Combinatorics, Probability and Computing (2013) 22, 282–293. © Cambridge University Press 2012
doi:10.1017/S0963548312000466

Sums of Dilates in $\mathbb{Z}_p$

GONZALO FIZ PONTIVEROS

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ, Brasil
(e-mail: gf232@cam.ac.uk)

Received 14 March 2012; revised 10 September 2012; first published online 30 October 2012

We consider the problem of sums of dilates in groups of prime order. It is well known that sets with small density and small sumset in $\mathbb{Z}_p$ behave like integer sets. Thus, given $A \subset \mathbb{Z}_p$ of sufficiently small density, it is straightforward to show that

$$|\lambda_1 A + \lambda_2 A + \cdots + \lambda_k A| \geqslant \left(\sum_i |\lambda_i|\right)|A| - o(|A|).$$

On the other hand, the behaviour for sets of large density turns out to be rather surprising. Indeed, for any $\epsilon > 0$, we construct subsets of density $1/2 - \epsilon$ such that $|A + \lambda A| \leqslant (1 - \delta)p$, showing that there is a very different behaviour for subsets of large density.

AMS 2010 *Mathematics subject classification*: Primary 11B13; 05B10

1. Introduction

Combinatorial number theory is a field of mathematics which deals with the additive and multiplicative structure of sets, and encompasses techniques from harmonic analysis, ergodic theory, graph theory and number theory (see Tao and Vu [16], for example). A central notion in this field is that of a *sumset*. Given an abelian group G, a finite subset $A \subset G$ and integers $\lambda_1, \ldots, \lambda_k \in \mathbb{Z}$, consider the set

$$X = \lambda_1 A + \lambda_2 A + \cdots + \lambda_k A,$$

where $A + B = \{a + b : a \in A,\ b \in B\}$ and $\lambda A = \{\lambda a : a \in A\}$. Finding lower bounds for the size of X (in terms of $|A|$) is a basic and central question in the field.

For the group $G = \mathbb{Z}$, the situation is fairly well understood. It is easy to prove that $|A + A| \geqslant 2|A| - 1$, with equality if and only if A is an arithmetic progression. Plagne and Hamidoune [8] gave the lower bound $|A + \lambda A| \geqslant 3|A| - 2$ for $\lambda \notin \{-1, 0, 1\}$. Later on, for $|\lambda| \geqslant 3$, Nathanson [11] showed that $|A + \lambda A| \geqslant \frac{7}{2}|A| - \frac{5}{2}$. The case $\lambda = 3$ was completely solved in [1] and [4]; in the latter, Cilleruelo, Silva and Vinuesa also characterize all of the extremal sets.

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

In the general case, the great breakthrough is due to Bukh, who gave an asymptotically sharp bound in [1].

Theorem 1.1 (Bukh, 2008). *For any k coprime integers $\lambda_1, \lambda_2, \ldots, \lambda_k$ and any finite set $A \subset \mathbb{Z}$, we have*

$$|\lambda_1 A + \lambda_2 A + \cdots + \lambda_k A| \geqslant \left(\sum_i |\lambda_i| \right) |A| - o(|A|), \tag{1.1}$$

where the error term $o(|A|)$ depends on $\lambda_1, \ldots, \lambda_k$.

However, in the finite field setting $G = \mathbb{Z}_p$, the problem turns out to be much harder, and almost nothing is known. The only exception is the case $k = 2$ and $\lambda_1 = \lambda_2 = 1$, for which we have the celebrated *Cauchy–Davenport theorem* [2, 5].

Cauchy–Davenport theorem. Let p be prime, and let $A, B \subset \mathbb{Z}_p$ be non-empty. Then

$$|A + B| \geqslant \min\{|A| + |B| - 1, p\}.$$

Thus, in particular, we have the bound

$$|A + A| \geqslant 2|A| - 1 \ \text{ for every } \ |A| \leqslant \frac{p-1}{2}.$$

Furthermore, Vosper [17] showed that, as in the case $G = \mathbb{Z}$, when $|A| + |B| \leqslant p$ equality holds if and only if both A and B are arithmetic progressions with the same common difference.

In light of the two theorems above, it seems natural to conjecture that if p is prime, $\lambda_1, \lambda_2 \in \mathbb{Z}$ and $A \subseteq \mathbb{Z}_p$, then

$$|\lambda_1 A + \lambda_2 A| \geqslant \min\{(|\lambda_1| + |\lambda_2|)|A| - O(1),\ p\}. \tag{1.2}$$

However, we have shown that (1.2) fails in a very strong sense. We prove the following result.

Theorem 1.2. *For every $\lambda \in \mathbb{Z}$ and $\epsilon > 0$, there exists $\delta = \delta(\lambda, \epsilon) > 0$ such that the following holds. If p is a sufficiently large prime, then there exists a set $A \subset \mathbb{Z}_p$ with $|A| \geqslant (1/2 - \epsilon)p$ such that $|A + \lambda A| \leqslant (1 - \delta)p$.*

Although the result is surprising, the proof of Theorem 1.2 is relatively straightforward. The key step is to transfer the problem to the continuous setting, where one can apply a known result on the dynamics of measure-preserving maps in the circle. In the general setting of [1] we make the following conjecture.

