LARGE SETS WITH SMALL DOUBLING MODULO $p$ ARE WELL COVERED BY AN ARITHMETIC PROGRESSION

ORIOL SERRA AND GILLES ZÉMOR

Abstract. We prove that there is $\epsilon > 0$ and $p_0 > 0$ such that for every prime $p > p_0$, every subset $S$ of $\mathbb{Z}/p\mathbb{Z}$ which satisfies $|2S| \leq (2 + \epsilon)|S|$ and $2(|2S|) - 2|S| + 3 \leq p$ is contained in an arithmetic progression of length $|2S| - |S| + 1$. This is the first result of this nature which places no unnecessary restrictions on the size of $S$.

1. Introduction

In 1959 Freiman [2] proved that if $S$ is a set of integers such that

$$|2S| \leq 3|S| - 4$$

then $S$ is contained in an arithmetic progression of length $|2S| - |S| + 1$.

This result is often known as Freiman’s $(3k - 4)$ Theorem. It has been conjectured that the same result also holds in the finite groups $\mathbb{Z}/p\mathbb{Z}$ of prime order. Working towards this conjecture, Freiman [2] proved (see Nathanson [13] for the following formulation of the result):

**Theorem 1** (Freiman [2]). Let $S \subset \mathbb{Z}/p\mathbb{Z}$ such that $3 \leq |S| \leq c_0 p$ and

$$|2S| \leq c_1 |S| - 3,$$

with $0 < c_0 \leq 1/12$, $c_1 > 2$ and $(2c_1 - 3)/3 < (1 - c_0 c_1)/c_1^{1/2}$. Then $S$ is contained in an arithmetic progression of length $|2S| - |S| + 1$.

The largest possible numerical value of $c_1$ given by this theorem is $c_1 \approx 2.45$, which falls somewhat short of the value predicted by the conjecture (namely 3). In addition, Theorem 1 only guarantees the result for sets $S$ that are small enough. For example, to guarantee $c_1 = 2.4$, the theorem needs the assumption $|S| \leq p/35$. This last assumption was improved to $|S| \leq p/10.7$ by Rødseth [14] but without improving the value of the constant $c_1$.

It follows from a recent result of Green and Rusza [4] on rectification of sets with small doubling in $\mathbb{Z}/p\mathbb{Z}$ that the value of $c_1$ can actually be pushed all the way to 3 while preserving the conclusion that $S$ is contained in a short arithmetic progression, but this comes at the expense of a stringent condition on the size of $S$: namely the the extra assumption $|S| < 10^{-180} p$. 
In the present paper, we shall work at the conjecture from a different direction. Rather than focusing on the best possible value for the constant $c_1$, we shall try to lift all restrictions on the size of $S$. First we need to formulate properly what should be the right version of Freiman’s $(3k - 4)$ theorem in $\mathbb{Z}/p\mathbb{Z}$.

For $-1 \leq m \leq |S| - 4$, we want the condition $|2S| = 2|S| + m$ to imply that $S$ is included in an arithmetic progression of length $|S| + m + 1$. One fact that has not been spelt out explicitly in the literature is that for such a result to hold, some lower bound on the size of the complement $\mathbb{Z}/p\mathbb{Z} \setminus 2S$ of $2S$ must be formulated. Indeed, if $p - |2S|$ is too small, the conclusion will not hold even if $m$ is small compared to $|S| - 4$. Consider in particular the following example. Let $S = \{0\} \cup \{m + 3, m + 4, \ldots, (p + 1)/2\}$. We have $|2S| = p - (m + 1) = 2|S| + m$, but it can be seen with a little thought that $S$ is not included in an arithmetic progression of length $|S| + m + 1$. For the desired result to hold, we must therefore add the condition $p - |2S| > m + 1$.

We conjecture that this extra condition is sufficient for a $\mathbb{Z}/p\mathbb{Z}$-version of Freiman’s $(3k - 4)$ theorem to hold. More precisely:

**Conjecture 2.** Let $S \subset \mathbb{Z}/p\mathbb{Z}$ and let $m = |2S| - 2|S|$. Suppose that $m$ satisfies:

$$-1 \leq m \leq \min\{|S| - 4, p - |2S| - 3\}.$$

Then $S$ is included in an arithmetic progression of length $|S| + m + 1$.

Note that $p - |2S| = p - 2|S| - m$ can not be equal to $m + 2$, otherwise $p$ would be an even number. Therefore condition (ii) of the conjecture is equivalent to $p - |2S| > m + 1$, as implied by the example above.

We remark that the cases $m = -1, 0, 1$ of this conjecture are known. They are implied by Vosper’s theorem [15] ($m = -1$), by a result of Hamidoune and Rødseth [9] ($m = 0$) and by a result of Hamidoune and the present authors [10] ($m = 1$). In the present paper we shall prove conjecture 2 for all values of $m$ up to $\epsilon|S|$, where $\epsilon$ is a fixed absolute constant. More precisely, our main result is:

**Theorem 3.** There exist positive numbers $p_0$ and $\epsilon$ such that, for all primes $p > p_0$, any subset $S$ of $\mathbb{Z}/p\mathbb{Z}$ such that

1. $|2S| < (2 + \epsilon)|S|$,
2. $m = |2S| - 2|S|$ satisfies $m \leq \min\{|S| - 4, p - |2S| - 3\}$,

is included in an arithmetic progression of length $|S| + m + 1$.

