ACADEMIC PRESS

Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT®

Journal of Functional Analysis 207 (2004) 444–460

JOURNAL OF Functional Analysis

http://www.elsevier.com/locate/jfa

On problems of Erdös and Rudin

Mei-Chu Chang

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

Received 1 November 2002; revised 5 January 2003; accepted 10 January 2003

Communicated by J. Bourgain

Abstract

A well-known conjecture of W. Rudin is that the set of squares is a $\wedge_p$-set for all $p>4$. In particular, this implies that for all $\varepsilon>0$, there exists a constant c_ε such that

$$\left(\int_\Pi \left|\sum_{j=1}^{k} e^{in_j^2\lambda}\right|^4 dx\right)^{\frac{1}{4}} \leqslant c_\varepsilon k^{\frac{1}{2}} + \varepsilon$$

for any k distinct integers $n_1 \ldots n_k$. In this article we give a combinatorial interpretation of the inequality above in the spirit of $\|q\|q$ sum and product sets along graphs as considered by P. Erdös and E. Szemeredi (Studies in Pure Mathematics, pp. 213–218). We also show that the left-hand side of the inequality is bounded by $C_\varepsilon \dfrac{k^{\frac{3}{4}}}{(\log k)^{\frac{1}{48}-\varepsilon}}$.

Keywords: Rudin; λ–p conjecture; Squares; Sumset; Product set

1. Introduction

Let

$$S = \{n^2 \mid n \in \mathbb{N}\} \tag{1.1}$$

be the set of squares. A well-known (and still open) conjecture of Rudin is that S is a $\wedge_p$-set for all $p<4$. This means that for all $p<4$, there is a constant c_p such that for

E-mail address: meichu.chang@ucr.edu.

0022-1236/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0022-1236(03)00073-9

all finite scalar sequences $(a_n)_{n\in\mathbb{N}}$,

$$\left(\int_{\Pi}\left|\sum a_n\, e^{in^2x}\right|^p dx\right)^{\frac{1}{p}} \leqslant c_p\left(\sum |a_n|^2\right)^{\frac{1}{2}}. \tag{1.2}$$

Here Π denotes the usual circle group. In 'Harmonic Analysis language', the problem is thus whether $L^p_S(\Pi) = L^2_S(\Pi)$ if $p<4$ and S as above. Presently, there is no exponent $p>2$ known for which $L^p_S(\Pi) = L^2_S(\Pi)$ holds; see [Ru].

Rudin's problem implies an affirmative answer to the following question:

For all $\varepsilon>0$, does there exist a constant c_ε such that

$$\left(\int_{\Pi}\left|\sum_{j=1}^{k} e^{in_j^2x}\right|^4 dx\right)^{\frac{1}{4}} < c_\varepsilon k^{\frac{1}{2}+\varepsilon} \tag{1.3}$$

for any k distinct integers $n_1, \ldots, n_k$?

Our purpose here is to give a combinatorial interpretation of (1.3) in the spirit of 'sum and product sets along graphs' as considered in [E-S] by Erdös and Szemeredi.

We recall the setup.

Let $A = \{a_i\in\mathbb{Z} \mid a_i<a_j, \text{ if } i<j\}$ be a set of n distinct integers and $G\subset\{(i,j) \mid i,j\in\mathbb{Z}, 1\leqslant i,j\leqslant n\}$ a graph.

Denote

$$Sum_G\, A = \{a_i + a_j \mid (i,j)\in G\}, \tag{1.4}$$

$$Diff_G\, A = \{a_i - a_j \mid (i,j)\in G\}, \tag{1.5}$$

$$Prod_G\, A = \{a_i a_j \mid (i,j)\in G\}. \tag{1.6}$$

In [E-S], there is the following:

Conjecture E-S. *For all $\alpha>0, \varepsilon>0$, $n\gg 0$, and $A\subset\mathbb{Z}$, one has the inequality*

$$|Sum_G\, A| + |Prod_G\, A| > c_\varepsilon |G|^{1-\varepsilon} \tag{1.7}$$

for some constant c_ε, provided $|G|>n^{1+\alpha}$ and n is large enough.

(*) We use the notation $|A|$ for the cardinality of a (finite) set A.

Remark 1. Conjecture E-S may be extended to the case of real numbers $a_1, \ldots, a_n\in\mathbb{R}$.

Erdös also formulates the following:

Conjecture E. *If* $|G|>cn$, *then for* $a_1,\dots,a_n\in\mathbb{Z}$, *inequality* (1.7) *holds*.

Remark 2. Conjecture E fails for $a_1,\dots,a_n\in\mathbb{R}$.