Conjecture 1.3. *For every $\lambda_1, \ldots, \lambda_k \in \mathbb{Z}$ and $\epsilon > 0$, there exists $\delta > 0$ such that the following holds. For every sufficiently large prime p, there exists a set $A \subset \mathbb{Z}_p$ with $|A| \geqslant (1/k - \epsilon)p$ such that*

$$|\lambda_1 A + \lambda_2 A + \cdots + \lambda_k A| \leqslant (1 - \delta)p.$$

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

On the other hand, if A is much smaller then we show that the bound (1.2) *does* hold.

Theorem 1.4. *For every coprime $\lambda_1, \dots, \lambda_k \in \mathbb{Z}$, there exists a constant $\alpha > 0$ such that*

$$|\lambda_1 A + \lambda_2 A + \cdots + \lambda_k A| \geqslant \left(\sum_i |\lambda_i| \right) |A| - o(|A|),$$

for every sufficiently large prime p, and every $A \subset \mathbb{Z}_p$ with $|A| \leqslant \alpha p$. The error term $o(|A|)$ depends on $\lambda_1, \dots, \lambda_k$, but not on p.

Very recently, Plagne [12] has given bounds in the same spirit for sumsets of the form $A + \lambda A$. Namely, it was shown that there exists a function f_λ and a constant $w(\lambda)$ such that the following holds.

Theorem 1.5 (Plagne [12]).

$$|A + \lambda A| \geqslant \min\big(f_\lambda(\alpha)|A| - w(\lambda), p\big), \tag{1.3}$$

where $|A| = \alpha p$ and $f_\lambda(\alpha)$ is defined to be the maximum of 2 and the unique solution to the equation

$$(\lambda + 1) \sin\left(\frac{\pi}{\lambda + 1}\right)(1 - \alpha x) = x^{3/2} \sin\left(\frac{\pi}{x}\right).$$

The range of densities α for which the function f_λ yields non-trivial bounds (*i.e.*, $f_\lambda(\alpha) > 2$) is much bigger than the one obtained in Theorem 1.4, and tends to $\frac{1}{2} - \frac{\sqrt{2}}{\pi} \approx 0.0498$ as $|\lambda| \to \infty$. However, the value of $f_\lambda(\alpha)$ is far from the truth when looking at very small α, indeed $f_\lambda(0) < 2.16$ for all $\lambda \in \mathbb{Z}$. So the bound for large λ is essentially

$$|A + \lambda A| \geqslant \min\big(2.16|A|, 0.0996p\big),$$

whereas our method yields

$$|A + \lambda A| \geqslant \min\big((\lambda + 1)|A|, \epsilon(\lambda)p\big).$$

Hence the bounds are in a sense complementary.

2. Small densities

The strategy for dealing with sets of small density is to show that they in fact behave like subsets of the integers with respect to addition, and thus we may transfer the relevant bounds. This kind of rectification technique was first introduced by Freiman [6] in the proof of his well-known theorem, and is now a standard tool for transferring results from the integer setting.

Suppose that the set A is contained in the interval $\{0, \dots, \lfloor \frac{p}{M} \rfloor\}$, where $M = \sum |\lambda_i|$. It is clear that when looking at sums of the form $\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_k a_k$ for $a_1, a_2, \dots, a_k \in A$, there is never any 'wrap-around', and hence we might as well think of them as happening in $\mathbb{Z}$. We make this statement more precise as follows.

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

Definition. **(1)** The *diameter* of a subset $A \subseteq \mathbb{Z}_p$, denoted by $l(A)$, is defined as the smallest integer l for which there exists some $x, d \in \mathbb{Z}_p$ such that $A \subseteq \{x, x+d, \ldots, x+(l-1)d\}$: in words, the length of the shortest arithmetic progression containing A.

(2) Let A and B be finite subsets of abelian groups (not necessarily equal). We say that A and B are *m-Freiman isomorphic* if there exists a bijection $f : A \to B$ such that

$$\begin{gathered} f(a_1) + f(a_2) + \cdots + f(a_m) = f(a_1') + f(a_2') + \cdots + f(a_m') \\ \Updownarrow \\ a_1 + \cdots + a_m = a_1' + \cdots + a_m'. \end{gathered}$$

Formally speaking, if $l(A) < \frac{1}{M}p$, where $M = \sum_1^k |\lambda_i|$, then there exists $\tilde{A} \subset \mathbb{Z}$ such that $\tilde{A}$ is M-Freiman isomorphic to A. In particular, it follows that

$$|\lambda_1\tilde{A} + \lambda_2\tilde{A} + \cdots + \lambda_k\tilde{A}| = |\lambda_1 A + \lambda_2 A + \cdots + \lambda_k A|,$$

and therefore we are free to apply any known bounds in the integer setting.

Corollary 2.1. *Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be relatively prime integers and let $A \subset \mathbb{Z}_p$ be a subset with diameter $l(A) < \frac{1}{M}p$ where $M = \sum_1^k |\lambda_i|$. Then*

$$|\lambda_1 A + \lambda_2 A + \cdots + \lambda_k A| \geqslant \big(|\lambda_1| + |\lambda_2| + \cdots + |\lambda_k|\big)|A| - o(|A|), \tag{2.1}$$

where the $o(|A|)$ is the error term given by Theorem 1.1.