We shall prove this result with the numerical values $\epsilon = 10^{-4}$ and $p_0 = 2^{94}$. 
In the past, the dominant strategy, already present in Freiman’s original proof of Theorem 1, has been to rectify the set \( S \), i.e. find an argument that enables one to claim that the sum \( S + S \) must behave as in \( \mathbb{Z} \), and then apply Freiman’s \((3k - 4)\) theorem. Rectifying \( S \) directly however, becomes more and more difficult when the size of \( S \) grows, hence the different upper bounds on \( S \) that one regularly encounters in the literature. In our case, without any upper bound on \( S \), rectifying \( S \) by studying its structure directly is a difficult challenge. Our method will be indirect. Our strategy is to use an auxiliary set \( A \) that minimizes the difference \(|S + A| - |S|\) among all sets such that \(|A| \geq m + 3\). The set \( A \) is called an \((m + 3)\)-atom of \( S \) and using such sets to derive properties of \( S \) is an instance of the isoperimetric (or atomic) method in additive number theory which was introduced by Hamidoune and developed in [6, 7, 8, 10, 11]. The point of introducing the set \( A \) is that we shall manage to prove that it is both significantly smaller than \( S \) and also has a small sumset \( 2A \). This will enable us to show that first the sum \( A + A \), and then the sum \( S + A \), must behave as in \( \mathbb{Z} \). Finally we will use Lev and Smelianski’s distinct set version [12] of Freiman’s \((3k - 4)\) Theorem to conclude.

The paper is organised as follows. The next section will introduce \( k \)-atoms and their properties that are relevant to our purposes. In Section 3 we will show how our method works proving Theorem 3 in the relatively easy case when \( m \) is an arbitrary constant or a slowly growing function of \( p \) (i.e. \( \log p \)). In Section 4 we will prove Theorem 3 in full when \( m \) is a linear function of \(|S|\).

2. Atoms

Let \( S \) be a subset of \( \mathbb{Z}/p\mathbb{Z} \) such that \( 0 \in S \). For a positive integer \( k \), we shall say that \( S \) is \( k \)-separable if there exists \( X \subset \mathbb{Z}/p\mathbb{Z} \) such that \(|X| \geq k \) and \(|X + S| \leq p - k\).

Suppose that \( S \) is \( k \)-separable. The \( k \)-th isoperimetric number of \( S \) is then defined by

\[
\kappa_k(S) = \min\{|X + S| - |X|, \quad X \subset \mathbb{Z}/p\mathbb{Z}, \quad |X| \geq k \text{ and } |X + S| \leq p - k\}. \tag{1}
\]

For a \( k \)-separable set \( S \), a subset \( X \) achieving the above minimum is called a \( k \)-fragment of \( S \). A \( k \)-fragment with minimal cardinality is called a \( k \)-atom.

What makes \( k \)-atoms interesting objects is the following lemma:

**Lemma 4** (The intersection property [6]). Let \( S \) be a subset of \( \mathbb{Z}/p\mathbb{Z} \) such that \( 0 \in S \), and suppose \( S \) is \( k \)-separable. Let \( A \) be a \( k \)-atom of \( S \). Let \( F \) be a \( k \)-fragment of \( S \) such that \( A \not\subseteq F \). Then \(|A \cap F| \leq k - 1\).

The following Lemma is proved in [11]:
Lemma 5. Let $S \subset \mathbb{Z}/p\mathbb{Z}$ with $|S| \geq 3$ and $0 \in S$. Suppose $S$ is 2-separable and $\kappa_2(S) \leq |S| + m$. Let $A$ be a 2–atom of $S$. Then $|A| \leq m + 3$.

Lemma 5 implies the following upper bound on the size of atoms.

Lemma 6. Let $k \geq 3$ and let $A$ be a $k$–atom of a $k$–separable set $S \subset \mathbb{Z}/p\mathbb{Z}$ with $0 \in S$, $|S| \geq 2$ and $\kappa_k(S) \leq |S| + m$. Then $|A| \leq 2m + k + 2$.

Proof. The set $A$ is clearly 2–separable. Let $B$ be a 2–atom of $A$ with $0 \in B$, so that $|B + A| \leq |B| + |A| + m$. Let $b \in B$, $b \neq 0$. By Lemma 5 we have $|B| \leq m + 3$. Therefore,

$$|A \cup (b + A)| = |\{0, b\} + A| \leq |B + A| \leq |A| + 2m + 3.$$  \hspace{1cm} (2)

But $b + A$ is also a $k$–atom of $S$. By the intersection property, it follows that $|A \cap (b + A)| \leq k - 1$. Hence $2|A| - (k - 1) \leq |A \cup (b + A)|$ which together with (2) gives the result. \hspace{1cm} $\square$

From now on $S$ will refer to a subset of $\mathbb{Z}/p\mathbb{Z}$ satisfying conditions (i) and (ii) of Theorem 3 for a fixed $\epsilon > 0$ to be determined later, and $m$ always denotes the integer $m = |2S| - |S|$. Without loss of generality we will also assume $0 \in S$.

