Astérisque

YURI BILU Structure of sets with small sumset

Astérisque, tome 258 (1999), p. 77-108 <http://www.numdam.org/item?id=AST_1999_258_77_0>

© Société mathématique de France, 1999, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/ Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

\mathcal{N} umdam

Article numérisé dans le cadre du programme Numérisation de documents anciens mathématiques http://www.numdam.org/ Astérisque 258, 1999, p. 77-108

STRUCTURE OF SETS WITH SMALL SUMSET

by

Yuri Bilu

Abstract. — Freiman proved that a finite set of integers K satisfying $|K+K| \leq \sigma |K|$ is a subset of a "small" *m*-dimensional arithmetical progression, where $m \leq \lfloor \sigma - 1 \rfloor$. We give a complete self-contained exposition of this result, together with some refinements, and explicitly compute the constants involved.

1. Introduction

This is an exposition of the fundamental theorem due to G. A. Freiman on the addition of finite sets. (It will be referred to as *Main theorem*). Let K be a finite set of integers (more generally, a finite subset of a torsion-free abelian group) of cardinality k. The Main Theorem states that if the sumset K + K is "small", then K possesses a rigid structure. An example of a statement of this type is the following

Proposition 1.1

- (i) Any K satisfies $|K + K| \ge 2k 1$ and the equality |K + K| = 2k 1 implies that K is an arithmetical progression.
- (ii) Assume that |K + K| = 2k 1 + t, where $0 \le t \le k 3$. Then K is a subset of an arithmetical progression of length k + t.
- (iii) Assume that |K + K| = 3k 3 and $k \ge 7$. Then either K is a subset of an arithmetical progression of length 2k 1, or K is a union of two arithmetical progressions with the same difference.

Here (i) is trivial, for (ii) and (iii) see [12, Theorems 1.9 and 1.11], where the result is obtained for subsets of integers. The case of subsets of an arbitrary torsion-free abelian group follows from [12, Lemma 1.14], which is Lemma 4.3 of the present paper.

Let us deviate for a while from our main subject, and make a short (and very incomplete) historical account. Item (i) easily generalizes to distinct summands: if K

¹⁹⁹¹ Mathematics Subject Classification. - 11B25, 11B05.

Key words and phrases. — Addition of finite sets; generalized arithmetical progressions; inverse additive theorems.

Y. BILU

and L are finite subsets of a torsion-free abelian group, then $|K+L| \ge |K|+|L|-1$, and the equality |K+L| = |K|+|L|-1 implies that K and L are arithmetical progressions with the same difference. Freiman [10] extended item (ii) to two distinct summands; see also [15, 23, 32, 35]. An important generalization to several (equal or distinct) summands was obtained by Lev [22]. Concerning item (iii) see also Hamidoune [17].

Item (i) extends to torsion-free non-abelian groups (Brailovski and Freiman [4]). It also has an analogue for cyclic groups of prime order (Cauchy [6], Davenport [7, 8], Vosper [36]). Hamidoune [16] gave short and conceptual proofs of the theorems of Brailovski-Freiman and Vosper. For general finite (abelian and/or non-abelian) groups see [20, 18, 37, 38]. However, we do not know non-commutative analogues of items (ii) and (iii), and we know only partial analogues of these items for cyclic groups of prime order [11, 12, 2].

The first part of item (i) has various continuous analogues, for instance for connected unimodular locally compact groups [19, 29]. Item (ii) has a partial analogue for real tori [1].

Many of the results mentioned above are proved in the books of Mann [24] and Nathanson [26], where the reader can also find further references.

The Main Theorem, however, develops Proposition 1.1 in a completely different direction. Reformulate item (ii) as follows:

Let $\sigma < 3$ be a positive number. Assume that $|K + K| \leq \sigma k$ and $k > 3/(3 - \sigma)$. Then K is a subset of an arithmetical progression of length $(\sigma - 1)k + 1$.

