Combinatorics, Probability and Computing (2013) 22, 351–365. © Cambridge University Press 2013
doi:10.1017/S0963548313000060

Arithmetic Progressions in Sumsets and L^p-Almost-Periodicity

ERNIE CROOT[1†], IZABELLA ŁABA[2‡] and OLOF SISASK[3§]

[1]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA
(e-mail: ecroot@math.gatech.edu)
[2]Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada
(e-mail: ilaba@math.ubc.ca)
[3]School of Mathematical Sciences, Queen Mary, University of London,
Mile End Road, London E1 4NS, UK
(e-mail: O.Sisask@qmul.ac.uk)

Received 18 October 2012; revised 31 January 2013; first published online 19 March 2013

We prove results about the L^p-almost-periodicity of convolutions. One of these follows from a simple but rather general lemma about approximating a sum of functions in L^p, and gives a very short proof of a theorem of Green that if A and B are subsets of $\{1,\dots,N\}$ of sizes αN and βN then $A+B$ contains an arithmetic progression of length at least

$$\exp(c(\alpha\beta\log N)^{1/2} - \log\log N).$$

Another almost-periodicity result improves this bound for densities decreasing with N: we show that under the above hypotheses the sumset $A+B$ contains an arithmetic progression of length at least

$$\exp\left(c\left(\frac{\alpha\log N}{\log^3 2\beta^{-1}}\right)^{1/2} - \log(\beta^{-1}\log N)\right).$$

AMS 2010 *Mathematics subject classification*: Primary 11B30
Secondary 11B25

1. Introduction

Let A and B be subsets of an abelian group. What can one say about the structure of the sumset $A+B$? One of the main endeavours of additive combinatorics is to answer questions along these lines, and one particularly appealing goal is to find long arithmetic progressions in $A+B$ for particular kinds of sets A and B. Under the assumption that A and B are 'large', a remarkable result in this direction was established by Bourgain [1], who showed that if A and B are subsets of

† Supported by NSF grant DMS-1001111.
‡ Supported by an NSERC Discovery Grant.
§ Partly supported by an NSERC Discovery Grant and an EPSRC Postdoctoral Fellowship.

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

$\{1,\dots,N\}$ of sizes αN and βN, then $A+B$ contains an arithmetic progression of length about[1] $\exp(c(\alpha\beta\log N)^{1/3})$. This was subsequently improved by Green [4] to the following.

Theorem 1.1 (Green). *Suppose A and B are subsets of $\{1,\dots,N\}$ with densities α and β. Then $A+B$ contains an arithmetic progression of length at least*

$$\exp(c(\alpha\beta\log N)^{1/2}-\log\log N).$$

Green's method of proof was quite different to Bourgain's, which involved establishing a particular kind of almost-periodicity result, and a third strategy was employed by Sanders [12] in giving an alternative proof of the theorem. In this paper we prove almost-periodicity results along similar lines to Bourgain's and those in [3], and give applications in the direction of Theorem 1.1. Let us first state a particularly clean almost-periodicity result; this yields precisely Green's bounds when applied in the direction of finding arithmetic progressions in sumsets.

Theorem 1.2. *Let $p\geqslant 2$ and $\epsilon\in(0,1)$ be parameters. Let G be a finite abelian group and suppose $f:G\to\mathbb{C}$ is a function. Then there is a Bohr set $T\subseteq G$ of rank at most Cp/ϵ^2 and radius $c\epsilon$ such that, for each $t\in T$,*

$$\|f(x+t)-f(x)\|_{L^p(x)}\leqslant\epsilon\|\widehat{f}\|_{\ell^1}.$$

(We give definitions in the next section.) Taking $f=1_A*1_B$, the proof of Theorem 1.1 from this result is very quick. What is perhaps surprising is that the proof of this theorem is also very short and involves hardly any Fourier analysis. Instead it relies on a simple and quite general probabilistic argument similar to that employed in [3]; this is contained in Section 3.

Theorem 1.2 is quite general and makes no assumptions on f; however, in order for it to be effective we must assume that $\|\widehat{f}\|_{\ell^1}$ is sufficiently small. This will indeed be the case in the application to proving Theorem 1.1, where $f=1_A*1_B$ and $\|\widehat{f}\|_{\ell^1}$ is well controlled by the Cauchy–Schwarz inequality.

By a different and somewhat more involved argument we are also able to establish the following strengthening of Theorem 1.1.

Theorem 1.3. *Suppose A and B are subsets[2] of $\{1,\dots,N\}$ with densities α and β. Then $A+B$ contains an arithmetic progression of length at least*

$$\exp\left(c\left(\frac{\alpha\log N}{(\log 2\beta^{-1})^3}\right)^{1/2}-\log(\beta^{-1}\log N)\right).$$

The improvement of this over Theorem 1.1 arises when one of the sets has density decreasing with N. For example, note that for Theorem 1.1 to yield a non-trivial conclusion at least one

[1] Here, and throughout the paper, we employ the convenient device of letting the letters c and C denote positive absolute constants that may change from occurrence to occurrence.

[2] We shall always assume that A and B are non-empty, and that $N\geqslant 3$, to avoid trivialities.

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

of the sets needs to have density at least $\log\log N/(\log N)^{1/2}$, and each set must have density at least $(\log\log N)^2/\log N$, whereas Theorem 1.3 allows both sets to have density about $(\log\log N)^C/\log N$, and in fact one of the sets can have density as low as $\exp(-(\log N)^c)$.

We shall in fact deduce this theorem from the following small-doubling version, where the sets involved are not assumed to be dense in an interval but rather satisfy a small sumset hypothesis.

Theorem 1.4. *Suppose A and B are sets of integers satisfying $|A+B| \leqslant K_A|A|, K_B|B|$. Then $A+B$ contains an arithmetic progression of length at least*

$$\exp\left(c\left(\frac{\log|A+B|}{K_B(\log 2K_A)^3}\right)^{1/2} - \log(2K_A\log 2|A|)\right).$$

Theorem 1.3 follows on noting that if A and B are subsets of $\{1,\dots,N\}$ then $|A+B| \leqslant 2N$, whence we may take $K_A = 2\alpha^{-1}$ and $K_B = 2\beta^{-1}$.