Note that condition (ii) implies that $S$ is $(m + 3)$–separable so that $(m + 3)$-atoms of $S$ exist. Note that by the definition of an atom, if $X$ is an atom of $S$ then so is $x + X$ for any $x \in \mathbb{Z}/p\mathbb{Z}$. Therefore there are atoms containing the zero element.

In the sequel $A$ will denote an $(m + 3)$–atom of $S$ with $0 \in A$. We will regularly call upon the following two inequalities:

$$|S + A| \leq |S| + |A| + m$$ \hspace{1cm} (3)

which follows from the definition of an atom, and

$$|A| \leq 3m + 5.$$ \hspace{1cm} (4)

which follows from Lemma 6 with $k = m + 3$.

The reader should also bear in mind that for all practical purposes, inequality (4) means that we will only be dealing with cases when $|A|$ is significantly smaller than $|S|$. Indeed, we shall prove Theorem 3 for a small value of $\epsilon$, namely $\epsilon = 10^{-4}$, so that $3m$ is very much smaller than $|S|$. We can also freely assume that $|S| \geq p/35$, since otherwise Freiman’s Theorem 1 gives the result with $\epsilon = 0.4$. The prime $p$ will also be assumed to be larger than some fixed value $p_0$ to be determined later.
3. The case $m \leq \log p$

In this section we will deal with the case when $m$ is a very small quantity, i.e. smaller than a logarithmic function of $p$. This will allow us to introduce, without technical difficulties to hinder us, the general idea of the method which is to first show that $A$ must be contained in a short arithmetic progression and then to transfer the structure of $A$ to the larger set $S$. It will also serve the additional purpose of allowing us to suppose $m \geq 6$ when we switch to the looser condition $m \leq \epsilon |S|$.

We start by stating some results that we shall call upon. The first is a generalization of Freiman’s theorem in $\mathbb{Z}$ to sums of different sets and is proved by Lev and Smelianski in [12], we give it here somewhat reworded (see also [17, Th. 5.12]).

**Theorem 7** (Lev and Smelianski [12]). Let $X$ and $Y$ be two nonempty finite sets of integers with

$$|X + Y| = |X| + |Y| + \mu.$$ 

Assume that $\mu \leq \min\{|X|, |Y|\} - 3$ and that one of the two sets $X, Y$ has size at least $\mu + 4$. Then $X$ is contained in an arithmetic progression of length $|X| + \mu + 1$ and $Y$ is contained in an arithmetic progression of length $|Y| + \mu + 1$.

The second result we shall use is due to Bilu, Lev and Ruzsa [1, Theorem 3.1] and gives a bound on the length of small sets in $\mathbb{Z}/p\mathbb{Z}$. By the length $\ell(X)$ of a set $X \subset \mathbb{Z}/p\mathbb{Z}$ we mean the length (cardinality) of the shortest arithmetic progression which contains $X$.

**Theorem 8** (Bilu, Lev, Ruzsa [1]). Let $X \subset \mathbb{Z}/p\mathbb{Z}$ with $|X| \leq \log_4 p$. Then $\ell(X) < p/2$.

Theorem 8 will be used to show that, when $m$ is small enough, then the atom $A$ is contained in a short arithmetic progression.

**Lemma 9.** Suppose that $6m + 11 \leq \log_4 p$. Then $A$ is contained in an arithmetic progression of length $2(|A| - 1)$.

**Proof.** Since we assume $|S| \geq p/35$, it follows from (3) and (4) that $A$ is an $(m + 4)$--separable set. Let therefore $B$ be an $(m + 4)$--atom of $A$ containing 0, so that $|B + A| \leq |B| + |A| + m$. By Lemma 6 we have $|B| \leq 3m + 6$ so that $|A \cup B| \leq 6m + 11$. By the present lemma’s hypothesis, it follows from Theorem 8 that $A \cup B$ is contained in an arithmetic progression of length less than $p/2$. The sum $A + B$ can therefore be considered as a sum of integers, so that Theorem 7 applies and $A$ is contained in an arithmetic progression of length $|A| + m + 1 \leq 2|A| - 2$. $\square$
We now proceed to deduce from Lemma 9 the structure of $S$. It will be convenient to introduce the following notation.

Recall that we denote by $\ell(X)$ the length of the smallest arithmetic progression containing $X$. By $\ell_X(Y)$ we shall denote the length of a smallest arithmetic progression of difference $x$ containing $Y$, where $x$ is the difference of a shortest arithmetic progression containing $X$.

The point of the above definition is that if we have $\ell_A(S) + \ell(A) \leq p$ then the sum $S + A$ can be considered as a sum in $\mathbb{Z}$, so that (3) and Theorem 7 applied to $S$ and $A$ imply Theorem 3. We summarize this point in the next Lemma for future reference.

**Lemma 10.** If we can assume $\ell_A(S) + \ell(A) \leq p$ then Theorem 3 holds.

Whenever we will wish transfer the structure of $A$ to $S$ we will assume that $\ell_A(S) + \ell(A) > p$ and look for a contradiction. We can think of this hypothesis as $S$ having no ‘holes’ of length $\ell(A)$. In the present case of very small $m$, the desired result on $S$ follows with very little effort.