Remark. In the case $k = 2$, $\lambda_1 \in \{1, 2\}$ and $\lambda_2 \in \{\pm q\}$ for q a prime, very precise error terms of the form $g(\lambda_1, \lambda_2)$ are given in [3] and [9] for the integer case and hence are also transferable to $\mathbb{Z}_p$.

A priori, we can only guarantee that $l(A) \leqslant \frac{1}{k}p$ for sets with $|A| < \frac{\log p}{\log k}$ (using the pigeonhole principle), and this bound is best possible up to a constant: for example, a random subset of $\mathbb{Z}_p$ of size $4 \log p$ has diameter larger than $p/2$ with high probability.

However, if we are given the additional information that the sumset $A + A$ is small, then it is possible to bound well the diameter of the set A. Freiman [6] showed that for all $C > 0$ there exists $\epsilon(C) > 0$ such that, if $|A + A| \leqslant C|A|$ and $|A| < \epsilon(C)p$, then the set A is Freiman isomorphic to a subset of the integers. More recently, Green and Ruzsa [7] showed the same result with a different approach that yielded much better bounds. We will use their formulation.

Theorem 2.2 (Green and Ruzsa [7]). *Let A be a subset of $\mathbb{Z}_p$, p a prime, with $|A| = \alpha N$, and assume that $\min\{|A + A|, |A - A|\} \leqslant K|A|$. If $\alpha \leqslant (16K)^{-12K^2}$, then the diameter of A is at most*

$$12\alpha^{1/4}\sqrt{\log(1/\alpha)}N. \tag{2.2}$$

Equipped with this tool, it is now a simple task to deduce a bound for sets of small density by transferring the known $\mathbb{Z}$ bounds.

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

Corollary 2.3. *For every $\bar{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_k)$, there exists some $\alpha(\bar{\lambda}) > 0$ such that, for all $A \subset \mathbb{Z}_p$ with $|A| \leqslant \alpha p$,*

$$|\lambda_1 A + \lambda_2 A + \cdots + \lambda_k A| \geqslant \big(|\lambda_1| + |\lambda_2| + \cdots + |\lambda_k|\big)|A| - o(|A|).$$

Proof. Let $M = \sum |\lambda_i|$. We may assume that $|\lambda_1 A + \lambda_2 A + \cdots + \lambda_k A| \leqslant M|A|$ as otherwise we are done. We make use of the following variation of an elementary inequality due to Ruzsa [15].

Ruzsa's triangle inequality (sum form: see [14]). Given any B, C, D finite subsets of some abelian group G, we have

$$|B + D| \leqslant \frac{|B + C||C + D|}{|C|}. \tag{2.3}$$

Setting $B = D = \lambda_1 A$ and $C = \lambda_2 A + \cdots + \lambda_k A$, we have that

$$|A + A| = |\lambda_1 A + \lambda_1 A| \leqslant \frac{|\lambda_1 A + \cdots + \lambda_k A|^2}{|\lambda_2 A + \cdots + \lambda_k A|} \leqslant M^2|A|. \tag{2.4}$$

Now pick $\alpha > 0$ such that $12\alpha^{1/4}\sqrt{\log(1/\alpha)} < M^{-1}$ and $\alpha < (16M^2)^{-12M^4}$. By the Green–Rusza theorem, any set $A \subset \mathbb{Z}_p$ with $|A| \leqslant \alpha p$ and $|A + A| \leqslant M^2|A|$ will have diameter at most $M^{-1}p$, and hence it follows from Corollary 2.1 that

$$|\lambda_1 A + \cdots + \lambda_k A| \geqslant \big(|\lambda_1| + \cdots + |\lambda_k|\big)|A| - o(|A|),$$

as required. □

3. Large densities

We begin with a simple lemma that already shows that the behaviour of sets of large density is indeed quite different from the integers. Indeed, if (1.2) were true, then any set with density greater than $\frac{1}{|\lambda_1|+|\lambda_2|}$ ought to satisfy $\lambda_1 A + \lambda_2 A = \mathbb{Z}_p$. Instead we have the following result.

Lemma 3.1. *For any prime p, we can find a set $A \subset \mathbb{Z}_p$ such that $|A| = \frac{1}{2}p - o(p)$ and $0 \notin \lambda_1 A + \lambda_2 A$.*

Proof. Assume that $p > |\lambda_1| + |\lambda_2|$ and define a digraph Γ on the vertex set $\mathbb{Z}_p \setminus \{0\}$, where $x \to y$ if and only if $\lambda_1 x + \lambda_2 y = 0$. We are done if we find a large independent set in Γ. Now, for any $x \in \mathbb{Z}_p$ we have that $d_-(x) = d_+(x) = 1$, and hence Γ is the disjoint union of directed cycles.