Behind this theorem lies another almost-periodicity result; this time the driving force is a combination of the main result from [3] with an observation from the remarkable paper [14] by Sanders. The result, Theorem 6.2, is more naturally stated in the context of other results and so we postpone it to Section 6.

The remainder of this paper is laid out as follows. In the next section we fix notation and note some preliminary lemmas. In Section 3 we prove a general lemma on approximating functions in L^p; this forms the heart of our proof of Theorem 1.2, which we deduce together with Theorem 1.1 in Section 4. Together these sections provide all the material needed for the reader interested purely in a short proof of Theorem 1.1. The proof of Theorem 1.4 proceeds by different means, starting in Section 5 where we note a variant of Ruzsa's model lemma that works efficiently with two sets. In Section 6 we state and prove Theorem 6.2, the almost-periodicity result required for Theorem 1.4, and deduce the latter theorem. Finally, in Section 7 we make some further remarks, including details about the finite field versions of our results.

2. Notation, definitions and preliminaries

We now fix notation and give definitions. Most of what we use is standard in the additive combinatorial literature; as a consequence we shall be rather brief, referring the unfamiliar reader to the book [15] by Tao and Vu for more information about any of the following notions.

For a finite abelian group G we write $\widehat{G} = \{\gamma : G \to \mathbb{C}^\times : \gamma \text{ a homomorphism}\}$ for its dual group, and we define the Fourier transform $\widehat{f} : \widehat{G} \to \mathbb{C}$ of a function $f : G \to \mathbb{C}$ by $\widehat{f}(\gamma) = \mathbb{E}_{x\in G} f(x)\overline{\gamma(x)}$. Here and throughout the paper, $\mathbb{E}_{x\in X} = |X|^{-1}\sum_{x\in X}$ denotes the average over a finite set X; when X is clear from the context we may write simply $\mathbb{E}_x$. We define the *convolution* of two functions $f, g : G \to \mathbb{C}$ by $f * g(x) = \mathbb{E}_y f(y)g(x-y)$. With these normalizations, the Fourier inversion formula, Parseval's identity and the convolution identity take the form

$$f(x) = \sum_{\gamma\in\widehat{G}} \widehat{f}(\gamma)\gamma(x),$$

$$\mathbb{E}_{x\in G}|f(x)|^2 = \sum_{\gamma\in\widehat{G}} |\widehat{f}(\gamma)|^2, \quad \text{and}$$

$$\widehat{f * g}(\gamma) = \widehat{f}(\gamma)\widehat{g}(\gamma).$$

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

We plainly need the notion of a Bohr set; our definition here is not identical to that in [15] but is essentially equivalent.

Definition. Let G be a finite abelian group. For a set $\Gamma \subseteq \widehat{G}$ of characters and a number $\delta \geqslant 0$ we define

$$\mathrm{Bohr}_G(\Gamma, \delta) = \{x \in G : |\gamma(x) - 1| \leqslant \delta \text{ for all } \gamma \in \Gamma\},$$

and call this[3] a *Bohr set* of *rank* $|\Gamma|$ and *radius* δ.

A Bohr set is thus a set of approximate annihilators of a set of characters, and today it is well known that such sets play an important role in additive combinatorics. It is of critical importance that Bohr sets are highly structured.

Lemma 2.1. *Suppose T is a Bohr set of rank $d \geqslant 1$ and radius $\delta \in (0, 2]$ in a finite abelian group G. Then $|T| \geqslant (\delta/2\pi)^d |G|$. Furthermore, if $G = \mathbb{Z}_N$ is a group of residues modulo a prime N, then T contains an arithmetic progression of size at least $\frac{1}{2\pi}\delta N^{1/d}$.*

See for example Section 4.4 of [15] for proofs.

For a finite abelian group G we write μ_G for the normalized counting measure $\mu_G(A) = |A|/|G|$ on G, and for a subset X of G we write $\mu_X = 1_X/\mu_G(X)$. For a function $f : X \to \mathbb{C}$ we use the notation $\|f\|_{\ell^p} = (\sum_{x\in X}|f(x)|^p)^{1/p}$ for the ℓ^p norm of f and $\|f\|_{L^p} = \|f\|_p = (\mathbb{E}_x|f(x)|^p)^{1/p}$ for the L^p norm of f.

Finally, we write $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ for the unit circle in $\mathbb{C}$.

3. Approximating a linear combination of functions in L^p

In this section we present a rather general lemma about approximating a linear combination of functions by a combination of just a few of those functions. Theorem 1.2, the simple almost-periodicity result, will be a simple deduction from this lemma. To illustrate the lemma, let us sketch how it works as applied in the proof of Theorem 1.2. To prove Theorem 1.2 it suffices to approximate f in L^p by a linear combination of a few characters. The Fourier inversion formula gives us an expression

$$f = \sum_{\gamma\in\widehat{G}} \widehat{f}(\gamma)\gamma \tag{3.1}$$

for f as a linear combination of characters, but too many to be of use. The idea, then, is to reduce the number of characters used by sampling some at random. Which distribution should we use? An answer is suggested directly by (3.1): we can view this expression essentially as the expectation of a random character χ, where χ is chosen to be the character γ with probability proportional to $\widehat{f}(\gamma)$ (ignoring technicalities). Taking an average of a few such random characters then approximates f well in L^p by the law of large numbers, yielding what we want.

[3] There is an abuse of notation here; technically one should specify the triple (G, Γ, δ) and not just the set $\mathrm{Bohr}_G(\Gamma, \delta)$ in order to discuss the rank and radius of a Bohr set.

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

The version of the law of large numbers we shall use is the following inequality of Marcinkiewicz and Zygmund.

Lemma 3.1 (Marcinkiewicz–Zygmund inequality). *Suppose $X_1, \dots, X_n$ are independent, mean zero (complex-valued) random variables with $\mathbb{E}|X_i|^p < \infty$. Then*

$$\mathbb{E}\left|\sum_{j=1}^{k} X_j\right|^p \leqslant (Cp)^{p/2}\mathbb{E}\left(\left(\sum_{j=1}^{k}|X_j|^2\right)^{p/2}\right).$$

This is proved in many places, though attention is often not paid to the dependence on p in the multiplicative constant. One route to get this particular dependence on p is to note that one has this form for the bound for Rademacher random variables via Khintchine's inequality [8], and that one can deduce the result for general real-valued random variables from this by the techniques of symmetrization and randomization; see for example [7, § 3.8]. The complex-valued case follows easily from the real case.