Concerning the validity of (1.7), only partial results are known. It can, for instance, be shown that (1.7) holds, if we assume $|G|>\delta n^2$ and $|Sum_G A|<Cn$ for $A\subset\mathbb{Z}$ or $\mathbb{R}$ (see [Ch3]), or $|Prod_G A|<C'n$ for $A\subset\mathbb{Z}$ (see [Ch2]), where $0<\delta$, and $C, C'<\infty$ are arbitrary constants. One may also obtain some information (again assuming $|G|$ large) from Elekes' method based on the Szemeredi–Trotter theorem (see [E1]).

We are able to show the following:

Proposition 1. *Conjecture* E *implies* (1.3).

Proposition 2. *Inequality* (1.3) *is equivalent to the following statement*:

$$|Sum_G A|\cdot|Diff_G A|\cdot|Prod_G A|>c_\varepsilon|G|^{2-\varepsilon}\quad \text{for all } \varepsilon>0,\ \text{and } A\subset\mathbb{Z}. \tag{1.8}$$

Remark 3. The following bound is obvious:

$$\left(\int_\Pi\left|\sum_{j=1}^k e^{in_j^2x}\right|^4 dx\right)^{\frac14}\leqslant k^{\frac34}. \tag{1.9}$$

Indeed,

$$\int_\Pi\left|\sum_{j=1}^k e^{in_j^2x}\right|^4 dx\leqslant k^2\int_\Pi\left|\sum_{j=1}^k e^{in_j^2x}\right|^2 dx=k^2k,$$

because of $|\sum_{j=1}^k e^{in_j^2x}|\leqslant k$ and Parseval's identity.

We will show the following slight (but far from trivial) improvement of (1.9):

Proposition 3. *For any* k *distinct integers* $n_1,\dots,n_k$, *there is the bound*

$$\left(\int_\Pi\left|\sum_{j=1}^k e^{in_j^2x}\right|^4 dx\right)^{\frac14}<C_\varepsilon\frac{k^{\frac34}}{(\log k)^{\frac{1}{48}-\varepsilon}} \tag{1.10}$$

for all $\varepsilon>0$.

Along the lines of Proposition 2, (1.10) implies

Corollary 4. *For* $A \subset \mathbb{Z}$,

$$\max\{|Sum_G A|, |Diff_G A|, |Prod_G A|\} > c_\varepsilon |G|^{\frac{1}{2}} (\log |G|)^{\frac{1}{48}-\varepsilon}. \tag{1.11}$$

Remark 4. Corollary 4 fails, if $A \subset \mathbb{R}$. This is shown by the following example: Let A be the set

$$\{\sqrt{i} \pm \sqrt{j} \mid 1 \leqslant i < j \leqslant m, \text{ and } i,j \text{ are square free}\}$$

on which we consider the graph

$$G = \{(\sqrt{i} + \sqrt{j}, \sqrt{i} - \sqrt{j}) \mid 1 \leqslant i < j \leqslant m, \text{ and } i,j \text{ are square free}\}.$$

Thus $k = |G| \sim m^2$, while clearly

$$Sum_G A \subset \{2\sqrt{i} \mid 1 \leqslant i \leqslant m\},$$

$$Diff_G A \subset \{2\sqrt{j} \mid 1 \leqslant j \leqslant m\},$$

$$Prod_G A \subset \{i - j \mid 1 \leqslant i, j \leqslant m\}.$$

Hence, all sets have cardinality $< 2m$. □

2. Proofs of Propositions 1 and 2

To show (1.3), we first define

$$F_m = \{(j_1, j_2) \mid 1 \leqslant j_1, j_2 \leqslant k, \ m = n_{j_1}^2 - n_{j_2}^2\}. \tag{2.1}$$

Applying Parseval's identity to the left-hand side of (1.3), we have

$$\int_\Pi \left| \sum_{j=1}^{k} e^{in_j^2 x} \right|^4 dx = \int_\Pi \left| \sum_{m \in \mathbb{Z}} |F_m| e^{imx} \right|^2 dx = \sum_{m \in \mathbb{Z}} |F_m|^2. \tag{2.2}$$

Then, to relate $\{n_j\}_j$ to the notations (1.4)–(1.6), we define

$$A = \{n_{j_1} \pm n_{j_2} \mid 1 \leqslant j_1, j_2 \leqslant k\} = \{a_1 < a_2 < \cdots < a_n\},$$

where

$$n = |A|.$$

For integers $0 \leqslant l \leqslant \log_2 k$, denote

$$M_l = \{m \in \mathbb{Z} \mid 2^l \leqslant |F_m| < 2^{l+1}\}. \tag{2.3}$$

Thus,

$$\sum_l |M_l| = |\{m \varepsilon \mathbb{Z} | F_m \neq \phi\}| \leqslant k^2. \tag{2.4}$$

(The inequality follows from definition (2.1) of F_m.)