**Lemma 11.** Suppose that $6m + 11 \leq \log_4 p$. Then $S$ is contained in an arithmetic progression of length $|S| + m + 1$.

**Proof.** By Lemma 9, $A$ is contained in an arithmetic progression of difference $r$, that we can assume to equal $r = 1$, and of length $2(|A| − 1)$. In particular $A$ has two consecutive elements. Without loss of generality we may replace $A$ by a translate of $A$ and assume that $\{0, 1\} \subset A$. Let $S = S_1 \cup \cdots \cup S_k$ be the decomposition of $S$ into maximal arithmetic progressions of difference one, so that

$$|S + A| \geq |S| + k.$$  

Because of (3) we have $k \leq |A| + m$. By Lemma 10 we can assume every maximal arithmetic progression in the complement of $S$ to have length at most $\ell(A)$. Therefore,

$$\ell_A(S) + \ell(A) \leq |S| + k\ell(A) \leq |S| + (|A| + m)2(|A| − 1).$$

Now by (4) we get

$$\ell_A(S) + \ell(A) \leq |S| + (4m + 5)(6m + 8) < |S| + (\log_4 p)^2 < \frac{p}{2} + (\log_4 p)^2$$

since $|S| < p/2$. We have $\log_4^2 p < p/2$ for all $p$ therefore we get $\ell_A(S) + \ell(A) < p$, a contradiction. \hfill $\Box$

## 4. The general case

### 4.1. Overview.

When $m$ grows we encounter two difficulties. First, Theorem 8 will not apply anymore to any set containing $A$, and we need an alternative method to argue that $A$ is contained
in a short arithmetic progression. Second, even if we do manage to prove that \( A \) is contained in a short arithmetic progression, we will not be able to deduce the structure of \( S \) from (3) by the simple technique of the preceding section.

We will now use an extra tool, namely the Plünecke-Ruzsa estimates for sumsets; see e.g. \([15, 13]\).

**Theorem 12** (Plünecke-Ruzsa \([15]\)). Let \( S \) and \( T \) be finite subsets of an abelian group with \(|S + T| \leq c|S|\). There is a nonempty subset \( S' \subset S \) such that

\[
|S' + jT| \leq c^j|S'|
\]

The Plünecke-Ruzsa inequalities applied to \( S \) and \( A \) will give us that there exists a positive \( \delta \) such that either \( A \) is contained in a progression of length \((2 - \delta)(|A| - 1)\) or \( 2A \) is contained in an arithmetic progression of length \((2 - \delta)(|2A| - 1)\) (Lemma \([15]\)). We will then proceed to transfer the structure of \( A \) or \( 2A \) to \( S \).

Again we shall use Lemma \([10]\) to assume that \( S \) does not contain a “gap” of length \( \ell(A) \) or \( \ell(2A) \). We define the density of a set \( X \subset \mathbb{Z}/p\mathbb{Z} \) as \( \rho(X) = (|X| - 1)/\ell(X) \). If \( \ell(A) \leq (2 - \delta)(|A| - 1) \) we will argue that the sum \( S + A \) must have a density at least that of \( A \) and get a contradiction with the upper bound on \(|S + A|\). The details will be given in Subsection \([4.3]\).

We will not be quite done however, because we can not guarantee that \( \ell(A) \leq (2 - \delta)(|A| - 1) \) holds. In that case we have to fall back on the condition \( \ell(2A) \leq (2 - \delta)(|2A| - 1) \), meaning that it is the set \( 2A \), rather than \( A \), that has large enough density. In this case we have to work a little harder. We proceed in two steps: we first apply the Plünecke-Ruzsa inequalities again to show that there exists a large subset \( T \) of \( S \) such that \(|T + 2A|\) is small. We then apply the density argument to show that \( T \) must be contained in an arithmetic progression with few missing elements. We then focus on the remaining elements of \( S \), i.e. the set \( S \setminus T \). We will again argue that if this set has a gap of length \( \ell(A) \) the desired result holds and otherwise the density argument will give us that \( S + A \) is too large. This analysis is detailed in Subsection \([4.4]\) and will conclude our proof of Theorem \([3]\).

4.2. **Structure of \( A \).**

**Lemma 13.** Suppose \( 6 \leq m \leq \epsilon|S| \) with \( \epsilon \leq 10^{-4} \). Then for any positive integer \( k \leq 32 \) we have

\[
|kA| \leq k(|A| + m) \left( 1 + \frac{5k\epsilon}{2} \right) + 1.
\]

**Proof.** Rewrite (3) as

\[
|S + A| \leq |S| + |A| + m = c|S|,
\]
with \( c = 1 + \frac{|A| + m}{|S|} \). By Theorem 12 (Plünnecke–Ruzsa), for each \( k \) there is a subset \( S' = S'(k) \) such that

\[
|S' + kA| \leq c^k |S'|
\]  
(5)