With this in hand, the following lemma implements the idea outlined above in a slightly more general set-up. For a non-zero complex number λ we write $\lambda^\circ = \lambda/|\lambda| \in \mathbb{T}$ for the direction on the unit circle determined by λ, and we set $0^\circ = 0$.

Lemma 3.2. *Let $p \geqslant 2$ and $\epsilon \in (0,1)$ be parameters, and let (X, μ) be a probability space. Suppose*

$$f = \lambda_1 g_1 + \cdots + \lambda_N g_N,$$

where each $\lambda_j \in \mathbb{C}$ and each $g_j : X \to \mathbb{C}$ is a measurable function with $\|g_j\|_{L^p(\mu)} \leqslant 1$. Then there is a positive integer $k \leqslant Cp/\epsilon^2$ and a k-tuple $\sigma \in [N]^k$ such that

$$\|\tfrac{1}{\|\lambda\|_{\ell^1}} f - \tfrac{1}{k}(\lambda^\circ_{\sigma_1} g_{\sigma_1} + \cdots + \lambda^\circ_{\sigma_k} g_{\sigma_k})\|_{L^p(\mu)} \leqslant \epsilon.$$

In fact, the inequality holds with probability at least 0.99 if the tuple $(\sigma_1, \dots, \sigma_k)$ is chosen with probability $|\lambda_{\sigma_1} \cdots \lambda_{\sigma_k}|/\|\lambda\|_{\ell^1}^k$.

Remark. *In the applications we have in mind one generally has X being a finite set and μ being uniform measure on some subset of X.*

Proof. Let h be chosen randomly from $\lambda_1^\circ g_1, \dots, \lambda_N^\circ g_N$, with $\lambda_j^\circ g_j$ being chosen with probability $|\lambda_j|/\|\lambda\|_{\ell^1}$. Then we have

$$\mathbb{E}h = \sum_{j=1}^{N} \frac{\lambda_j}{\|\lambda\|_{\ell^1}} g_j = f/\|\lambda\|_{\ell^1} =: f_0.$$

Let $h_1, \dots, h_k$ be i.i.d. copies of h. The average $k^{-1}(h_1 + \cdots + h_k)$ ought then to approximate its expectation f_0 provided k is not too small; and indeed, by the Marcinkiewicz–Zygmund

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

inequality coupled with two interchanges of orders of integration, we have

$$\mathbb{E}\int_X \left|\frac{1}{k}\sum_{j=1}^k h_j(x) - f_0(x)\right|^p \,\mathrm{d}\mu(x) \leqslant \frac{(Cp)^{p/2}}{k^{p/2}}\mathbb{E}\int_X \left(\frac{1}{k}\sum_{j=1}^k |h_j(x) - f_0(x)|^2\right)^{p/2} \mathrm{d}\mu(x)$$
$$\leqslant \frac{(Cp)^{p/2}}{k^{p/2}}, \tag{3.2}$$

where the second inequality follows from the nesting of L^q norms and the assumption that $\|g_j(x)\|_{L^p(\mu)} \leqslant 1$ for each j. Choosing $k = \lceil Cp/\epsilon^2 \rceil$ gives that this is at most $c\epsilon^p$ for some small constant c, and so

$$\mathbb{E}\left\|\frac{1}{k}\sum_{j=1}^k h_j(x) - f_0(x)\right\|_{L^p(\mu)}^p \leqslant c\epsilon^p.$$

Hence, by Markov's inequality, we have that

$$\left\|\frac{1}{k}\sum_{j=1}^k h_j(x) - f_0(x)\right\|_{L^p(\mu)} \leqslant \epsilon$$

holds with probability at least $1 - c$, as required. □

Remarks. **(1)** In obtaining (3.2) we bounded the integrals rather crudely, using the bounds $\|h_j(x)\|_{L^p(\mu)}, \|f_0\|_{L^p(\mu)} \leqslant 1$. One can in certain situations make some additional savings here; see the remarks at the end of the next section.

(2) Since the completion of this paper we learned that abstract random sampling ideas similar to the above have been used in other contexts before. In particular, our proof of Lemma 3.2 essentially constitutes a proof of Maurey's lemma in Banach space theory, as recorded in [10]. We thank Assaf Naor for directing us to this paper.

Applied to the Fourier expansion (3.1), the lemma has the following immediate corollary.

Corollary 3.3. *Let G be a finite abelian group and let $\epsilon \in (0,1)$ and $p \geqslant 2$ be parameters. For any non-zero function $f : G \to \mathbb{C}$ there are characters $\gamma_1, \dots, \gamma_k \in \widehat{G}$ with $k \leqslant Cp/\epsilon^2$ and coefficients $c_1, \dots, c_k \in \mathbb{C}$ with $|c_j| = 1$ such that*

$$\|f/\|\widehat{f}\|_{\ell^1} - \tfrac{1}{k}(c_1\gamma_1 + \cdots + c_k\gamma_k)\|_{L^p} \leqslant \epsilon.$$

4. Deducing almost-periodicity and finding progressions in sumsets

We are now ready to prove Theorems 1.1 and 1.2. We restate these here so that this section provides a self-contained account of Green's bounds, modulo Lemma 2.1 and Corollary 3.3.

