0$ and

$$A \subset \left\{ \begin{pmatrix} \varepsilon & \beta \\ 0 & \varepsilon \end{pmatrix} : \varepsilon = \pm 1 \right\}$$

contradicting the failure of (i). □

Thus we take $\xi \in A^{[2]}$ with Tr $\xi \neq \pm 2$ and choose a basis over a quadratic extension field K of $\mathbb{Q}$ as to make ξ diagonal

$$\xi = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \qquad \lambda \neq \pm 1.$$

We will work over this basis.

Fix another element $\zeta = \begin{pmatrix} \zeta_{11} & \zeta_{12} \\ \zeta_{21} & \zeta_{22} \end{pmatrix} \in A$ which is not diagonal.

Claim 2.

$$\max(|Tr\, \xi A|, |Tr\, \xi^2 A|, |Tr\, \zeta A|) \gtrsim |A|^{1/3}$$

Proof. Assume say $\zeta_{12} \neq 0$.

For $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in A$

$$\text{Tr } \xi g = \lambda a + \lambda^{-1} d \tag{5.2}$$

$$\text{Tr } \xi^2 g = \lambda^2 a + \lambda^{-2} d \tag{5.3}$$

$$\text{Tr } \zeta g = \zeta_{11} a + \zeta_{12} c + \zeta_{21} b + \zeta_{22} d. \tag{5.4}$$

Assume Tr ξg, Tr $\xi^2 g$, Tr ζg given. From (5.2), (5.3), a and d are specified and from (5.4), we obtain $\zeta_{12} c + \zeta_{21} b$, hence b and c (up to multiplicity 2), since $ad - bc = 1$.□

Consequently we reached (4.1) with $\theta = \frac{1}{3}$ and $\tilde{g} \in A^{[4]}$.

Next, apply again Helfgott's argument to produce a set $D \subset A^{-1}A$ of simultaneously diagonalizable elements over a quadratic extension field K of $\mathbb{Q}$, $|D| \gtrsim |A|^{1/3}$. Proceeding as before for $A \subset SL_3(\mathbb{Z})$, use Lemma 3.3 and the subsequent construction to contradict the assumption $|A^3| < |A|^{1+\delta}$. The only additional ingredient needed is an element $h \in A$ with at least three nonzero entries. If there is no such element, then A would be contained in the virtually abelian group

$$\left\{ \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix} : \lambda \in U_K \right\} \cup \left\{ \begin{pmatrix} 0 & \lambda \\ -\frac{1}{\lambda} & 0 \end{pmatrix} : \lambda \in U_K \right\}$$

contradicting the failure of (i).

This proves Theorem 5.1.

Let F_k be the free group generated on k generators. Since $SL_2(\mathbb{Z})$ contains a subgroup isomorphic to F_2 (in fact of finite index) and F_2 has a subgroup isomorphic to F_k for all $k \geq 1$, Theorem 5.1 has the following implication.

Corollary 5.2. *There is an absolute constant $\delta > 0$ such that the following holds.*

Let A be a finite subset of the free group F_2 (or $F_k, k \geq 2$) which is not contained in a cyclic group. Then

$$|A^3| > c|A|^{1+\delta}. \tag{5.5}$$

It would be interesting to have a direct combinatorial proof of this fact.

§6. The proof of Theorem 2.

In the present situation, it is not clear how to involve the Subspace Theorem. Rather, for most of the proof, we will follow Helfgott's $SL_2(\mathbb{Z}_p)$ argument. The main digression in the preceding argument, compared with Helfgott's approach, was the use of the Subspace Theorem rather than the trace-amplification technique from [H].

Assume (i), (ii) both fail. Returning to the proof of Theorem 5.1, Claim 1 and Claim 2 may be reproduced also in the present situation. Thus there is $\tilde{g} \in A^{[4]}$ such that

$$|Tr\, \tilde{g} A| \gtrsim |A|^{1/3}. \tag{6.1}$$

This gives again a subset $D \subset A^{-1}A$ of diagonal elements (in the same basis), with

$$|D| > |A|^{1/3}. \tag{6.2}$$

Let

$$\mathcal{D} = \left\{ \lambda : \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix} \in D \right\} \tag{6.3}$$

Take further an element $h \in A$ which is neither diagonal nor off-diagonal in this basis (which is possible since we assume (i) fails).

Let

$$h = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}.$$

We distinguish several cases.