For each M_l we define the corresponding graph G_l on $\{1, 2, \ldots, n\}$ as follows:

$$\begin{aligned}(i_1, i_2) \in G_l \;\Leftrightarrow\; &\text{there exist } 1 \leqslant j_1,\ j_2 \leqslant k,\\ &\text{such that } a_{i_1} = n_{j_1} + n_{j_2}, a_{i_2} = n_{j_1} - n_{j_2}, n_{j_1}^2 - n_{j_2}^2 \in M_l.\end{aligned} \tag{2.5}$$

Hence,

$$|Sum_{G_l} A| \leqslant k, \quad |Diff_{G_l} A| \leqslant k, \quad |Prod_{G_l} A| \leqslant |M_l|, \tag{2.6}$$

while clearly

$$|G_l| = \sum_{m \in M_l} |F_m| \tag{2.7}$$

and

$$|G_l| \leqslant k^2. \tag{2.8}$$

Claim. (1.7) *implies* (1.3).

Proof. Applying (1.7) to the graph G_l with $\frac{\varepsilon}{2}$, and using (2.6) and (2.8), we have that for all $\varepsilon > 0$,

$$k + |M_l| > c_\varepsilon |G_l|^{1-\frac{\varepsilon}{2}} > c_\varepsilon k^{-\varepsilon} |G_l|. \tag{2.9}$$

From (2.7) and (2.9), we have

$$\sum_{m \in M_l} |F_m| < C_\varepsilon k^{1+\varepsilon} + C_\varepsilon k^\varepsilon |M_l|, \tag{2.10}$$

where $C_\varepsilon = (c_\varepsilon)^{-1}$.

On the other hand, (2.3) and (2.10) give

$$2^l |M_l| \leqslant \sum_{m \in M_l} |F_m| \leqslant C_\varepsilon k^{1+\varepsilon} + C_\varepsilon k^\varepsilon |M_l|, \tag{2.11}$$

hence,

$$2^l|M_l| < 2C_\varepsilon k^{1+\varepsilon}, \quad \text{if } 2^l > 2C_\varepsilon k^\varepsilon. \tag{2.12}$$

Therefore, we have

$$\begin{aligned}\sum_{m\in\mathbb{Z}} |F_m|^2 &\leqslant \sum_{0\leqslant l\leqslant \log_2 k} |M_l|(2^{l+1})^2 = \sum_{2^l\leqslant 2C_\varepsilon k^\varepsilon} 4(2^l)^2|M_l| + \sum_{2^l > 2C_\varepsilon k^\varepsilon} 4(2^l|M_l|)2^l \\ &\leqslant 4(2C_\varepsilon k^\varepsilon)^2 k^2 + 4(2C_\varepsilon k^{1+\varepsilon}) \sum_{0\leqslant l\leqslant \log_2 k} 2^l < 16C_\varepsilon^2 k^{2+2\varepsilon}. \end{aligned}\tag{2.13}$$

The inequality for the first summation is by (2.4), while the second one is by (2.12). □

This proves Proposition 1.

Claim. (1.8) *implies* (1.3).

Proof. Applying (1.8) to (2.6) with $\frac{\varepsilon}{2}$, we have

$$k^2|M_l| > c_\varepsilon |G_l|^{2-\frac{\varepsilon}{2}}. \tag{2.14}$$

Putting together (2.14) with (2.7) and (2.8), we have

$$\left(\sum_{m\in M_l} |F_m|\right)^2 < C_\varepsilon k^{2+\varepsilon}|M_l|, \tag{2.15}$$

where $C_\varepsilon = (c_\varepsilon)^{-1}$.

Thus (combining with (2.3))

$$4^l|M_l| < C_\varepsilon k^{2+\varepsilon} \quad \text{for each } 0\leqslant l\leqslant \log_2 k \tag{2.16}$$

and

$$\sum_m |F_m|^2 \leqslant \sum_l |M_l|(2^{l+1})^2 \leqslant C_\varepsilon k^{2+\varepsilon} \sum_l 4 \leqslant 4C_\varepsilon k^{2+\varepsilon}(\log_2 k), \tag{2.17}$$

which is (1.3) (cf. (2.2)). □

Claim. (1.3) *implies* (1.8).

Proof. Let $A = \{a_i \mid i = 1, \ldots, n\}$ and G a graph on A. Define for each $m\in\mathbb{Z}$

$$G_m = \{(j,k)\in G \mid a_j a_k = m\}. \tag{2.18}$$

Since $4a_ja_k=(a_j+a_k)^2-(a_j-a_k)^2$ and a_ja_k is uniquely determined by $u=a_j+a_k$, $v=a_j-a_k$, we may write

$$\begin{aligned}\sum_m |G_m|^2 &= \int_\Pi \left|\sum_{(j,k)\in G} e^{4ia_ja_kx}\right|^2 dx\\ &\leqslant \int \left|\left(\sum_{u\in Sum_G\ A} e^{iu^2x}\right)\left(\sum_{v\in Diff_G\ A} e^{-iv^2x}\right)\right|^2 dx\\ &\leqslant \left(\int \left|\sum_{u\in Sum_G\ A} e^{iu^2x}\right|^4 dx\right)^{\frac{1}{2}}\left(\int \left|\sum_{v\in Diff_G\ A} e^{iv^2x}\right|^4 dx\right)^{\frac{1}{2}}\\ &< C_\varepsilon n^\varepsilon |Sum_G\ A||Diff_G\ A|. \end{aligned} \tag{2.19}$$

Here we used Parseval's identity, Cauchy–Schwartz and (1.3) (Note that $|Sum_G\ A|$, $|Diff_G\ A|$ are both less than n^2).