Suppose we have a cycle $x_1 \to x_2 \to \cdots \to x_k \to x_1$. By definition of Γ we have that $\lambda_2 x_{i+1} = (-\lambda_1)x_i \bmod p$ for all $1 \geqslant i \geqslant k - 1$. Hence k can only be the smallest positive integer for which we have that

$$(-\lambda_1)^k \equiv \lambda_2^k \bmod p.$$

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

Certainly the above can only happen if $k \geqslant \lfloor \log_{\max\{|\lambda_1|,|\lambda_2|\}} p \rfloor - 1$. Thus all cycles in Γ have the same length,

$$k = \Omega(\log p).$$

This is good news since it is simple to find independent sets in cycles by picking alternative vertices, that is, as many as $\lfloor k/2 \rfloor$ vertices in each cycle. In this way we obtain an independent set A of density

$$\frac{\lfloor k/2 \rfloor}{k} \geqslant \frac{1}{2} - O\left(\frac{1}{\log p}\right),$$

and therefore

$$|A| \geqslant \frac{1}{2}p - O\left(\frac{p}{\log p}\right),$$

as claimed. □

Note that, on the other hand, the Cauchy–Davenport theorem gives us the lower bound

$$|\lambda_1 A + \lambda_2 A| \geqslant \min\{2|A| - 1, p\},$$

and thus any set of density above $1/2$ must satisfy that $\lambda_1 A + \lambda_2 A = \mathbb{Z}_p$.

One might wonder if this is some kind of exception, and if instead we demand that $|\lambda_1 A + \lambda_2 A| \leqslant p - 2$ then we have a non-trivial upper bound on the density of A, meaning that it is at most $1/2 - \epsilon$ for some $\epsilon > 0$. We shall show that this is not the case: our strategy is to transfer the problem into a continuous setting. The following proposition makes this transference explicit.

Lemma 3.2. *Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, equipped with the Lebesgue measure μ and $\lambda_1, \lambda_2 \in \mathbb{Z}$. For any $\alpha \in [0,1]$ the following are equivalent.*

(i) *Given any $\epsilon_1 > 0$, there exists a measurable set $A \subset \mathbb{T}$ with $\mu(A) \geqslant \alpha - \epsilon_1$ such that $\lambda_1 A \cap \lambda_2 A = \emptyset$.*

(ii) *Given any $\epsilon_2 > 0$, there exists p_0 and $\delta > 0$ such that, for all primes $p \geqslant p_0$ and any $0 \leqslant t \leqslant \delta p$, there exists $A \subset \mathbb{Z}_p$ with $|A| \geqslant (\alpha - \epsilon_2)p$, and satisfying $\{-t, -t+1, \ldots, 0, 1, \ldots, t\} \notin \lambda_1 A - \lambda_2 A$.*

Proof. **(ii) ⇒ (i)** We begin with the easier implication (ii) ⇒ (i). By assumption, given any $\epsilon > 0$ we may find a prime p and a set $A \subset \mathbb{Z}_p$ such that $|A| \geqslant (\alpha - \epsilon)p$ and such that the interval

$$\left\{ -\left\lceil \frac{|\lambda_1| + |\lambda_2|}{2} \right\rceil, \ldots, \left\lceil \frac{|\lambda_1| + |\lambda_2|}{2} \right\rceil \right\} \notin \lambda_1 A - \lambda_2 A.$$

We embed $\mathbb{Z}_p$ into $\mathbb{T}$ in the natural way, namely via the homomorphism $\chi_p : x \mapsto x/p$. Let $\|\cdot\|$ denote the distance to the closest integer, and define

$$\tilde{A} = \left\{ \theta \in \mathbb{T} : \exists\, x \in A \text{ s.t. } \left\| \theta - \frac{x}{p} \right\| < \frac{1}{2p} \right\}. \tag{3.1}$$

In words, $\tilde{A}$ is the disjoint union of open intervals of length $1/p$, centred at $\chi_p(x)$ for $x \in A$.

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

It is clear that $\mu(\tilde{A}) = |A|(1/p) \geqslant \alpha - \epsilon$. It remains to show that $\lambda_1\tilde{A} \cap \lambda_2\tilde{A} = \emptyset$. Suppose that $\theta \in \lambda_1\tilde{A} \cap \lambda_2\tilde{A}$. Then there exist $x, y \in A$ such that

$$\left\|\theta - \frac{\lambda_1 x}{p}\right\| < \frac{|\lambda_1|}{2p} \quad \text{and} \quad \left\|\theta - \frac{\lambda_2 y}{p}\right\| < \frac{|\lambda_2|}{2p}.$$

Thus, by the triangle inequality,

$$\left\|\frac{\lambda_1 x - \lambda_2 y}{p}\right\| < \frac{|\lambda_1| + |\lambda_2|}{2p},$$

but since x and y are integers this can only happen if there is a solution in A to the equation $\lambda_1 x - \lambda_2 y = r \mod p$, for some

$$r \in \left\{-\left\lceil \frac{|\lambda_1| + |\lambda_2|}{2} \right\rceil, \ldots, \left\lceil \frac{|\lambda_1| + |\lambda_2|}{2} \right\rceil\right\}.$$

This is a contradiction to our choice of A and thus $\lambda_1\tilde{A} \cap \lambda_2\tilde{A}$ must be empty.