Case 1: $h_{11}h_{22} = 1, h_{21} = 0 \qquad$ (or $h_{12} = 0$).

Hence h is upper triangular

$$h = \begin{pmatrix} a & b \\ 0 & \frac{1}{a} \end{pmatrix}, \quad \text{where } ab \neq 0.$$

For any $\mu_1, \dots, \mu_{r-1} \in \mathcal{D}^r$, we write the following element in $hD^r(hD^rD^{-r})^{r-2}hD^{-r}$ as

$$h \begin{pmatrix} \mu_{r-1} & 0 \\ 0 & \frac{1}{\mu_{r-1}} \end{pmatrix} h \begin{pmatrix} \frac{\mu_{r-2}}{\mu_{r-1}} & 0 \\ 0 & \frac{\mu_{r-1}}{\mu_{r-2}} \end{pmatrix} h \begin{pmatrix} \frac{\mu_{r-3}}{\mu_{r-2}} & 0 \\ 0 & \frac{\mu_{r-2}}{\mu_{r-3}} \end{pmatrix} \cdots h \begin{pmatrix} \frac{\mu_1}{\mu_2} & 0 \\ 0 & \frac{\mu_2}{\mu_1} \end{pmatrix} h \begin{pmatrix} \frac{1}{\mu_1} & 0 \\ 0 & \mu_1 \end{pmatrix}$$

$$= \begin{pmatrix} a^r & b\big(a^{r-1}\mu_1^2 + a^{r-3}\mu_2^2 + \cdots + \frac{1}{a^{r-3}}\, \mu_{r-1}^2 + \frac{1}{a^{r-1}}\big) \\ 0 & \frac{1}{a^r} \end{pmatrix}. \tag{6.4}$$

We see that

$$\begin{aligned} |A|^{1+\delta_{4r^2-3r}} \geq& |AD^r(AD^rD^{-r})^{r-2}AD^{-r}| \\ \geq& \Big| \big\{ a^{r-1}\mu_1^2 + a^{r-3}\mu_2^2 + \cdots + \frac{1}{a^{r-3}}\mu_{r-1}^2 : \mu_1, \dots, \mu_{r-1} \in \mathcal{D}^r \big\} \Big|. \end{aligned} \tag{6.5}$$

Since $|\mathcal{D}| > |A|^{1/3}$, (6.5) clearly contradicts Proposition 1.8. (e.g. we first choose r large enough such that $\frac{1}{3}c(r) > 2$ then δ small such that $\delta_{4r^2-3r} < 1$.)

Case 2: $h_{12}h_{21} = -1, h_{22} = 0$ (or $h_{11} = 0$). Thus

$$h = \begin{pmatrix} a & b \\ -\frac{1}{b} & 0 \end{pmatrix}, \quad \text{where } ab \neq 0.$$

Taking some $\lambda \in \mathcal{D}$, we write

$$h \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix} h = \begin{pmatrix} \lambda a^2 - \frac{1}{\lambda} & \lambda ab \\ -\lambda \frac{a}{b} & -\lambda \end{pmatrix}.$$

Appropriate choice of λ will provide an element $h' \in A^{[4]}$ with four nonzero entries. This brings us to

Case 3: h has four nonzero entries.

In this situation, we apply Helfgott's trace amplification argument.

Denote $D_1 = D \cup D^{-1}$ and consider the subset of $D_1^4 h D_1^4 h$ of elements

$$g_{xy} = \begin{pmatrix} xy & 0 \\ 0 & \frac{1}{xy} \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} \frac{x}{y} & 0 \\ 0 & \frac{y}{x} \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$$

$$= \begin{pmatrix} x^2 h_{11}^2 + y^2 h_{12} h_{21} & * \\ * & \frac{1}{y^2} h_{12} h_{21} + \frac{1}{x^2} h_{22}^2 \end{pmatrix}$$

with $x, y \in \left(\mathcal{D} \cup \mathcal{D}^{-1}\right)^2$.