Next, Cauchy–Schwartz again gives

$$|G| = \sum_{m\in Prod_G\ A} |G_m| \leqslant |Prod_G\ A|^{\frac{1}{2}}\left(\sum |G_m|^2\right)^{\frac{1}{2}}. \tag{2.20}$$

Substitution of (2.19) yields (1.8). □

This proves Proposition 2.

3. Proof of Proposition 3

The proof is based on the following three ingredients:

(i) *T. Gowers's version of the Balog–Szemeredi Theorem* [G]: The Balog–Szemeredi theorem states that if $A\subset\mathbb{Z}$ is a set of k integers such that for some constant $\alpha>0$,

$$|\{(a_1,a_2,a_3,a_4)\mid a_i\in A,\ a_1-a_2+a_3-a_4=0\}|>\alpha|A|^3=\alpha k^3, \tag{3.1}$$

then there is a subset $A'\subset A$ satisfying

$$|A'|>\beta(\alpha)|A| \tag{3.2}$$

and

$$|A'-A'|<C(\alpha)|A| \tag{3.3}$$

for some $0<\beta(\alpha), C(\alpha)$. The original proof (see [Na]) relies on the Szemeredi uniformity lemma, leading to a very poor dependence of $\beta(\alpha), C(\alpha)$ on α when α goes to 0. Recently, Gowers [G] discovered an argument providing powerlike bounds of $\beta(\alpha), C(\alpha)$ in α. We will reproduce this argument here in a bit simplified and economical form.

(ii) *Freiman's Theorem*: This result states that if A is a finite set of integers with small doubling set, i.e. $|A+A|<C|A|$ with C a constant, then A is contained in a proper d-dimensional generalized arithmetic progression P satisfying

$$d\leqslant d(C), \tag{3.4}$$

$$|P|\leqslant C_1(C)|A|. \tag{3.5}$$

See [Na] for definitions, details and a proof. We will use in this context the estimates obtained in [Ch1], where (3.4), (3.5) are shown to hold with

$$d(C)\leqslant [C-1], \tag{3.6}$$

$$\log\frac{|P|}{|A|}\leqslant C^2(\log C)^3. \tag{3.7}$$

(iii) *Density of the set S of squares in arithmetic progressions*: Observe that Proposition 3 implies in particular that if P is a finite arithmetic progression, then $|P\cap S|<o(|P|)$. Already this statement (conjectured by Erdös and proved by Szemeredi [S2]), is far from obvious. It follows for instance from the fact that there are no 4 squares in arithmetic progression (a result due to Fermat) and Szemeredi's theorem on arithmetic progressions in sets of positive density.

Rudin made the natural conjecture that always $|P\cap S|\leqslant|P|^{\frac{1}{2}}$.
In [B-G-P], the following bound is obtained:

$$|P\cap S|\leqslant|P|^{\frac{2}{3}}(\log|P|)^C \tag{3.8}$$

for any finite (1-dimensional) arithmetic progression P.

Very recently, further improvement was given in Bombieri:

$$|P\cap S|<|P|^{\frac{3}{5}+\varepsilon}. \tag{3.9}$$

However, for our purpose, the specific exponent of $|P|$ in (3.8) and (3.9) is of little importance since the main weakness of our argument lies in the use of Freiman's theorem.

Using these three ingredients, it was observed in [Bo] that one has a nontrivial estimate

$$\int_{\Pi} \left| \sum_{j=1}^{k} e^{in_j^2 x} \right|^4 dx < o(k^3) \tag{3.10}$$

whenever $n_1 < n_2 < \cdots < n_k$. The bound (1.9) is obtained from a more careful quantitative analysis of that argument, using in particular the results from [Ch1].

Returning first to the Balog–Szemerédi theorem, we prove

Proposition 5. *Let $A \subset \mathbb{Z}$ be a set of k integers, and let*

$$T = \{(a_1, a_2, a_3, a_4) \mid a_i \in A, a_1 - a_2 = a_3 - a_4\}. \tag{3.11}$$

If there is a positive constant α such that

$$|T| > \alpha k^3, \tag{3.12}$$

then there is a subset $A' \subset A$ satisfying

$$|A'| > \alpha^{1-\varepsilon} k, \tag{3.13}$$

$$|A' - A'| < 2^{21} \alpha^{-5} \left(1 + \left\lceil \log_2 \frac{1}{\alpha} \right\rceil\right)^5 k. \tag{3.14}$$