Theorem 1.2. *Let $p \geqslant 2$ and $\epsilon \in (0,1)$ be parameters. Let G be a finite abelian group and suppose $f : G \to \mathbb{C}$ is a function. Then there is a Bohr set $T \subseteq G$ of rank at most Cp/ϵ^2 and*

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

radius $c\epsilon$ such that, for each $t \in T$,

$$\|f(x+t) - f(x)\|_{L^p(x)} \leqslant \epsilon \|\widehat{f}\|_{\ell^1}.$$

Proof. Applying Corollary 3.3 gives us characters $\gamma_1, \ldots, \gamma_k$ with $k \leqslant Cp/\epsilon^2$ and complex numbers $c_1, \ldots, c_k$ with $|c_j| = 1$, such that $g = (c_1\gamma_1 + \cdots + c_k\gamma_k)/k$ satisfies

$$\|f/\|\widehat{f}\|_{\ell^1} - g\|_{L^p} \leqslant \epsilon/3.$$

But g is L^∞-almost-periodic: for any $t \in \mathrm{Bohr}_G(\{\gamma_1, \ldots, \gamma_k\}, \epsilon/3)$, the triangle inequality yields

$$|g(x+t) - g(x)| \leqslant \frac{1}{k}\sum_{j=1}^{k} |\gamma_j(t) - 1| \leqslant \epsilon/3$$

for any $x \in G$, whence the theorem follows by the L^p triangle inequality. □

We can now deduce Theorem 1.1.

Theorem 1.1. *Suppose A and B are subsets of $\{1, \ldots, N\}$ with densities α and β. Then $A + B$ contains an arithmetic progression of length at least*

$$\exp(c(\alpha\beta \log N)^{1/2} - \log\log N).$$

Proof. By embedding the interval $\{1, \ldots, N\}$ in the cyclic group $\mathbb{Z}_{N'}$, where $4N \leqslant N' \leqslant 8N$ is a prime, it suffices to establish the theorem for cyclic groups of prime order instead of intervals; we shall thus prove the theorem for $A, B \subseteq \mathbb{Z}_N$ with densities α, β.

Let $f = 1_A * 1_B$ and apply Theorem 1.2 to f with parameters to be determined; this gives us a Bohr set T of rank $d \leqslant Cp/\epsilon^2$ and radius $c\epsilon$ such that, for any $t \in T$,

$$\|1_A * 1_B(x+t) - 1_A * 1_B(x)\|_{L^p(x)} < \epsilon \|\widehat{1_A}\widehat{1_B}\|_{\ell^1} \leqslant \epsilon(\alpha\beta)^{1/2},$$

where we have used the convolution identity, the Cauchy–Schwarz inequality and Parseval's identity. Let P be a long arithmetic progression in T; we shall determine how long we can pick it in a moment. We then have

$$\begin{aligned}\mathbb{E}_{x\in G} \sup_{t\in P} |1_A * 1_B(x+t) - 1_A * 1_B(x)| &\leqslant \mathbb{E}_x \left(\sum_{t\in P} |1_A * 1_B(x+t) - 1_A * 1_B(x)|^p\right)^{1/p} \\ &< |P|^{1/p}\epsilon(\alpha\beta)^{1/2}\end{aligned}$$

by the nesting of L^p norms. If this were less than $\alpha\beta = \mathbb{E}_x 1_A * 1_B(x)$ then we would be done, for then there would be an element x such that $1_A * 1_B(x+t) > 0$ for each $t \in P$, whence $x + P \subseteq A + B$. Let us pick $\epsilon = (\alpha\beta)^{1/2}/e$; we shall then be done if $|P| \leqslant e^p$. By Lemma 2.1 we can find an arithmetic progression P in T of length the integer part of

$$c\epsilon N^{1/d} = \exp(cp^{-1}\alpha\beta \log N - \log(C(\alpha\beta)^{-1})),$$

so we pick $p = C\sqrt{\alpha\beta \log N}$ in order to ensure that $|P| \leqslant e^p$. Noting that the theorem is trivial if $\alpha\beta \leqslant C/\log N$, we are done. □

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

Fourier space versus physical space

We should note that one can also obtain physical-side almost-periodicity results via Lemma 3.2 by applying it to the expansion

$$1_A * 1_B = \sum_{y \in G} \frac{1_A(y)}{|G|} 1_{y+B}.$$

So applied, the lemma essentially lets one approximate $1_A * 1_B$ by a multiple of $1_C * 1_B$ for many small sets $C \subseteq A$, and from this one can deduce almost-periodicity results: see [3]. One can in fact obtain slightly stronger estimates for the expectations involved in the proof of Lemma 3.2 in this set-up than those concluded above. In particular, by specializing the proof to this expansion at (3.2), one naturally obtains the following version of the main result of [3]; see also [14, Lemma 3.3].

Theorem 4.1. *Let $p \geqslant 2$, $\epsilon \in (0,1)$ and let $k \in \mathbb{N}$ be parameters. Let G be a group and let A, B and S be finite subsets of G with $|S \cdot A| \leqslant K|A|$. Then there is a set $T \subseteq S$ of size*

$$|T| \geqslant \frac{0.99|S|}{K^{Cpk^2/\epsilon^2}}$$

such that, for each $t \in (T^{-1}T)^k$,

$$\|1_A * 1_B(tx) - 1_A * 1_B(x)\|_{\ell^p(x)} \leqslant \epsilon |A||B|^{1/p}.$$

Here we have used product notation for the group operation and unnormalized convolutions $f * g(x) = \sum_{y \in G} f(y)g(y^{-1}x)$.

By making use of more detailed still distributional information in the random sampling, one can obtain estimates of a different form for the almost-periodicity of $1_A * 1_B$. This was done in [3], and it is the use of a particular form of these estimates that leads to the improved density dependence in Theorem 1.3.

5. A variant of Ruzsa's model lemma

We now turn to the proof of Theorem 1.4, which finds long arithmetic progressions in $A + B$ if A and B are sets of integers with $|A + B|$ small. Before we can embark on the proof proper, we need to transfer the problem to one in a cyclic group of prime order, where we have more effective tools at our disposal. If A and B were assumed to be dense subsets of $\{1, \dots, N\}$, then we could simply embed the sets in $\mathbb{Z}_{N'}$ for a small prime N' and work there instead of in the integers, as in the proof of Theorem 1.1. In the small-doubling set-up, however, we need to argue more subtly; the following modification of a modelling lemma of Ruzsa [11, Theorem 2.3.5] will allow us to proceed efficiently.