Hence

$$\text{Tr } g_{xy} = h_{11}^2 x^2 + h_{22}^2 x^{-2} + h_{12} h_{21} (y^2 + y^{-2})$$

and

$$Tr(D_1^4 h D_1^4 h) \supset \left\{ h_{11}^2 x^2 + h_{22}^2 x^{-2} + h_{12} h_{21}(y^2 + y^{-2}) : x, y \in \left(\mathcal{D} \cup \mathcal{D}^{-1}\right)^2 \right\}.$$

We claim that

$$\left| Tr\Big(\big((A^{-1}A)^4 A\big)^2 \Big) \right| \geq |Tr(D_1^4 h D_1^4 h)| > |D|^{1+\gamma} \tag{6.6}$$

for some absolute constant $\gamma > 0$. This is a consequence of the sum-product theorem in $\mathbb{C}$. Assume (6.6) fails. it would follow that

$$\left| \left\{ h_{11}^2 x^2 + h_{22}^2 x^{-2} : x \in \left(\mathcal{D} \cup \mathcal{D}^{-1}\right)^2 \right\} + h_{12} h_{21} \left\{ y^2 + \frac{1}{y^2} : y \in \left(\mathcal{D} \cup \mathcal{D}^{-1}\right)^2 \right\} \right| < |\mathcal{D}|^{1+\gamma}, \tag{6.7}$$

for any $\gamma > 0$.

Denote

$$S_1 = \left\{ y^2 + \frac{1}{y^2} : y \in \mathcal{D} \cup \mathcal{D}^{-1} \right\}$$

and

$$S_2 = \left\{ y^2 + \frac{1}{y^2} : y \in \left(\mathcal{D} \cup \mathcal{D}^{-1} \right)^2 \right\}.$$

Then

$$|S_1| \sim |\mathcal{D}| \tag{6.8}$$

and the Plunnecke-Ruzsa inequality and (6.7) imply that

$$|S_2 + S_2| < |\mathcal{D}|^{1+3\gamma}. \tag{6.9}$$

Since clearly

$$S_1 S_1 \subset S_2 + S_2$$

and

$$|S_1 + S_1| \leq |S_2 + S_2|,$$

(6.8) and (6.9) indeed contradict the sum-product theorem in $\mathbb{C}$.

Hence (6.6) holds.

Replacing A by $\tilde{A} = \left((A^{-1}A)^4 A \right)^2$, we obtain a new set $\tilde{D} \subset (\tilde{A})^{-1}\tilde{A}$ of simultaneously diagonal elements (in another basis), for which

$$|\tilde{\mathcal{D}}| > |D|^{1+\gamma} > |A|^{\frac{1}{3}+\frac{\gamma}{3}}.$$

Go again through Cases 1, 2, 3.

In Case 1, we obtain a contradiction.

In Cases 2 and 3, a further trace amplification is achieved. Eventually a contradiction is reached. This proves Theorem 2.

References

[BIW]. B. Barak, R. Impagliazzo, A. Wigderson, *Extracting randomness using few independent sources*, Proc of the 45th FOCS (2004), 384-393.

[BKSSW]. B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, A. Wigderson, *Simulating Independence: New Constructions of Condensers, Ramsey Graphs, Dispersers, and Extractors*, STOC (to appear).

[BC]. J. Bourgain, M-C. Chang, *On the size of k-fold Sum and Product Sets of Integers*, J. Amer. Math. Soc., 17, No. 2, (2003), 473-497.

[BGK]. J. Bourgain, A. Glibichuk, S. Konyagin, *Estimate for the number of sums and products and for exponential sums in fields of prime order*, submitted to J. London MS.

[B]. E. Breuillard, *On Uniform exponential growth for solvable groups*, (preprint).

[C]. M-C. Chang, *Sum and product of different sets*, Contributions to Discrete Math, Vol 1, 1 (2006), 57-67.

[EMO]. A. Eskin, S. Mozes, H. Oh, *On Uniform exponential growth for linear groups*, Invent. 160, (2005), 1-30.

[ESS]. J.-H. Evertse, H. Schlickewei, W. Schmidt, *Linear equations in variables which lie in a multiplicative group*, Annals Math 155, (2002), 807-836.

[G]. M. Gromov, *Groups of polymonial growth and expanding maps*, IHES, 53, (1981), 53-73.

[H]. H. Helfgott, *Growth and generation in* $SL_2(\mathbb{Z}/\mathbb{Z}_p)$, Annals (to appear).

[T]. T. Tao, *Product set estimates in non-commutative groups*, math.CO/0601431.

[TV]. T. Tao, V. Vu, *Additive Combinatorics,*, Cambridge University Press (to appear).

[Ti]. J. Tits, *Free subgroups in linear groups*, J. Algebra 20, (1972), 250-270.

Mathematics Department, University of California, Riverside CA 92521

E-mail address: `mcc@math.ucr.edu`