Apply (4) and \( m \geq 6 \) to get \(|A| \leq 3m + 5 \leq 4m \). Since \( m \leq \varepsilon|S| \) we obtain for the constant \( c \) just defined \( c \leq 1 + 5\varepsilon \). We clearly have

\[
c^k |S'| \leq c^k |S| \leq (1 + 5\varepsilon)^k |S| < 2|S| < p
\]

for \( k \leq 32 \). Now apply the Cauchy–Davenport Theorem to \( S' + kA \) in (5) to obtain

\[
|kA| \leq (c^k - 1)|S'| + 1 \leq (c^k - 1)|S| + 1.
\]  
(6)

Numerical computations give that

\[
(1 + x)^k \leq 1 + kx + \frac{k^2}{2}x^2
\]

for any positive real number \( x \leq 5.10^{-4} \) and for \( k \leq 32 \). Hence, since \( c = 1 + (|A| + m)/|S| \leq 1 + 5\varepsilon \), we can write, for \( k \leq 32 \),

\[
c^k = \left(1 + \frac{|A| + m}{|S|}\right)^k \leq 1 + k\frac{|A| + m}{|S|} + k^2 \left(\frac{|A| + m}{|S|}\right)^2.
\]

Applied to (6) we get

\[
|kA| \leq k(|A| + m) + \frac{k^2}{2} \left(\frac{|A| + m}{|S|}\right)^2 + 1
\]

\[
\leq k(|A| + m) \left(1 + \frac{k}{2} \frac{|A| + m}{|S|}\right) + 1
\]

\[
\leq k(|A| + m) \left(1 + \frac{5k\varepsilon}{2}\right) + 1,
\]

as claimed. \( \square \)

**Lemma 14.** If \( 6 \leq m \leq \varepsilon|S| \) with \( \varepsilon \leq 10^{-4} \), then \( A \) and \( 2A \) are contained in an arithmetic progression of length less than \( p/2 \).

**Proof.** Put \( k = 2^j \) and \( c_1 = 2.44 \). Suppose that \(|2^jA| \geq c_1 |2^{j-1}A| - 3 \) for each \( 1 \leq j \leq 5 \). Then,

\[
|32A| \geq c_1^5 |A| - 3(c_1^5 - 1)/(c_1 - 1) \geq 86|A| - 179 \geq 65|A| + 10,
\]

where in the last inequality we have used \(|A| \geq m + 3 \geq 9 \). On the other hand, by Lemma 13 we have

\[
|kA| \leq k(|A| + m) \left(1 + \frac{5k\varepsilon}{2}\right) + 1 \leq 2k(1 + \frac{5k\varepsilon}{2})|A|,
\]  
(7)

which, for \( k = 32 \), gives \(|32A| \leq 64(1 + 80\varepsilon)|A| \leq 65|A| \), a contradiction.
Hence $|2^j A| \leq c_1 |2^{j-1} A| - 3$ for some $1 \leq j \leq 5$. Since 
$$|2^{j-1} A| \leq |16 A| \leq 32(1 + 40\epsilon)|A| \leq 64(1 + 40\epsilon)c_p < 8 \cdot 10^{-3}p,$$
where again we have used inequality (7) for $k = 16$ and $|A| \leq 4m \leq 4\epsilon|S| \leq 2ep$. It follows from Freiman’s Theorem (with $c_0 = 8 \cdot 10^{-3}$ and $c_1 = 2.44$) that $A \subset 2^{j-1} A$ is contained in an arithmetic progression of length at most 
$$|2^j A| - |2^{j-1} A| + 1 < 1.44|2^{j-1} A| \leq (1.44)8 \cdot 10^{-3}p.$$
In particular, $A$ and $2A$ are included in arithmetic progressions of lengths less than $p/2$. □

Now that we know that $A$ and $2A$ are contained in an arithmetic progression of length smaller than $p/2$, we can apply to them the Freiman’s $(3k - 4)$ Theorem to get the following result.

**Lemma 15.** Suppose $6 \leq m \leq \epsilon|S|$ with $\epsilon \leq 10^{-4}$, and let $0 < \delta \leq 10^{-1}$. If $A$ is not contained in an arithmetic progression of length $(2 - \delta)(|A| - 1)$ then $2A$ is contained in an arithmetic progression of length $(2 - \delta)(|2A| - 1)$.

**Proof.** Suppose first that $|2A| \geq (3 - \delta)(|A| - 1)$ and $|4A| \geq (3 - \delta)(|2A| - 1)$. Then 
$$|4A| \geq (3 - \delta)^2|A| - (3 - \delta)^2 - (3 - \delta) \geq (3 - \delta)^2|A| - 12. \quad (8)$$
On the other hand, Lemma 13 for $k = 4$ and $\epsilon = 10^{-4}$ gives $|4A| \leq 4(1 + 10\epsilon)(|A| + m) + 1$. By using (8) and $m \leq |A| - 3$ we get 
$$(3 - \delta)^2|A| - 12 \leq 8(1 + 10\epsilon)|A| - 12(1 + 10\epsilon) + 1.$$ 
Since $m \geq 6$, we have $|A| \geq m + 3 \geq 9$. Therefore we obtain 
$$(3 - \delta)^2|A| < \left(8(1 + 10\epsilon) + \frac{1}{9}\right)|A|,$$
a contradiction for $\delta \leq 0.1$.