(i) ⇒ (ii) Suppose that there exists a measurable set A of $\mathbb{T}$ with $\mu(A) \geqslant \alpha - \epsilon$ such that $\lambda_1 A \cap \lambda_2 A = \emptyset$. Since A is measurable, we may find a $A' = \bigcup_{i=1}^{K} [a_i, b_i]$ such that $\mu(A \setminus A') + \mu(A' \setminus A) < \epsilon$. Certainly $\mu(A') \geqslant \alpha - 2\epsilon$, and furthermore

$$\begin{aligned}\mu(\lambda_1 A' \cap \lambda_2 A') &\leqslant \mu\big(\lambda_1(A' \setminus A) \cap \lambda_2 A'\big) + \mu\big(\lambda_1 A' \cap \lambda_2(A' \setminus A)\big) \\ &\leqslant |\lambda_1|\mu\big(A' \setminus A\big) + |\lambda_2|\big(A' \setminus A\big) < \big(|\lambda_1| + |\lambda_2|\big)\epsilon.\end{aligned}$$

Now we 'discretize' the circle in the obvious way,

$$A_p = \{x : x \in \{0, \ldots, p-1\}, x/p \in A'\},$$

and view it as a subset of $\mathbb{Z}_p$. As A' is the union of K intervals, we also have the bounds

$$|A_p| - K \leqslant \mu(A')p \leqslant |A_p| + K. \tag{3.2}$$

Next, we claim that the set A_p is close to satisfying the desired property.

Claim.

$$\left|\bigcup_{-t}^{t} \lambda_1 A_p \cap (\lambda_2 A_p + i)\right| \leqslant \big(|\lambda_1| + |\lambda_2|\big)\epsilon p + (2t+1)K^2 + 2tK.$$

First of all, note that $\lambda_1 A'$ and $\lambda_2 A'$ are also the union of at most K intervals, and hence the set $\lambda_1 A' \cap \lambda_2 A'$ may certainly be expressed as the union of at most K^2 intervals.

We also have the following inequality:

$$\mu\left(\lambda_2 A' + \left[-\frac{t}{p}, \frac{t}{p}\right] \setminus \lambda_2 A'\right) \leqslant 2tK/p. \tag{3.3}$$

Thus,

$$\begin{aligned}\mu(\lambda_1 A' \cap \lambda_2 A' + \{-t/p, \ldots, t/p\}) &\leqslant \mu(\lambda_1 A' \cap \lambda_2 A') + \mu(\lambda_2 A' + [-t/p, t/p] \setminus \lambda_2 A') \\ &< \big(|\lambda_1| + |\lambda_2|\big)\epsilon + 2tK/p.\end{aligned}$$

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

Now, applying equation (3.2) to the set $\lambda_1 A' \cap (\lambda_2 A' + \{-t/p, \dots, t/p\})$, which is the union of at most $(2t+1)K^2$ intervals, we obtain

$$|\lambda_1 A_p \cap \big(\lambda_2 A_p + \{-t, \dots, t\}\big)| \leqslant \big(|\lambda_1| + |\lambda_2|\big)\epsilon p + (2t+1)K^2 + 2tK.$$

Hence, by deleting at most $\big(|\lambda_1| + |\lambda_2|\big)\epsilon p + (2t+1)K^2 + 2tK$ points from A_p (namely λ_1^{-1} times those in $\lambda_1 A_p \cap (\lambda_2 A_p + \{-t, \dots, t\})$, we can find a set A in $\mathbb{Z}_p$ such that $-t, \dots, t \notin \lambda_1 A - \lambda_2 A$ and

$$|A| \geqslant \big(\alpha - (|\lambda_1| + |\lambda_2|)\epsilon\big)p - O_K(\delta)p \geqslant \big(\alpha - 2(|\lambda_1| + |\lambda_2|)\epsilon\big)p,$$

provided we have chosen δ sufficiently small. □

The task now is that of finding sets of large (Lebesgue) measure in $\mathbb{T}$ such that $\lambda_1 A \cap \lambda_2 A = \emptyset$. First, note that we trivially have the upper bound $\mu(A) \leqslant 1/2$ since $\mu(\lambda_i A) \geqslant \mu(A)$.

We address below the special case where $\lambda_1 = 1$, and begin by showing that in fact we cannot have that $\mu(A) = 1/2$.

Proposition 3.3. *If $\lambda \in \mathbb{Z}$ with $|\lambda| \geqslant 2$ and A is a subset of $\mathbb{T}$ such that $A \cap \lambda A = \emptyset$, then $\mu(A) < 1/2$.* □

Proof. Suppose that there exists such a subset $A \subset \mathbb{T}$ with $\mu(A) = 1/2$. Then $(-\lambda)A = \bar{A} = \mathbb{T} \setminus A$ (modulo a set of null measure) and, since multiplication by $-\lambda$ is a measure-preserving transformation of the circle, it also follows that $\lambda^2 A = A$. However, it is well known that multiplication by an integer is an ergodic transformation, that is to say, the only invariant subsets have either null or full measure. This is a contradiction. □

We will again consider the action of multiplication by $-\lambda$ as an ergodic, measure-preserving transformation of $\mathbb{T}$. The following result was first proved by Rokhlin in [13]. We will not need the full generality of Rokhlin's lemma as the transformations we are interested in are ergodic. We include here a proof in the setting of ergodic transformations of the circle, as it is very short and simple, and it also contains the underlying ideas of the 'quantitative' version we give in the following section. For the more general result, we recommend the reader to see [10], which presents a remarkably simple proof of Rokhlin's lemma, even when the transformations are not invertible.