Proof. Let $\pi : A \times A \times A \times A \to A \times A$ be the projection on the first two coordinates, and let $T_{a_1 a_2} = \pi^{-1}(a_1, a_2)$ be the fiber of T at (a_1, a_2):

$$T_{a_1 a_2} = \{\bar{a} = (a_1, a_2, a_3, a_4) \mid \bar{a} \in T\}. \tag{3.15}$$

Claim 1. *There exist $l \in \mathbb{Z}$, $0 \leqslant l \leqslant \lceil \log_2 \frac{1}{\alpha} \rceil$, and $R \subset A \times A$ such that*

$$R = \left\{ (a_1, a_2) \middle| \frac{k}{2^l} > |T_{a_1 a_2}| > \frac{k}{2^{l+1}} \right\} \tag{3.16}$$

and

$$|R| > \rho k^2, \tag{3.17}$$

where

$$\rho = \frac{2^{l-1} \alpha}{1 + \lceil \log_2 \frac{1}{\alpha} \rceil}. \tag{3.18}$$

Note. For any $(a_1, a_2) \in R$, there are at least $\frac{k}{2^{l+1}}$ ways to represent $a_1 - a_2$ as the difference of two elements in A.

Proof of Claim 1. Bound (3.1) and the fact that $\sum_{|T_{a_1a_2}|<\frac{\alpha}{2}k} |T_{a_1a_2}| < \frac{\alpha}{2}k^3$ imply

$$\sum_{|T_{a_1a_2}|>\frac{\alpha}{2}k} |T_{a_1a_2}| > \frac{\alpha}{2}k^3. \tag{3.19}$$

We refine the summation in (3.19) as

$$\sum_{l=0}^{\lceil \log_2 \frac{1}{\alpha} \rceil} \sum_{\frac{k}{2^l} > |T_{a_1a_2}| > \frac{k}{2^{l+1}}} |T_{a_1a_2}| > \frac{\alpha}{2}k^3. \tag{3.20}$$

So there is l, $0 \leqslant l \leqslant \lceil \log_2 \frac{1}{\alpha} \rceil$, such that

$$\sum_{\frac{k}{2^l} > |T_{a_1a_2}| > \frac{k}{2^{l+1}}} |T_{a_1a_2}| > \frac{\alpha}{2(1 + \lceil \log_2 \frac{1}{\alpha} \rceil)}k^3. \tag{3.21}$$

Let R be the set of points (a_1, a_2) that occur in (3.21):

$$R = \left\{ (a_1, a_2) \middle| \frac{k}{2^l} > |T_{a_1a_2}| > \frac{k}{2^{l+1}} \right\}. \tag{3.22}$$

Then

$$\frac{k}{2^l}|R| > \frac{\alpha}{2(1 + \lceil \log_2 \frac{1}{\alpha} \rceil)}k^3. \tag{3.23}$$

This is (3.17). □

We view R as a (symmetric) relation on A. Denote

$$R_a = \{a' \in A | (a, a') \in R\}. \tag{3.24}$$

Inequality (3.17) is

$$\sum_{a \in A} |R_a| > \rho k^2. \tag{3.25}$$

Next, we define the set

$$Y = \left\{ (a, b) \in R \mid |R_a \cap R_b| < \frac{\rho^2}{32}k \right\}. \tag{3.26}$$

Claim 2. *There exists $c\in A$ such that*

$$|R_c|>\frac{\rho}{2}k \tag{3.27}$$

and

$$|R_c|>\frac{16}{\rho k}|Y\cap(R_c\times R_c)|. \tag{3.28}$$

Proof. Since for any $Y\subset A\times A$, we have

$$\sum_{c\in A}|Y\cap(R_c\times R_c)|=\sum_{(a,b)\in Y}|R_a\cap R_b|,$$

definition (3.26) of Y gives

$$\sum_{c\in A}|Y\cap(R_c\times R_c)|<\frac{\rho^2}{32}k^3. \tag{3.29}$$

Namely,

$$\frac{16}{\rho k}\sum_{c\in A}|Y\cap(R_c\times R_c)|<\frac{\rho}{2}k^2. \tag{3.30}$$

Combining (3.25) and (3.30), we have

$$\sum_{c\in A}|R_c|>\frac{\rho}{2}k^2+\frac{\rho}{2}k^2>\frac{\rho}{2}k^2+\frac{16}{\rho k}\sum_{c\in A}|Y\cap(R_c\times R_c)|$$

which permits us to fix some $c\in A$ such that

$$|R_c|>\frac{\rho}{2}k+\frac{16}{\rho k}|Y\cap(R_c\times R_c)|.$$

This implies (3.27) and (3.28). □

Let

$$Y_a=\{b\,|\,(a,b)\in Y\}, \tag{3.31}$$

and let

$$A'=\{a\in R_c\,|\,|Y_a\cap R_c|<\tfrac{1}{4}|R_c|\}. \tag{3.32}$$

Claim 3.