Proposition 5.1. *Let A and B be finite sets of integers and suppose $k \geqslant 2$ is an integer. Then, for any integer $N \geqslant |kA - kA + kB - kB|$ there are subsets $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geqslant |A|/2k$, $|B'| \geqslant |B|/2k$ such that $A' + B'$ is k-isomorphic to a subset of $\mathbb{Z}_N$.*

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

In this context, two subsets A and A' of two abelian groups are said to be *Freiman k-isomorphic* if there is a bijection $\phi : A \to A'$ such that

$$a_1 + \cdots + a_k = a'_1 + \cdots + a'_k \iff \phi(a_1) + \cdots + \phi(a_k) = \phi(a'_1) + \cdots + \phi(a'_k)$$

whenever $a_j \in A$; such a map is called a *k-isomorphism*. The relevance of this property here is that a 2-isomorphism preserves the structure of a sumset, as well as the structure of an arithmetic progression.

In the following proof we use the notation $\{x\} = x - \lfloor x \rfloor$ to represent the fractional part of a real number x.

Proof. We follow Ruzsa's proof with minor modifications. Let ξ be a real number in the interval $[0, N]$, to be fixed later, and define the map $\phi : \mathbb{Z} \to \mathbb{Z}_N$ by

$$\phi(a) = \lfloor \xi a \rfloor \pmod N.$$

We shall show that if ξ is chosen appropriately then one can find A' and B' as required such that the map $\psi(a + b) := \phi(a) + \phi(b)$ is well-defined and a k-isomorphism on $A' + B'$.

To this end, let us define

$$A_j = \left\{ a \in A : \frac{j-1}{2k} \leqslant \{\xi a\} < \frac{j}{2k} \right\} \text{ for } j = 1, \ldots, 2k,$$

and similarly for B_j. We claim that if ξ is chosen appropriately then any A_r and B_s will lead to a good ψ above; the sets A' and B' can thus simply be taken to be the largest A_r and B_s, which will clearly have sizes at least $|A|/2k$ and $|B|/2k$.

That ψ is well-defined and a k-isomorphism on each $A_r + B_s$ will follow if we can show that there is a ξ such that, for arbitrary $a_1, \ldots, a_k, a'_1, \ldots, a'_k \in A_r$ and $b_1, \ldots, b_k, b'_1, \ldots, b'_k \in B_s$, the congruence

$$\begin{aligned} &\lfloor \xi a_1 \rfloor + \cdots + \lfloor \xi a_k \rfloor + \lfloor \xi b_1 \rfloor + \cdots + \lfloor \xi b_k \rfloor \\ &\quad \equiv \lfloor \xi a'_1 \rfloor + \cdots + \lfloor \xi a'_k \rfloor + \lfloor \xi b'_1 \rfloor + \cdots + \lfloor \xi b'_k \rfloor \pmod N \end{aligned} \tag{5.1}$$

holds if and only if the equality

$$a_1 + \cdots + a_k + b_1 + \cdots + b_k = a'_1 + \cdots + a'_k + b'_1 + \cdots + b'_k \tag{5.2}$$

holds. To see that such a ξ exists, consider the quantity

$$\begin{aligned} &\sum_{i=1}^{k} (\lfloor \xi a_i \rfloor - \lfloor \xi a'_i \rfloor) + \sum_{i=1}^{k} (\lfloor \xi b_i \rfloor - \lfloor \xi b'_i \rfloor) \\ &\quad = \xi \sum_{i=1}^{k} (a_i - a'_i) + \xi \sum_{i=1}^{k} (b_i - b'_i) - \sum_{i=1}^{k} (\{\xi a_i\} - \{\xi a'_i\}) - \sum_{i=1}^{k} (\{\xi b_i\} - \{\xi b'_i\}). \end{aligned} \tag{5.3}$$

Supposing first that (5.2) holds, we have that if all the a_i, a'_i and b_i, b'_i lie in A_r and B_s respectively, then the last two sums in (5.3) have absolute value strictly less than $1/2$ each, and so the left-hand side is an integer lying strictly between -1 and 1, and is hence 0. Thus (5.1) holds even as an equality of integers, not just a congruence, regardless of the choice of ξ.

Suppose now that (5.1) holds. The left-hand side of (5.3) is then a multiple of N, and the right-hand side is of the form $\xi t + \delta$, where $t \in kA - kA + kB - kB$ and $|\delta| < 1$. In this notation, we

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

want our choice of ξ to force $t = 0$. It thus suffices to choose $\xi \in [0, N]$ such that

$$\xi \notin \frac{1}{t} \cdot (dN + (-1, 1))$$

for any integer d and non-zero integer $t \in kA - kA + kB - kB$. For a fixed t, this condition excludes at most $|t| + 1$ intervals from $[0, N]$ (corresponding to $d \in \{0, \dots, t\}$), the lengths of which sum to 2. The number of integers t that need to be considered is $(|kA - kA + kB - kB| - 1)/2$, since we are omitting 0 and since t and $-t$ give rise to the same collection of excluded intervals. The sum total lengths of the excluded intervals is thus $|kA - kA + kB - kB| - 1$, and since $N \geqslant |kA - kA + kB - kB|$ this means that there is a $\xi \in [0, N]$ outside these intervals, completing the proof. □

Remark. *Clearly one can extend this statement to m sets instead of two, as long as one is willing to reduce the size of the sets by a factor of mk instead of $2k$.*

In order to apply this result we shall need an efficient bound on the size of the set $2A - 2A + 2B - 2B$ when $A + B$ is small. The following lemma provides this.

Lemma 5.2. *Suppose A and B are sets of integers with $|A + B| \leqslant K \min(|A|, |B|)$. Then $|2A - 2A + 2B - 2B| \leqslant K^C |B|$.*

Proof. This is an application of standard tools from additive combinatorics. There are many possible ways to argue; here is one. By averaging, there is an element x such that the set $D = (x - A) \cap B$ has size at least $|A||B|/|A + B| \geqslant |A|/K$. By the Ruzsa triangle inequality [15, Lemma 2.6] we then have

$$|2A - 2A + 2B - 2B| \leqslant \frac{|2A - 2A - D||D + 2B - 2B|}{|D|} \leqslant \frac{|3A - 2A||3B - 2B|}{|D|} \leqslant K^{11}|B|,$$

where at the last step we used the Plünnecke–Ruzsa inequality [15, Corollary 6.29] (see also [9] for Petridis's simple proof of this inequality). □

6. Improving the density dependence and using small doubling

As previously mentioned, in order to obtain the improved density dependence in Theorem 1.3 we shall make use of the more detailed moment estimates given in [3]. These are encoded in the following specialization of Proposition 3.3 of [3].