Lemma 3.4 (Rokhlin's lemma). *Let $\phi : \mathbb{T} \to \mathbb{T}$ be an ergodic, measure-preserving, measurable map. Given any $n \in \mathbb{N} \setminus \{0\}$ and any $\epsilon > 0$, there exists a measurable set $B \subset \mathbb{T}$ such that:*

(i) $\big(\phi^i(B)\big)_{0 \leqslant i \leqslant n-1}$ *are pairwise disjoint,*
(ii) $\mu(B) \geqslant \frac{1}{n} - \epsilon$.

Proof. Pick any measurable set $E \subset \mathbb{T}$ with $0 < \mu(E) \leqslant \epsilon$, for example the open ball centred at 0 of radius $\epsilon/2$. We construct a family of sets as follows:

$$E_0 = E; \qquad E_{i+1} = \phi^{-1}(E_i) \setminus \bigcup_{0 \leqslant j \leqslant i} E_j. \tag{3.4}$$

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

The sets $(E_i)_{i\geqslant 0}$ are, by construction, pairwise disjoint and for any $i \in \mathbb{N}$ we have the inclusion

$$\phi(E_{i+1}) \subseteq E_i. \tag{3.5}$$

Claim.

$$\bigcup_{i\geqslant 0} E_i = \mathbb{T} \qquad \mu\text{-almost everywhere.} \tag{3.6}$$

Set $F = \bigcup_{i\geqslant 0} E_i = \bigcup_{i\geqslant 0} \phi^{-i}(E_0)$, and note that

$$\phi^{-1}(F) = \bigcup_{i\geqslant 1} \phi^{-i}(E_0) \subseteq F.$$

On the other hand, as ϕ is measure-preserving, $\mu(\phi^{-1}(F)) = \mu(F) \Rightarrow \phi^{-1}(F) = F$. Thus, by ergodicity, either $\mu(F) = 0$ or $\mu(F) = 1$. But F contains E, which has a strictly positive measure, so it must be the case that $\mu(F) = 1$.

Let $B = \bigcup_{j\geqslant 1} E_{jn}$. We claim that B meets the requirements of Rokhlin's lemma.

(i) For convenience, let J_r denote the set $\{k \in \mathbb{N} \setminus \{0\} : k \equiv r \mod n\}$. From (3.5) we have that, for each $i = 0, 1, \ldots, n-1$,

$$\phi^i(B) = \phi^i\left(\bigcup_{j\in J_0} E_j\right) = \bigcup_{j\in J_0} \phi^i(E_j) \subseteq \bigcup_{j\in J_0} E_{j-i} = \bigcup_{j\in J_{n-i}} E_j, \tag{3.7}$$

and therefore the sets $\left(\phi^i(B)\right)_{i=0}^{n-1}$ are pairwise disjoint.

(ii) We have the inclusion

$$\bigcup_{i=0}^{n-1} \phi^{-i}(A) \supseteq \bigcup_{i\geqslant 0} E_i \setminus \bigcup_{i\geqslant 0}^{n-1} E_i = \mathbb{T} \setminus \bigcup_{i\geqslant 0}^{n-1} E_i. \tag{3.8}$$

Now recall that we chose E_0 with $\mu(E_0) \leqslant \epsilon$, hence

$$\mu\left(\bigcup_{i=0}^{n-1} \phi^{-i}(A)\right) = n\mu(B) \geqslant 1 - n\epsilon,$$

and therefore $\mu(B) \geqslant \frac{1}{n} - \epsilon$. □

As an immediate consequence of Rokhlin's lemma we have the following result.

Proposition 3.5. *For any given $\epsilon > 0$, there exists a set $A \subset \mathbb{T}$ such that $A \cap \lambda A = \emptyset$ and $\mu(A) > 1/2 - \epsilon$.* □

Theorem 1.2 now easily follows. By the above, for every $\lambda \in \mathbb{Z}$ and any $\epsilon > 0$, there exists $A \subset \mathbb{T}$ such that $\mu(A) > 1/2 - \epsilon$. Applying Lemma 3.2 to the set A we obtain $A' \subset \mathbb{Z}_p$ (for some sufficiently large prime p) such that $(A' + \lambda A') \cap [0, \delta p] = \emptyset$, and hence in particular

$$|A' + \lambda A'| \leqslant (1 - \delta)p,$$

as claimed.

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

4. A quantitative bound

Since we are considering a very specific family of maps, it is possible to make the construction in Rokhlin's lemma very explicit. Indeed, one can avoid appealing to measure theory or limiting arguments, which has the added advantage of providing an explicit relation between the constants involved. As a first instance we will assume that λ is a positive integer and later on we will show how one can remove this assumption.