$$|A'| > \tfrac{1}{2}|R_c| \tag{3.33}$$

and

$$|R_c \setminus (Y_a \cup Y_{a'})| > \tfrac{1}{2}|R_c| \quad \text{for all } a, a' \in A'. \tag{3.34}$$

Proof. To show (3.33), it is sufficient to give an upper bound on $|R_c \setminus A'|$:

$$|R_c \setminus A'| \leqslant \frac{4}{|R_c|} \sum_{a \in R_c} |Y_a \cap R_c| = \frac{4}{|R_c|} |Y \cap (R_c \times R_c)| < \frac{4}{|R_c|} \frac{\rho k}{16} |R_c| = \frac{\rho}{4} k < \frac{1}{2} |R_c|. \tag{3.35}$$

The last two inequalities are (3.28) and (3.27).

On the other hand, (3.34) follows trivially from (3.32). □

Claim 4. For all $a, a' \in A'$, there are at least

$$\frac{1}{2} |R_c| \left(\frac{\rho^2}{32} k\right)^2 \left(\frac{k}{2^{l+1}}\right)^4 \tag{3.36}$$

ways to represent $a - a'$ as

$$a - a' = (\xi_1 - \xi_2) - (\xi_3 - \xi_4) + (\xi_5 - \xi_6) - (\xi_7 - \xi_8) \tag{3.37}$$

with $\xi_i \in A$ for $1 \leqslant i \leqslant 8$.

Proof. Take any $b \in R_c \setminus (Y_a \cup Y_{a'})$. Since $(a, b), (a', b) \notin Y$, from definition (3.26) of Y, we have

$$\begin{aligned} |R_a \cap R_b| &> \frac{\rho^2}{32} k, \\ |R_{a'} \cap R_b| &> \frac{\rho^2}{32} k. \end{aligned} \tag{3.38}$$

Namely, there are at least $(\frac{\rho^2}{32} k)^2$ many choices of (u_1, u_2) such that

$$(a, u_1) \in R, (b, u_1) \in R, (a', u_2) \in R, (b, u_2) \in R. \tag{3.39}$$

Write

$$a - a' = (a - u_1) - (b - u_1) + (b - u_2) - (a' - u_2). \tag{3.40}$$

where, by (3.39) and definition (3.16) of R (see Note after Claim 1), each of the differences $a - u_1, b - u_1, b - u_2, a' - u_2$ may be written in at least $\frac{k}{2^{l+1}}$ ways as

difference of two elements from A. Putting this together with (3.34), (3.38) and (3.39), we have the lower bound (3.36). □

The bound (3.36) holds for any pair $a, a' \in A'$. Since $(\xi_1, \ldots, \xi_8) \in A \times \cdots \times A$ runs in a set of size k^8, we conclude that

$$\frac{1}{2}|R_c|\left(\frac{\rho^2}{32}k\right)^2\left(\frac{k}{2^{l+1}}\right)^4|A'-A'| \leqslant k^8. \tag{3.41}$$

The upper bound (3.14) on $|A'-A'|$ follows from (3.41), (3.27), and (3.18), while the lower bound (3.13) on $|A'|$ follows from (3.33), (3.27), and (3.18).

This proves Proposition 5. □

We will first produce a weaker version of (1.10) and then outline a more economical strategy that gives Proposition 3. We will use the same ε to denote various functions of ε.

Let $A \subset S$ be a subset of the squares, $|A| = k$. To prove (1.10) and to be able to apply Proposition 5, we assume

$$\int_{\Pi}\left|\sum_{a\in A} e^{iax}\right|^4 dx = |\{(a_1,a_2,a_3,a_4)|a_i \in A, a_1 - a_2 + a_3 - a_4 = 0\}| > \alpha k^3. \tag{3.42}$$

We want to find an upper bound on α.

Let $A' \subset A$ be the set obtained in Proposition 5.

Thus

$$|A'| \geqslant \alpha^{1-\varepsilon}k \tag{3.43}$$

and

$$|A'-A'| \leqslant \alpha^{-6-\varepsilon}|A'|, \tag{3.44}$$

where $\alpha^{-6-\varepsilon} = 2^{21}\alpha^{-5}(1+\lceil \log_2 \frac{1}{\alpha}\rceil)^5\alpha^{-1+\varepsilon}$.