Hence,

(a) either $|2A| < (3 - \delta)(|A| - 1) < 3|A| - 3$, but since $\ell(A) < p/2$ by Lemma 14 Freiman’s $(3k - 4)$ Theorem applies and $A$ is contained in an arithmetic progression of length $|2A| - (|A| - 1) \leq (2 - \delta)(|A| - 1)$.

(b) Or $|4A| < (3 - \delta)(|2A| - 1) < 3|2A| - 3$, but using Lemma 14 again, $(3k - 4)$–Freiman’s Theorem implies that $2A$ is contained in an arithmetic progression of length $(2 - \delta)(|2A| - 1)$.

□
4.3. **Structure of $S$ when $\ell(A)$ is small.** For a subset $B \subset \mathbb{Z}/p\mathbb{Z}$ define the *density* of $B$ by

$$\rho_B = \frac{|B| - 1}{\ell(B)}.$$  

The next lemma gives a lower bound for the cardinality of a sum set of two subsets $B, C \subset \mathbb{Z}/p\mathbb{Z}$ when $\ell(B) + \ell(C) > p$ in terms of their densities.

**Lemma 16.** Let $0 \in C \subset \mathbb{Z}/p\mathbb{Z}$ with $C \subset [0, \ell(C))$ and $\ell(C) < p/2$. Let $I_1, \ldots, I_{2t}$ be the sequence of intervals defined by $I_i = [(i - 1)c, ic)$, where $c = \ell(C)$ and $t < p/2c$. Let $B \subset \mathbb{Z}/p\mathbb{Z}$ such that for every $i = 1, \ldots, 2t$, we have $I_i \cap B \neq \emptyset$. Then,

$$|B + C| \geq |B \cup [(B + C) \cap I]| \geq |B| + (t - 1)\ell(C) \left(\rho_C - \frac{|B \cap I|}{(2t - 1)c}\right),$$

where $I = I_1 \cup \ldots \cup I_{2t}$.

**Proof.** Let $B' = B \cap I$. Let $B_0^i = B' \cap I_{2t-1}$ and $B_1^i = B' \cap I_{2i}$, and define $B_0^t = \bigcup_{i=1}^t B_0^i$, $B_1^t = \bigcup_{i=1}^t B_1^i$ so that $B' = B_0^t \cup B_1^t$. Note that, since $C \subset [0, c)$,

$$(B_0^i + C) \cap (B_0^j + C) = \emptyset$$

for $i \neq j$ and that $B_0^t + C \subset I_{2t-1} \cup I_{2t}$. Therefore $B_0^t + C$ can be written as the following union of disjoint sets.

$$B_0^t + C = \bigcup_{i=1}^t (B_0^i + C) \subset I_1 \cup \ldots \cup I_{2t}.$$  

Hence, since every set $B_0^i$ is nonempty, the Cauchy-Davenport Theorem implies

$$|B_0^t + C| \geq |B_0^t| + t(|C| - 1).$$  

(9)

In a similar manner we have

$$(B_1^1 + C) \cap I = \bigcup_{i=1}^{t-1} (B_1^i + C) \cup (B_1^{2t} + C) \cap I$$

$$\supset \bigcup_{i=1}^{t-1} (B_1^i + C) \cup B_1^{2t}$$

so that, applying the Cauchy-Davenport Theorem for $i = 1 \ldots t - 1$, we get

$$|(B_1^1 + C) \cap I| \geq |B_1^1| + (t - 1)(|C| - 1).$$  

(10)
Now we have $|B + C| \geq |B \setminus B'| + |(B'_0 + C) \cap I|$ and likewise $|B + C| \geq |B \setminus B'| + |(B'_1 + C) \cap I|$, hence, applying (9) and (10),

$$|B + C| \geq |B \setminus B'| + \frac{1}{2} (|(B'_0 + C) \cap I| + |(B'_1 + C) \cap I|)$$

$$\geq |B| - |B'|/2 + (t - \frac{1}{2})(|C| - 1)$$

$$\geq |B| + (t - \frac{1}{2})\epsilon (\rho C - \frac{|B'|}{(2t - 1)\epsilon})$$

which proves the result. $\square$

Lemma 16 allows us to conclude the proof when the $(m + 3)$-atom $A$ is contained in a short arithmetic progression.

**Lemma 17.** Suppose $6 \leq m \leq \epsilon |S|$ with $\epsilon \leq 10^{-4}$. Suppose furthermore that $\ell(A) \leq (2 - \delta)(|A| - 1)$. Then $\ell(S) \leq |S| + m + 1$.

**Proof.** Set $a = \ell(A)$. Write $p = 2ta + r$, $0 < r < 2a$ and let $I_1, \ldots, I_i, \ldots, I_{2t}$ be the partition of $[0, 2ta]$ into the intervals $I_i = [(i - 1)a, ia)$ and $I = \cup_{i=1}^{2t} I_i$. Let $S' = S \cap I$.