Let $m \in \mathbb{N}$ and set $E_0 = [0, \lambda^{-m})$. Then for each $i \in \mathbb{N}$ define

$$E_i = \{x \in \mathbb{T} : \lambda^i x \in E_0 \text{ but } \lambda^j x \notin E_0 \text{ for all } j < i\}.$$

More combinatorially, we may think of E_i as the set of $x \in \mathbb{T}$ such that the first occurrence of m consecutive zeros in the λ-ary expansion of x appears at position i.

Clearly, the sets E_i are disjoint and satisfy the inclusion $\lambda E_{i+1} \subseteq E_i$. Furthermore, their union almost covers $\mathbb{T}$.

Lemma 4.1. *For any $n \in \mathbb{N}$,*

$$\mu\left(\bigcap_{i=0}^{n-1} \bar{E}_i\right) \leqslant (1-\lambda^{-m})^{\frac{n}{m}} \leqslant e^{-\lambda^{-m}\frac{n}{m}}.$$

Proof. We turn to the combinatorial interpretation of the sets E_i: to estimate $\mu\big(\bigcap_{i=0}^{n-1} \bar{E}_i\big)$, it will be sufficient to bound the number of sequences in $[\lambda]^{n+m} = \{0, 1, \ldots, \lambda-1\}^{n+m}$ without any m consecutive zeros. In other words we are trying to bound above the probability that a uniformly chosen sequence in $[\lambda]^{n+m}$ contains m consecutive zeros.

To do so, split the sequence into at least $k = \lfloor \frac{n+m}{m} \rfloor$ disjoint 'blocks' of size m. The probability that each block is not a sequence of m consecutive zeros is $(1-\lambda^{-m})$. Hence, by independence, the probability that none of the blocks is a sequence of m consecutive zeros is at most

$$(1-\lambda^{-m})^k \leqslant (1-\lambda^{-m})^{\frac{n}{m}},$$

as claimed. □

Now for any $t \geqslant m$, let

$$A_t = E_1 \sqcup E_3 \sqcup \cdots \sqcup E_{2t+1}.$$

We have that $\lambda A_t \subseteq E_0 \sqcup E_2 \sqcup \cdots \sqcup E_{2t}$, and thus $A_t \cap \lambda A_t = \emptyset$. Furthermore, it is easy to see that

$$\begin{aligned}\mu(A_t) &\geqslant \frac{1}{2}\left(\sum_1^{2t} \mu(E_i)\right) \geqslant \frac{1}{2}\big(1-(1-\lambda^{-m})^{\frac{2t+1}{m}} - \lambda^{-m}\big) \\ &= \frac{1}{2} - \frac{1}{2}(1-\lambda^{-m})^{\frac{2t+1}{m}} - \frac{1}{2}\lambda^{-m}.\end{aligned}$$

Note that our set A_t is already the union of disjoint intervals, and therefore we do not need to approximate it as such in order to transfer it to $\mathbb{Z}_p$. Furthermore, is easy to see that each E_i is the union of at most λ^i intervals and hence A_t is the union of at most λ^{4t} intervals. Now, if we shorten

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

each of the intervals in A_t by removing segments of length $\epsilon\lambda^{-4t-1}/4$, we obtain a set A_t' such that $\mu(A_t') \geqslant \mu(A_t) - \epsilon/2$ and $d(A_t', \lambda A_t') \geqslant \epsilon\lambda^{-4t}$.

In order to remove the assumption that λ is positive, note that it was only used when estimating $\mu\big(\bigcup_0^{n-1} E_i^{(\lambda)}\big)$. If $\lambda < 0$, then one can obtain an estimate for the measure of the union by simply noting that applying the transformation twice is the same as multiplication by $\lambda^2 > 0$. It is clear that

$$\bigcup_{i=0}^{2(n-1)} E_i^{(\lambda)} \supset \bigcup_{i=0}^{n-1} E_i^{(\lambda^2)},$$

and thus we are well placed to use the previous estimate to bound the right-hand side of the inclusion.

Finally, setting $m = C_1 \log \epsilon^{-1}$ and $t = C_2\epsilon^{-1}(\log \epsilon^{-1})^2$, where C_1 and C_2 are some sufficiently large constants, and transferring to $\mathbb{Z}_p$ in the usual manner, we obtain the following quantitative version of Theorem 1.2.

Theorem 4.2. *Given any $\epsilon > 0$ and $\lambda \neq 0 \in \mathbb{Z}$, for all sufficiently large primes p, there exists a set $A \subset \mathbb{Z}_p$ with $|A| \geqslant 1/2 - \epsilon$ such that $|A - \lambda A| \leqslant (1 - \epsilon^{C\epsilon^{-1}})p$ for some universal constant C.*

5. Comments and further questions

Theorem 1.2 and Theorem 1.4 show that one cannot expect to obtain sharp bounds independent of the size of A. This motivates the following definition.