From general sumset estimates ([Na, Th 7.8], with $A = A'$, $B = -A'$, $k = 1$, and $l = 2$), it follows that

$$|A'+A'| \leqslant \alpha^{-18-\varepsilon}|A'|. \tag{3.45}$$

Next, we apply Freiman's theorem (3.6) and (3.7) to (3.45) and obtain a proper generalized d-dimensional arithmetic progression P satisfying

$$A' \subset P, \tag{3.46}$$

$$d \leqslant \alpha^{-18-\varepsilon}, \tag{3.47}$$

$$\log\frac{|P|}{|A'|} \leqslant \alpha^{-36-\varepsilon}. \tag{3.48}$$

Finally, use the density estimate (3.8) for the set S of squares in 1-dimensional arithmetic progressions. Notice that since P obtained above is a proper d-dimensional progression, we may clearly obtain P as a union of $\frac{|P|}{|P_0|}$ disjoint translates of a 1-dimensional progression P_0 with

$$|P_0| \geqslant |P|^{\frac{1}{d}}. \tag{3.49}$$

Thus, applying (3.8), we have

$$\max_{n\in\mathbb{Z}} |(n+P_0)\cap S| < |P_0|^{\frac{2}{3}+\varepsilon}. \tag{3.50}$$

Therefore, by (3.46) (note that $A'\subset A\subset S$), (3.50), and (3.49),

$$|A'| \leqslant |P\cap S| < \frac{|P|}{|P_0|}\,|P_0|^{\frac{2}{3}+\varepsilon} < |P|^{1-\frac{1}{3d}+\varepsilon}, \tag{3.51}$$

hence, from (3.48), (3.51), (3.46), (3.43) and (3.47), we have

$$\alpha^{-36-\varepsilon} > \log\frac{|P|}{|A'|} \geqslant \frac{1}{4d}\log|P| > \frac{1}{4d}\log(\alpha^{1-\varepsilon}k) > \alpha^{18+\varepsilon}\log(\alpha^{1-\varepsilon}k). \tag{3.52}$$

Namely,

$$\alpha < (\log k)^{-\frac{1}{54}+\varepsilon}. \tag{3.53}$$

Consequently, there is an upper bound $k^{\frac{3}{4}}(\log k)^{-\frac{1}{216}+\varepsilon}$ in (1.10). □

The procedure just described can be made more efficient.

The following property is from [Ch1]. The case for $|A'+A'|<K|A'|$ is the statement of Proposition 2.1 (which is the main step improving Freiman's theorem following Ruzsa's argument). However, in [Ch1] the assumption $|A'+A'|<K|A'|$ is only used when applying Lemma 3.3, and the proof of Lemma 3.3 uses Fact e, which has either $|A'+A'|<K|A'|$ or $|A'-A'|<K|A'|$ as hypothesis.

Proposition 6 (Chang [Ch1]). *Let $A'\subset\mathbb{Z}$ be a finite set satisfying $|A'+A'|<K|A'|$ or $|A'-A'|<K|A'|$. Then $A'-A'+A'-A'$ contains a proper d-dimensional progression P satisfying*

$$d < C(\log K)K, \tag{3.54}$$

$$\log\frac{|A'|}{|P|} < C(\log K)^2K. \tag{3.55}$$

In view of (3.44), *there is therefore a proper d-dimensional progression P in the set $A'\subset A$ obtained from Proposition* 5, *such that*

$$d < \alpha^{-6-\varepsilon}, \tag{3.56}$$

and

$$\log\frac{|A'|}{|P|}<\alpha^{-6-\varepsilon}. \tag{3.57}$$

Let $A''\subset A'$ be a maximal subset of A' such that all translates $\{a+P|a\in A''\}$ are disjoint. Then clearly

$$\bigcup_{a\in A''}(a+P)\subset 3A'-2A'. \tag{3.58}$$

Hence, again from sumset estimates ([Na, Thm 7.8], with $A=-A', B=A', k=3$, and $l=2$) and (3.44)

$$|A''|\leqslant\frac{|3A'-2A'|}{|P|}<\alpha^{-30-\varepsilon}\frac{|A'|}{|P|}. \tag{3.59}$$

On the other hand, our choice of A'' gives

$$A'\subset\bigcup_{a\in A''}(a+P-P). \tag{3.60}$$

Thus, using again the density estimate (3.51) (note that $A'\subset A\subset S$), we obtain

$$\begin{aligned}|A'|&\leqslant\sum_{a\in A''}|(a+P-P)\cap S|\\&\leqslant|A''|\,2^d\max_{n\in\mathbb{Z}}|(n+P)\cap S|\\&<\alpha^{-30-\varepsilon}\frac{|A'|}{|P|}\,2^d|P|^{1-\frac{1}{3d}+\varepsilon}\\&=\alpha^{-30-\varepsilon}|A'|\,2^d|P|^{-\frac{1}{3d}+\varepsilon}.\end{aligned} \tag{3.61}$$

Particularly, (3.61) implies, for some $C=C(\varepsilon)>0$,

$$\frac{1}{3d}\log|P|<d+C\log\frac{1}{\alpha}. \tag{3.62}$$

Together with (3.57) and (3.56), we have

$$\log|A'|-\alpha^{-6-\varepsilon}<\log|P|\leqslant 3d^2+3dC\log\frac{1}{\alpha}<C'\alpha^{-12-\varepsilon}. \tag{3.63}$$

Putting together (3.63) and (3.43), we have

$$\alpha<(\log k)^{-\frac{1}{12}+\varepsilon}. \tag{3.64}$$

This proves inequality (1.10) (Proposition 3).