Suppose that $\ell_A(S) + \ell(A) > p$. Then we have $I_i \cap S' \neq \emptyset$ for each $i = 1, \ldots, 2t$. By Lemma 16 with $B = S$ and $C = A$,

$$|S + A| \geq |S| + (t - \frac{1}{2})a (\rho A - \frac{|S'|}{(2t - 1)a}).$$

(11)

Now we have $(2t - 1)a > p - 3a$ by definition of $t$. Since $|A| \leq 3m + 5$ we have $a = \ell(A) \leq 2(|A| - 1) \leq 6m + 8$, and since we have supposed $m \geq 6$, we get $a \leq 8m$. We therefore have

$$2t - 1 > 2 - \frac{3a}{p} \geq 2 - \frac{3a}{|S|} \geq 2 - \frac{3a}{p/2} \geq 2 - \frac{3a}{p/2}.$$  

(12)

By the hypothesis of the Lemma we have $\rho A \geq 1/(2 - \delta)$. Together with (12) we get, writing $|S'| \leq |S| < p/2$,

$$\rho A - \frac{|S'|}{(2t - 1)a} > \frac{1}{2 - \delta} - \frac{1}{2 - 2\epsilon}.$$  

Finally, applying again (12), inequality (11) becomes

$$|S + A| > |S| + \frac{p}{2} (1 - 12\epsilon) \left(\frac{1}{2 - \delta} - \frac{1}{2 - 2\epsilon}\right).$$

(13)

Now recall that by definition of $A$ we have $|A| \geq m + 3$. We will therefore get that (13) contradicts (3) whenever the righthand side of (13) is greater than $|S| + 2|A|$. Since $|A| \leq 3m + 5 \leq 4m \leq 2\epsilon p$, a contradiction is obtained whenever

$$\frac{1}{2} (1 - 12\epsilon) \left(\frac{1}{2 - \delta} - \frac{1}{2 - 2\epsilon}\right) \geq 4\epsilon.$$  

(14)
For $\epsilon \leq 10^{-4}$ the inequality (14) is verified for every $\delta > 5 \cdot 10^{-3}$. Since Lemma 15 allows us to choose $\delta$ up to the value $10^{-1}$, the hypothesis $\ell_A(S) + \ell(A) > p$ can not hold, so that the result follows from Lemma 10.

4.4. Structure of $S$ when $\ell(2A)$ is small. To conclude the proof of Theorem 3 it remains to consider the case where $\ell(A) > (2 - \delta)(|A| - 1)$. We break up the proof into several lemmas.

**Lemma 18.** Suppose $6 \leq m \leq \epsilon|S|$ with $\epsilon \leq 10^{-4}$. Suppose furthermore that $\ell(A) > (2 - \delta)(|A| - 1)$. Then

(i) $|2A| \geq (3 - \delta)(|A| - 1)$.

(ii) $\ell(A) \leq (1 - \delta/2)|2A|$.

**Proof.** By point (a) of the final argument in the proof of Lemma 15 we know that we can not have $|2A| < (3 - \delta)(|A| - 1)$. This proves (i).

Since $A$ is contained in an arithmetic progression of length less than $p/2$ (Lemma 14) we have $\ell(A) \leq (\ell(2A) + 1)/2$. Now Lemma 15 implies $\ell(2A) \leq (2 - \delta)(|2A| - 1)$, hence $(\ell(2A) + 1)/2 \leq (1 - \delta/2)|2A|$. This proves (ii).

Next we apply the Plünnecke-Ruzsa inequalities to exhibit a subset $T$ of $S$ that sums to a small sumset with $2A$. We then show that this set $T$ must be contained in an arithmetic progression with few missing elements.

**Lemma 19.** Suppose $6 \leq m \leq \epsilon|S|$ with $\epsilon \leq 10^{-4}$. Suppose furthermore that $\ell(A) > (2 - \delta)(|A| - 1)$. Then there exists $T \subset S$ such that, denoting $\lambda = |T|/|S|$,

$$|2A| \leq \lambda(4 + 10\epsilon)(|A| - 1),$$

$$\ell(T) \leq |T| + 2\ell(A).$$

**Proof.** By Theorem 12 and (3), there is $T \subset S$ such that

$$|T + 2A| \leq (1 + \frac{|A| + m}{|S|})^2|T| \leq |T| + 2(|A| + m)|\frac{T}{|S|} + \frac{(|A| + m)^2}{|S|}|T|.$$ 

Writing $|A| + m \leq 3m + 5 + m \leq 5m \leq 5\epsilon|S|$ and $\lambda = |T|/|S|$ we get

$$|T + 2A| \leq |T| + \lambda(|A| + m)(2 + 5\epsilon) < p.$$ (17)

Now apply the Cauchy-Davenport Theorem $|T + 2A| \geq |T| + |2A| - 1$ in (17) to get, since $|A| \geq m + 3$,

$$|2A| - 1 \leq \lambda(2|A| - 3)(2 + 5\epsilon),$$

$$|2A| \leq 2\lambda(2 + 5\epsilon)(|A| - 1) - \lambda(2 + 5\epsilon) + 1.$$ (18)
Notice that if $\lambda(2 + 5\epsilon) < 1$ then (15) gives $|2A| < 2(|A| - 1) + 1$ which contradicts the Cauchy-Davenport Theorem. Therefore we have $1 - \lambda(2 + 5\epsilon) \leq 0$ and (18) yields (15).