The Main Theorem extends this to arbitrary σ , without the restriction $\sigma < 3$. To formulate it, we need some definitions. Let A, B be abelian groups, $K \subset A$ and $L \subset B$. The map $\varphi : K \to L$ is Freiman's homomorphism of order s or, in the terminology of [28], F_s -homomorphism, if for any $x_1, \ldots, x_s, y_1, \ldots, y_s \in K$ we have

$$x_1 + \dots + x_s = y_1 + \dots + y_s \Rightarrow \varphi(x_1) + \dots + \varphi(x_s) = \varphi(y_1) + \dots + \varphi(y_s)$$

In the other words, the map

$$\psi: \quad \overbrace{K+\cdots+K}^{s} \rightarrow \overbrace{L+\cdots+L}^{s}, \\ x_1+\cdots+x_s \quad \mapsto \quad \varphi(x_1)+\cdots+\varphi(x_s)$$

is well-defined. The F_s -homomorphism φ is an F_s -isomorphism if it is invertible and the inverse φ^{-1} is also an F_s -homomorphism; in other words, when both the maps φ and ψ are invertible. (In particular, F_1 -isomorphism is a synonym to bijection.)

It is easy to find an F_s -isomorphism not induced by a group-theoretic homomorphism $A \to B$. A typical example is the map

$$\begin{array}{rcl} \{0,a,\ldots,(k-1)a\} & \to & \{0,\ldots,k-1\},\\ & xa & \mapsto & x, \end{array}$$

where a generates an additive cyclic group of order p > (k-1)s.

A generalized arithmetical progression (further progression) of rank m in an abelian group A is a set of the form

$$P = P(x_0; x_1, \dots, x_m; b_1, \dots, b_m) = \{x_0 + \beta_1 x_1 + \dots + \beta_m x_m : \beta_i = 0, \dots, b_i - 1\},\$$

where x_0, \ldots, x_m are elements of the group and b_1, \ldots, b_m positive integers. We say that P is an F_s -progression if the map

(1.1)
$$\{0,\ldots,b_1-1\}\times\cdots\times\{0,\ldots,b_m-1\} \rightarrow P, \\ (\beta_1,\ldots,\beta_m) \mapsto x_0+\beta_1x_1+\cdots+\beta_mx_m \}$$

is an F_s -isomorphism. In particular, each F_s -progression is also an $F_{s'}$ -progression for any $s' \leq s$, and P is an F_1 -progression if and only if $|P| = b_1 \cdots b_m$.

Now we are ready to formulate the Main Theorem $^{(1)}$.

Theorem 1.2 (the Main Theorem). — Let σ be a positive real number, s a positive integer, and K a subset of a torsion-free abelian group such that

$$k := |K| > k_0(\sigma) := \frac{\lfloor \sigma \rfloor \lfloor \sigma + 1 \rfloor}{2(\lfloor \sigma + 1 \rfloor - \sigma)}$$

and

$$|K+K| \le \sigma k.$$

Then K is a subset of an F_s -progression P of rank $m \leq \lfloor \sigma - 1 \rfloor$ and cardinality

$$(1.2) |P| \le c_{11}(\sigma, s)k.$$

It must be pointed out that, unlike Proposition 1.1, this theorem has only very few known analogues for other types of groups, all of them being more or less direct consequences of the Main Theorem; see Chapter 3 of Freiman's book [12].

We also suggest the following more precise version of the Main Theorem, asserting that at most $\lfloor \log_2 \sigma \rfloor$ dimensions of the progression P can be "large"; the others are bounded by a constant, depending on σ .

Theorem 1.3. — Assuming the hypothesis of Theorem 1.2, write the F_s -progression P as $P(x_0; x_1, \ldots, x_m; b_1, \ldots, b_m)$, where $b_1 \geq \cdots \geq b_m$. Then

(1.3)
$$b_i \leq c_{12}(\sigma, s) \quad (i > \lfloor \log_2 \sigma \rfloor).$$

(See Subsection 5.5, where Theorem 1.3 is derived from Theorem 1.2.)