Theorem 6.1. *Let $p \geqslant 2$ and $k \in \mathbb{N}$ be parameters. Let G be a finite abelian group and let A, B and S be subsets of G with $|A + S| \leqslant K_A|A|$ and $|A + B| \leqslant K_B|B|$. Suppose $0 < \epsilon \leqslant k/\sqrt{K_B}$. Then there is a set $T \subseteq S$ of size*

$$|T| \geqslant \exp(-Ck^2 p \log(2K_A)/\epsilon^2)|S|$$

such that, for each $t \in kT - kT$,

$$\|\mu_A * 1_B(x + t) - \mu_A * 1_B(x)\|_{L^p(x)} \leqslant \epsilon \|\mu_A * 1_B\|_{p/2}^{1/2}.$$

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

(Recall that $\mu_A = 1_A/\mu_G(A)$.) The proof of this theorem in [3] was combinatorial, and various consequences of it were drawn by combinatorial means. It turns out that if we couple its conclusion with some relatively simple Fourier analysis then we can magnify its quantitative effectiveness dramatically: more specifically, following one of the key ideas of using Chang's theorem to handle a highly iterated convolution in the paper [14] of Sanders, we shall 'bootstrap' the iterated sumset of almost-periods given by the previous theorem to a Bohr set of almost-periods.

Theorem 6.2. *Let G be a finite abelian group and let $A, B \subseteq G$ be non-empty subsets with $|A+B| \leqslant K_A|A|, K_B|B|$. Let $p \geqslant 2$ and suppose $0 < \epsilon \leqslant 1/\sqrt{K_B}$. Then there is a Bohr set $T \subseteq G$ of rank $d \leqslant Cp(\log 2K_A/\epsilon)^2(\log 2K_A)/\epsilon^2 + C\log \mu_G(B)^{-1}$ and radius at least $c\epsilon/d\sqrt{K_A}$ such that*

$$\|\mu_A * 1_B(x+t) - \mu_A * 1_B(x)\|_{L^p(x)} \leqslant \epsilon \|\mu_A * 1_B\|_{p/2}^{1/2} \quad \text{for each } t \in T.$$

Proof. Apply the previous theorem with $S = B$ and parameter k to be determined, to get a set X with

$$\tau := |X|/|G| \geqslant \exp(-Ck^2 p(\log 2K_A)/\epsilon^2 - \log \mu_G(B)^{-1})$$

such that

$$\|\mu_A * 1_B(x+t) - \mu_A * 1_B(x)\|_{L^p(x)} \leqslant \tfrac{1}{3}\epsilon \|\mu_A * 1_B\|_{p/2}^{1/2} \text{ for each } t \in kX.$$

By the triangle inequality we thus have

$$\|\mu_A * 1_B * \mu_X^{(k)} - \mu_A * 1_B\|_p \leqslant \tfrac{1}{3}\epsilon \|\mu_A * 1_B\|_{p/2}^{1/2},$$

where $\mu_X^{(k)} = \mu_X * \mu_X * \cdots * \mu_X$ denotes the k-fold convolution of μ_X with itself.

Now let $t \in G$ be arbitrary. Another application of the triangle inequality yields

$$\begin{aligned}\|\mu_A * 1_B(x+t) - \mu_A * 1_B(x)\|_{L^p(x)} &\leqslant \tfrac{2}{3}\epsilon \|\mu_A * 1_B\|_{p/2}^{1/2} \\ &\quad + \|\mu_A * 1_B * \mu_X^{(k)}(x+t)1_S(x) - \mu_A * 1_B * \mu_X^{(k)}(x)1_S(x)\|_{L^p(x)}, \qquad (6.1)\end{aligned}$$

where $S = (A+B) \cup (A+B-t)$ contains the support of $\mu_A * 1_B(x+t) - \mu_A * 1_B(x)$. We shall bound the last term here by restricting t to an appropriate Bohr set. Indeed, bounding the term by its maximum on S and then using Fourier inversion, we have that it is at most

$$\begin{aligned}&2^{1/p}\mu_G(A+B)^{1/p}\|\mu_A * 1_B * \mu_X^{(k)}(x+t) - \mu_A * 1_B * \mu_X^{(k)}(x)\|_{L^\infty(x)} \\ &\leqslant 2^{1/p}\mu_G(A+B)^{1/p} \sum_{\gamma \in \widehat{G}} |\widehat{\mu_A}(\gamma)||\widehat{1_B}(\gamma)||\widehat{\mu_X}(\gamma)|^k |\gamma(t) - 1|.\end{aligned}$$

Now let $\Gamma = \{\gamma \in \widehat{G} : |\widehat{\mu_X}(\gamma)| \geqslant 1/e\}$ and set $\delta = \epsilon/5\sqrt{K_A}$ and $k = \lceil \log 2/\delta \rceil$. Then, for any $t \in \mathrm{Bohr}_G(\Gamma, \delta)$ we have

$$\begin{aligned}\sum_{\gamma \in \widehat{G}} |\widehat{\mu_A}(\gamma)||\widehat{1_B}(\gamma)||\widehat{\mu_X}(\gamma)|^k |\gamma(t) - 1| &\leqslant \delta \sum_{\gamma \in \Gamma} |\widehat{\mu_A}(\gamma)||\widehat{1_B}(\gamma)| + \frac{2}{e^k} \sum_{\gamma \notin \Gamma} |\widehat{\mu_A}(\gamma)||\widehat{1_B}(\gamma)| \\ &\leqslant \delta \mu_G(B)^{1/2} \mu_G(A)^{-1/2}.\end{aligned}$$