Definition. Given a prime p, a sequence of integers $\bar{\lambda} = (\lambda_1, \dots, \lambda_k)$, and $\alpha \in [0, 1]$, set

$$\mathrm{ex}(\mathbb{Z}_p, \bar{\lambda}, \alpha) = \inf\left\{ \frac{|\lambda_1 A + \cdots + \lambda_k A|}{|A|} : A \subset \mathbb{Z}_p \text{ with } |A| \leqslant \alpha p \right\}.$$

Thus, when $\bar{\lambda} = (1, -2)$, for example, we have $\mathrm{ex}(\mathbb{Z}_p, \bar{\lambda}, \alpha) = 3$ if $\alpha > 0$ is sufficiently small, and $\mathrm{ex}(\mathbb{Z}_p, \bar{\lambda}, \alpha) \leqslant 1/\alpha$ for every $\alpha \leqslant 1/2$. (So, in particular, $\mathrm{ex}(\mathbb{Z}_p, \bar{\lambda}, \alpha) \to 2$ as $\alpha \to 1/2$ and $p \to \infty$.) This suggests that $\mathrm{ex}(\mathbb{Z}_p, \bar{\lambda}, \alpha)$ might exhibit some non-trivial behaviour between the two extremes. The main open question, of course, is to understand the asymptotic behaviour of $\mathrm{ex}(\mathbb{Z}_p, \bar{\lambda}, \alpha)$ for all values of $\alpha \in [0, 1]$.

Question 1. Determine the value of $\mathrm{ex}(\mathbb{Z}_p, \bar{\lambda}, \alpha)$ as $p \to \infty$ for all $\alpha \in [0, 1]$.

In this work, we have shown that Conjecture 1.3 holds in the case where $\bar{\lambda} = (1, \lambda)$. Nonetheless, the transference principle still applies to an arbitrary $\bar{\lambda} = (\lambda_1, \lambda_2)$ and hence it suffices to construct large sets $A \subset \mathbb{T}$ (of measure $1/2 - \epsilon$) for which $\lambda_1(A) \cap \lambda_2(A) = \emptyset$.

Question 2. Let $\lambda_1, \lambda_2 \in \mathbb{Z}$ and suppose $A \subseteq \mathbb{T}$ is a Lebesgue-measurable set such that $\lambda_1 A \cap \lambda_2 A = \emptyset$. How large can $\mu(A)$ be?

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press

However, when dealing with higher-order sumsets, for instance $\bar{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$, it looks likely that substantial new ideas will be required.

Acknowledgements

The work was inspired by a series of conversations with Pablo Candela and David Saxton at the University of Cambridge. The author is very grateful to Ben Green, who pointed out the fact that one could remove the measure theory to obtain a quantitative version of the theorem. Special thanks are due to Robert Morris for his many suggestions to improve the clarity and exposition of this paper and last but not least the referees for their helpful comments.

References

[1] Bukh, B. (2008) Sums of dilates. *Combin. Probab. Comput.* **17** 627–639.
[2] Cauchy, A. (1813) Recherches sur les nombres. *J. École Polytechnique* **9** 99–116.
[3] Cilleruelo, J., Hamidoune, Y. O. and Serra, O. (2009) On sums of dilates. *Combin. Probab. Comput.* **18** 871–880.
[4] Cilleruelo, J., Silva, M. and Vinuesa, C. (2010) A sumset problem. *J. Combin. Number Theory* **2**.
[5] Davenport, H. (1935) On the addition of residue classes. *J. London Math. Soc.* **10** 30–32.
[6] Freiman, G. (1973) *Foundations of a Structural Theory of Set Addition*, Vol. 37 of *Translations of Mathematical Monographs*, AMS.
[7] Green, B. and Ruzsa, I. (2006) Sets with small sumset and rectification. *Bull. London Math. Soc.* **38** 43–52.
[8] Hamidoune, Y. O. and Plagne, A. (2002) A generalization of Freiman's $3k-3$ theorem. *Acta Arithmetica* **103** 147–156.
[9] Hamidoune, Y. O. and Rué, J. (2011) A lower bound for the size of a Minkowski sum of dilates. *Combin. Probab. Comput.* **20** 249–256.
[10] Heinemann, S. M. and Schmitt, O. (2001) Rokhlin's lemma for non-invertible maps. *Dyn. Syst. Appl.* **10** 201–213.
[11] Nathanson, M. (2008) Inverse problems for linear forms over finite sets of integers. *J. Ramanujan Math. Soc.* **23** 151–165.
[12] Plagne, A. (2011) Sums of dilates in groups of prime order. *Combin. Probab. Comput.* **20** 867–873.
[13] Rokhlin, V. A. (1963) Generators in ergodic theory. *Vestnik Leningrad. Univ. Math.* **18** 26–32.
[14] Ruzsa, I. (1989) An application of graph theory to additive number theory. *Scientia Ser. A* **3** 97–109.
[15] Ruzsa, I. (1996) Sums of finite sets. In *Number Theory: New York Seminar*, Springer, pp. 281–293.
[16] Tao, T. and Vu, V. (2006) *Additive Combinatorics*, Cambridge University Press.
[17] Vosper, A. G. (1956) The critical pairs of subsets of a group of prime order. *J. London Math. Soc.* (2) **31** 200–205.

https://doi.org/10.1017/S0963548312000466 Published online by Cambridge University Press