Finally, Corollary 4 is deduced from Proposition 3.

Proof of Corollary 4. In (2.19), we have

$$\sum_m |G_m|^2 \leqslant \left(\int \left| \sum_{u \in Sum_G\, A} e^{iu^2 x} \right|^4 dx \right)^{\frac{1}{2}} \left(\int \left| \sum_{v \in Diff_G\, A} e^{iv^2 x} \right|^4 dx \right)^{\frac{1}{2}}. \tag{3.65}$$

Applying Proposition 3, we get

$$\sum_m |G_m|^2 \leqslant c_\varepsilon \frac{|Sum_G\, A|^{\frac{3}{2}}}{(\log |Sum_G\, A|)^{\frac{1}{24}-\varepsilon}} \frac{|Diff_G\, A|^{\frac{3}{2}}}{(\log |Diff_G\, A|)^{\frac{1}{24}} - \varepsilon}. \tag{3.66}$$

On the other hand, (2.20) gives

$$|G| \leqslant |Prod_G\, A|^{\frac{1}{2}} \Big(\sum |G_m|^2 \Big)^{\frac{1}{2}}. \tag{3.67}$$

Putting together (3.66) and (3.67), we have

$$|G| \leqslant C_\varepsilon |Prod_G\, A|^{\frac{1}{2}} \frac{|Sum_G\, A|^{\frac{3}{4}}}{(\log |Sum_G\, A|)^{\frac{1}{48}-\varepsilon}} \frac{|Diff_G\, A|^{\frac{3}{4}}}{(\log |Diff_G\, A|)^{\frac{1}{48}-\varepsilon}}. \tag{3.68}$$

There are two cases:

Case 1. $\min\{\log |Sum_G\, A|, \log |Diff_G\, A|\} > \frac{1}{10} \log |G|$.

Inequality (3.68) implies

$$|G| \leqslant C'_\varepsilon \frac{(\max\{|Sum_G\, A|, |Diff_G\, A|, |Prod_G\, A|\})^2}{(\log |G|)^{\frac{1}{24}-\varepsilon}}. \tag{3.69}$$

This is (1.11).

Case 2. $\min\{\log |Sum_G\, A|, \log |Diff_G\, A|\} \leqslant \frac{1}{10} \log |G|$.

We may assume $\log |Sum_G\, A| \leqslant \frac{1}{10} \log |G|$. (The other case is similar.) Then

$$|Sum_G\, A| \leqslant |G|^{\frac{1}{10}}. \tag{3.70}$$

(3.70) and (3.68) give

$$|G| < \max\{|Diff_G\, A|, |Prod_G\, A|\}^{\frac{5}{4}} |G|^{\frac{3}{40}}.$$

Namely,

$$|G|^{\frac{37}{40}} < \max\{|Sum_G\, A|, |Diff_G\, A|, |Prod_G\, A|\}^{\frac{5}{4}},$$

which implies (1.11). □

Acknowledgments

The author thanks J. Bourgain for explaining [Bo], and E. Gulu for helpful discussions. Partially supported by NSA grant number MDA904–03–1–0045.

References

[B-G-P] E. Bombieri, A. Granville, J. Pintz, Squares in arithmetic progressions, Duke Math. J. 66 (3) (1992) 369–385.

[Bo] J. Bourgain, Lambda-*P* sets in analysis: results, problems and related aspects, in: Handbook of the Geometry of Banach Spaces, Vol. I, Elsevier, Amsterdam.

[Ch1] M. Chang, Polynomial bounds in Freiman's theorem, Duke Math. J. 13 (2002) 399–419.

[Ch2] M. Chang, Erdös–Szemeredi problems on sum set and product set, Ann. Math., to appear.

[Ch3] M. Chang, Factorization in generalized arithmetical progressions and applications to the Erdös–Szemeredi sum–product problems, preprint 2002 (to appear in GAFA).

[E1] G. Elekes, On the number of sums and products, Acta Arith. 81 (4) (1997) 365–367.

[E-S] P. Erdös, E. Szemeredi, On sums and products of integers, in: P. Erdös, L. Alpàr, G. Halàz (Eds.), Studies in Pure Mathematics, to the Memory of P. Turàn, Birkhauser, Basel, p. 213–218.

[G] T. Gowers, A new proof of Szemeredi's theorem for arithmetic progressions of length four, GAFFA 8 (3) (1998) 529–551.

[Na] M. Nathanson, Additive Number Theory, Springer, Berlin, 1996.

[Ru] W. Rudin, Trigonometric series with gaps, J. Math. Mech. 9 (1960) 203–227.

[S2] E. Szemeredi, The number of squares in an arithmetic progression, Studia Sci. Math. Hungar. 9 (3–4) (1974); 417 (1975).