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

To be of use in (6.1) we thus want $2^{1/p}\mu_G(A+B)^{1/p}\delta\mu_G(B)^{1/2}\mu_G(A)^{-1/2}$ to be at most $\frac{1}{3}\epsilon\|\mu_A * 1_B\|_{p/2}^{1/2}$. A quick calculation using Hölder's inequality and the relationship $\|\mu_A * 1_B\|_1 = \mu_G(B)$ reveals that our choice of δ ensures this. Hence we have the conclusion we want for any $t \in \mathrm{Bohr}_G(\Gamma,\delta)$. To obtain the low-rank conclusion, we apply Chang's theorem [2] (see also [15, Lemma 4.36]): this says that $\mathrm{Bohr}_G(\Gamma,\delta)$, being a Bohr set associated to a set of large Fourier coefficients, contains a low-rank Bohr set $\mathrm{Bohr}_G(\Lambda,\delta/d)$ where

$$d := |\Lambda| \leqslant C\log 1/\tau \leqslant Cp(\log 2K_A/\epsilon)^2(\log 2K_A)/\epsilon^2 + C\mu_G(B)^{-1},$$

which completes the proof. □

Proof of Theorem 1.4. We first need to embed the sets A and B into a finite abelian group so that we can effectively use the properties of Bohr sets. As mentioned in Section 5, this is simple in the set-up of Theorem 1.3, where A and B are dense subsets of $\{1,\dots,N\}$: we may simply embed the sets in $\mathbb{Z}_{N'}$, where N' is a prime between $4N$ and $8N$. For the small doubling case we use Proposition 5.1. Specifically, let us assume without loss of generality that $K_B \leqslant K_A$. Applying Proposition 5.1 to A and B with $k=2$ and N a prime less than $CK_A^C|B|$, as allowed by Lemma 5.2, we obtain sets $A' \subseteq A$, $B' \subseteq B$ with $|A'| \geqslant |A|/4$, $|B'| \geqslant |B|/4$ and a Freiman 2-isomorphism ϕ from $A'+B'$ to a subset of $\mathbb{Z}_N$. We shall find our desired arithmetic progression in $A'+B'$. We may assume, by translating if necessary, that A' and B' both contain 0 and that ϕ takes 0 to 0. Hence we have $\phi(A'+B') = \phi(A')+\phi(B')$ and so $|\phi(A')+\phi(B')| \leqslant 4K_A|\phi(A')|, 4K_B|\phi(B')|$. Since the property of containing an arithmetic progression of a given size is preserved under 2-isomorphism, it suffices (after adjusting the constants) to establish the theorem for $\phi(A')$ and $\phi(B')$, which we now relabel as A and B.

The argument is now similar to that in the proof of Theorem 1.1 given in Section 3, the main difference being that we take into account the higher-order energy $\|\mu_A * 1_B\|_{p/2}$ present in the conclusion of Theorem 6.2. Let us apply Theorem 6.2 with $\epsilon = 1/e\sqrt{K_B}$ and a parameter p to be specified later; since B has density at least cK_A^{-C} in $\mathbb{Z}_N$, this gives us a Bohr set T of rank $d \leqslant CpK_B(\log 2K_A)^3$ and radius at least $c/K_A d$ such that

$$\|\mu_A * 1_B(x+t) - \mu_A * 1_B(x)\|_{L^p(x)} < \epsilon\|\mu_A * 1_B\|_{p/2}^{1/2} \quad \text{for each } t \in T.$$

Let P be an arithmetic progression in T. Assume for a contradiction that $A+B$ does not contain a translate of P, so that for each $x \in \mathbb{Z}_N$ there is some $t \in P$ for which $\mu_A * 1_B(x+t) = 0$. Then

$$\begin{aligned}|P|\epsilon^p\|\mu_A * 1_B\|_{p/2}^{p/2} &> \sum_{t\in P}\|\mu_A * 1_B(x+t) - \mu_A * 1_B(x)\|_{L^p(x)}^p \\ &\geqslant \|\mu_A * 1_B\|_p^p \geqslant \|\mu_A * 1_B\|_{p/2}^p/\mu_G(A+B),\end{aligned}$$

the last inequality being an instance of the Cauchy–Schwarz inequality. We shall thus have a contradiction if $|P|$ and p are chosen so that

$$|P| \leqslant e^p K_B^{p/2}\|\mu_A * 1_B\|_{p/2}^{p/2}/\mu_G(A+B).$$

Since $\|\mu_A * 1_B\|_{p/2} \geqslant \|\mu_A * 1_B\|_1/\mu_G(A+B)^{1-2/p}$ by Hölder's inequality, it suffices to choose $|P| \leqslant e^p$ to obtain the contradiction. Now, by Lemma 2.1 we can find an arithmetic progression

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

P in T of size the integer part of

$$\exp\left(c\frac{\log N}{pK_B(\log 2K_A)^3} - C\log(pK_A)\right).$$

Choosing $p = C\sqrt{\log N/K_B(\log 2K_A)^3}$ then yields the conclusion, since $|A+B| \leqslant N \leqslant CK_A^C|A|$. □

7. Further remarks

7.1. Vector spaces over finite fields

A very useful philosophy in additive combinatorics, particularly as expounded and well described by Green [6], is that it can often be simpler to study a given additive problem, perhaps not in its original setting but rather in a particular family of groups, namely the vector spaces over finite fields $\mathbb{F}_q$. One of the simplifications that this setting allows over that of general abelian groups is the replacement of Bohr sets with subspaces, which are exact rather than approximate annihilators of sets of characters, and this makes many arguments simpler; this is the case in [4] and [12], for example. Breaking the trend, the methods of this paper hardly change when looked at in vector spaces. The only significant change occurs in the proofs of Theorems 1.1 and 1.3: instead of taking P to be an arithmetic progression in the Bohr set T, one can let P be a *subspace* of the *subspace* T. Making this change gives the following analogue of Theorem 1.1, a result originally established (in the $q = 2$ case) by Green [5] and then by Sanders [12].

Theorem 7.1. *Suppose A and B are subsets of $\mathbb{F}_q^n$ of densities α and β. Then $A + B$ contains a translate of a subspace of dimension at least $c\alpha\beta n/\log q$.*

The argument used in the proof of Theorem 1.3 similarly yields the following version with improved density dependence.