The quantitative estimates for the constants involve the function $fr(n,\varepsilon)$, defined in Subsection 5.3. We obtain the estimates

$$_{11}(\sigma,s) \leq (2c_{13}(\sigma)s)^{\sigma^{30\sigma}c_{13}(\sigma)}, \quad c_{12}(\sigma,s) \leq 2c_{11}(\sigma,s')fr\left(\lfloor \log_2 \sigma \rfloor + 1, \varepsilon_0\right),$$

where

c

$$c_{13}(\sigma) = fr\left(\left\lceil 8\sigma \log(2\sigma) \right\rceil, 1\right), \quad \varepsilon_0 = \left\lfloor \log_2 \sigma \right\rfloor + 1 - \log_2 \sigma, \quad s' = \min(s, 2).$$

At present, only a very poor estimate is known (see Subsection 5.3):

$$fr(n,\varepsilon) \leq \left(2 + \varepsilon^{-1}\right)^{\exp \exp n}$$

Therefore we have only

(1.4)
$$c_{11} \leq (2s)^{\exp \exp \exp(9\sigma \log(2\sigma))}$$

⁽¹⁾With a few exceptions, we write explicit constants as c_{ij} , where *i* is the number of the section where the constant is defined, and *j* is the number of the constant in Section *i*.

Freiman published two expositions [12, 13] of his proof. Recently a new proof of Freiman's theorem, simpler and more transparent than the original, was found by Ruzsa [30]. Ruzsa's argument implies the estimate $c_{11} \leq (2s)^{\exp(\sigma^c)}$, which is better than (1.4) (here c is an absolute constant). In the final section we briefly review the main points of Ruzsa's proof. A detailed self-contained exposition of Ruzsa's proof is given in [26, Chapter 8]

Our exposition is based on the same principles as Freiman's original proof [12, 13], though the technical details are different. The most substantial innovations are in Subsection 5.1, where we suggest a simpler proof of the Cube Lemma, and in Subsection 8.3, where we apply the Bombieri–Vaaler theorem instead of Freiman's sophisticated elementary argument. We believe that the original argument of Freiman is still of great interest, even after Ruzsa's work.

We tried to make the exposition self-contained. Only three standard results from the Geometry of Numbers, namely, the theorems of Minkowski, Mahler and Bombieri-Vaaler, are quoted without proofs (but with exact references). The other auxiliary facts are provided with complete proofs even if they are available in the literature.

In Section 2 we introduce the notation used throughout the paper. In Sections 3 and 4 we reduce the Main Theorem to certain more technical statements. At the end of Section 4 we give a plan of the remaining part of the article.

Acknowledgments. Gregory Freiman drew my attention to his theorem. Daniel Berend and Henrietta Dickinson read the drafts of the paper at different stages of its preparation and made a number of valuable remarks. Peter Pleasants sent me his unpublished notes on Freiman's theorem and Mel Nathanson put at my disposal a preliminary version of his book [26]. It is a pleasure to thank all of them.

My special gratitude is to Imre Ruzsa, who carefully studied the (pre)final version of this paper. I found his numerous comments and suggestions very useful. Many thanks for the hard job he has done.

The main part of this job was done in Bordeaux and was supported by the *Bourse Chateaubriand du gouvernement français.* I am grateful to Prof. J.-M. Deshouillers, Mrs D. Cooke and Mrs F. Duquesnoy for having done their best to make my work in Bordeaux pleasant and successful.

I must also acknowledge support of IMPA (Rio de Janeiro), Forschungsinstitut für Mathematik (ETH Zürich) and Lise Meitner Fellowship (Austria), during the final stage of my work on this paper.

2. Notation and conventions

For $B, C \subseteq \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ put

$$B \pm C = \{b \pm c : b \in B, c \in C\}, \quad \alpha B = \{\alpha b : b \in B\},\$$

etc.

A plane $\mathcal{L} \subseteq \mathbb{R}^n$ is a set of the form $v + \mathcal{L}'$, where $v \in \mathbb{R}^n$ and \mathcal{L}' is a linear subspace of \mathbb{R}^n . By (x, y) we denote the standard inner product in \mathbb{R}^n . The Lebesgue measure in \mathbb{R}^n is referred to as *volume* and is denoted by Vol or Vol_n. The standard