Theorem 7.2. *Suppose A and B are subsets of $\mathbb{F}_q^n$ of densities α and β. Then $A + B$ contains a translate of a subspace of dimension at least $c\alpha(\log 2\beta^{-1})^{-3}n/\log q$.*

In the case when $B = -A$, it is known from further work of Sanders [13] that one can remove the logarithmic factor. It has furthermore been suggested by Green [6] that the true nature of the bound should be one of an upper bound on codimension rather than a lower bound on dimension, so Theorem 7.2 might be quite far from the truth.

One need not specialize P to be a subspace in the proofs, of course; the more general result one obtains is then as follows.

Theorem 7.3. *Suppose A and B are subsets of $\mathbb{F}_q^n$ of densities α and β. Then, for any $d \in \mathbb{N}$, there is a subspace $V \leqslant \mathbb{F}_q^n$ of codimension d such that $A + B$ contains a translate of any subset of V of size at most $\exp\big(c\alpha(\log 2\beta^{-1})^{-3}d\big)$.*

7.2. Remarks on the bootstrapping procedure

Exactly the same smoothing-by-convolution argument that deduced Theorem 6.2 from Theorem 6.1 can be used to deduce the following result from Theorem 4.1.

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

Theorem 7.4. *Let $p \geqslant 2$ and $\epsilon \in (0,1)$. Let G be a finite abelian group and let A, B and S be finite subsets of G with $|A+S| \leqslant K_1|A|$ and $|A+B| \leqslant K_2|B|$. Then there is a Bohr set T of rank at most*

$$d = Cp(\log(1/\delta))^2\epsilon^{-2}\log(2K_1) + C\log(\mu_G(S)^{-1})$$

and radius at least δ/d, where

$$\delta = c\epsilon\sqrt{\mu_G(A)/\mu_G(B)}K_2^{-1/p},$$

such that, for each $t \in T$,

$$\|\mu_A * 1_B(x+t) - \mu_A * 1_B(x)\|_{L^p(x)} \leqslant \epsilon\mu_G(B)^{1/p}.$$

Although it looks considerably more technical, this theorem has the advantage over Theorem 4.1 that the set of almost-periods it gives is more easily seen to be structured than the iterated sumset given by Theorem 4.1. Applying this with $S = A$, $B = A - A$, $\epsilon = 1/2$ and $p = C\log(2K)$, where K is the doubling constant for A, it is a quick deduction to recover Theorems 11.1 and 11.3 of Sanders [14], namely that $2A - 2A$ contains a low-rank coset progression in groups with 'good modelling', like $\mathbb{Z}$ and $\mathbb{F}_p^n$. Indeed, for any almost-period $t \in T$ we have

$$\begin{aligned}|\mu_{-A} * \mu_A * 1_{A-A}(t) - 1| &= |\mu_{-A} * \mu_A * 1_{A-A}(t) - \mu_{-A} * \mu_A * 1_{A-A}(0)| \\ &\leqslant \|\mu_A\|_{L^q}\|\mu_A * 1_{A-A}(x+t) - \mu_A * 1_{A-A}(x)\|_{L^p(x)} \\ &\leqslant c,\end{aligned}$$

whence $t \in 2A - 2A$. Since T contains a large coset progression, we are done. We stress that this does not offer a quantitative improvement over Sanders' result, but we include the remark for clarification and since it follows a slightly different route, more direct than that in [14], avoiding the need to deal with correlations and deducing containment from this. We finish by recording the rather clean 'density version' of the result obtained in this fashion.

Theorem 7.5. *Let G be a finite abelian group and suppose $A \subseteq G$ has size $\alpha|G|$. Then $2A - 2A$ contains a Bohr set of radius at least $c\alpha^{1/2}$ and rank at most $C(\log 1/\alpha)^4$.*

Acknowledgement

The third author would like to acknowledge the hospitality of the University of British Columbia, where part of this work was done.

References

[1] Bourgain, J. (1990) On arithmetic progressions in sums of sets of integers. In *A Tribute to Paul Erdős*, Cambridge University Press, pp. 105–109.
[2] Chang, M.-C. (2002) A polynomial bound in Freiman's theorem. *Duke Math. J.* **113** 399–419.
[3] Croot, E. and Sisask, O. (2010) A probabilistic technique for finding almost-periods of convolutions. *Geom. Funct. Anal.* **20** 1367–1396.
[4] Green, B. (2002) Arithmetic progressions in sumsets. *Geom. Funct. Anal.* **12** 584–597.

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press

[5] Green, B. (2002) *Restriction and Kakeya Phenomena*, lecture notes. http://www.dpmms.cam.ac.uk/~bjg23/rkp.html
[6] Green, B. (2005) Finite field models in additive combinatorics. In *Surveys in Combinatorics 2005*, Vol. 327 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, pp. 1–27.
[7] Gut, A. (2005) *Probability: A Graduate Course*, Springer.
[8] Khintchine, A. (1923) Über dyadische Brüche. *Math. Z.* **18** 109–116.
[9] Petridis, G. (2013) New proofs of Plünnecke-type estimates for product sets in groups. *Combinatorica*, doi: 10.1007/s00493-012-2818-5.
[10] Pisier, G. (1981) Remarques sur un résultat non publié de B. Maurey. In *Seminar on Functional Analysis, 1980–1981*, exp. no. 5, École Polytechnique, Palaiseau. http://www.numdam.org/item?id=SAF_1980-1981____A5_0
[11] Ruzsa, I. Z. (2009) Sumsets and structure. In *Combinatorial Number Theory and Additive Group Theory*, Springer, pp. 87–210.
[12] Sanders, T. (2008) Additive structures in sumsets. *Math. Proc. Cambridge Philos. Soc.* **144** 289–316.
[13] Sanders, T. (2011) Green's sumset problem at density one half. *Acta Arith.* **146** 91–101.
[14] Sanders, T. (2012) On the Bogolyubov–Ruzsa lemma. *Anal. PDE* **5** 627–655.
[15] Tao, T. and Vu, V. H. (2006) *Additive Combinatorics*, Cambridge University Press.

https://doi.org/10.1017/S0963548313000060 Published online by Cambridge University Press