In the remaining part we prove (16). Recall that the hypothesis of the present lemma together with Lemma 15 imply

$$\ell(2A) \leq (2 - \delta)(|2A| - 1).$$

Suppose first that

$$\ell(2A)(T) + \ell(2A) > p.$$  

Set $a_2 = \ell(2A)$ and $p = 2ta_2 + r$ with $0 < r < 2a_2$. Let $I = I_1 \cup \cdots \cup I_2t$ with $I_i = [(i - 1)a_2, ia_2)$. By (20) we have $T \cap I_i \neq \emptyset$ for each $i = 1, \ldots, 2t$. By Lemma 16 with $B = T$ and $C = 2A$,

$$|T + 2A| \geq |T| + \left( t - \frac{1}{2} a_2 \right) \left( \rho(2A) - \frac{|T'|}{(2t - 1)a_2} \right)$$

where $T' = T \cap I$. By (19) we have $a_2 \leq 2|2A|$, so that by using (15) and $\lambda \leq 1$ we obtain the following rough upper bound

$$a_2 \leq (8 + 20\epsilon)|A| \leq 9(3m + 5) \leq 36m$$

where we have used $\epsilon \leq 1/20$.

As in the proof of Lemma 17, we have, by definition of $t$,

$$(2t - 1)a_2 \geq p - 3a_2 \geq p - 108m \geq p(1 - 54\epsilon)$$

so that, writing $|T'| \leq |T| \leq |S| \leq p/2$, and applying (19) we have

$$\rho(2A) - \frac{|T'|}{(2t - 1)a_2} \geq \frac{1}{2 - \delta} - \frac{1}{2 - 108\epsilon}.$$ 

Applying again (22), inequality (21) becomes

$$|T + 2A| \geq |T| + \frac{p}{2}(1 - 54\epsilon) \left( \frac{1}{2 - \delta} - \frac{1}{2 - 108\epsilon} \right).$$

On the other hand, (17) implies

$$|T + 2A| \leq |T| + 10m + 25\epsilon m \leq |T| + p(5\epsilon + 25\epsilon^2/2)$$

which together with (23) gives

$$5\epsilon + 25\epsilon^2/2 \geq \frac{1}{2}(1 - 54\epsilon) \left( \frac{1}{2 - \delta} - \frac{1}{2 - 108\epsilon} \right).$$

For $\epsilon = 10^{-4}$ the inequality (24) fails to hold for each $\delta \geq 2 \cdot 10^{-2}$. Since (19) holds for every $\delta \leq 10^{-1}$, the hypothesis (20) can not hold, so that the sumset $T + 2A$ behaves like a sum of integers. Let us write

$$|T + 2A| = |T| + |2A| + \mu$$
and check that the conditions of Theorem 7 hold. By Lemma 18 (i) we have

\[ |2A| \geq (3 - \delta)(|A| - 1) \geq (2 + 5\varepsilon)|A| + (1 - \delta - 5\varepsilon)|A| - 3 \geq (2 + 5\varepsilon)|A| + \frac{3}{2} \]

since \( m \geq 6 \) and \( |A| \geq m + 3 \geq 9 \). Therefore

\[ 2|2A| \geq 2(2 + 5\varepsilon)|A| + 3 \geq (2 + 5\varepsilon)(|A| + m) + 3, \]

which, since \( \mu \leq (|A| + m)(2 + 5\varepsilon) - |2A| \) by (17), leads to

\[ |2A| \geq \mu + 3. \quad (25) \]

Now by definition of \( \lambda \) we have \( |T| = \lambda |S| \) and we also have \( |S| \geq 11\varepsilon|S| \), so that

\[ |T| \geq \lambda 11\varepsilon|S| \geq \lambda 11m \geq \lambda (2 + 5\varepsilon)5m \geq \lambda (2 + 5\varepsilon)(|A| + m) \]

and, since \( \mu \leq \lambda (|A| + m)(2 + 5\varepsilon) - |2A| \) by (17), we obtain

\[ |T| \geq \mu + |2A| \geq \mu + 4. \quad (26) \]

Inequalities (25) and (26) mean that Theorem 7 holds and we have:

\[ \ell(T) \leq |T| + \mu + 1 \leq |T| + |2A| \leq |T| + \ell(2A) \leq |T| + 2\ell(A). \]

This proves (16) and concludes the lemma. \( \square \)

**Lemma 20.** Suppose \( 6 \leq m \leq \varepsilon|S| \) with \( \varepsilon \leq 10^{-4} \). Suppose furthermore that \( \ell(A) > (2 - \delta)(|A| - 1) \). Then \( \ell(S) \leq |S| + m + 1 \).

**Proof.** Let \( T \) be the set guaranteed by Lemma 19. Let \( \mathbf{T} = S \setminus T \), which belongs to an interval of length \( p - \ell(T) \). Set \( a = \ell(A) \). Let us apply again Lemma 16, this time with \( B = S \), \( C = A \), and \( t \) defined so as to have \( p - \ell(T) = 2ta + r \), \( 0 \leq r < 2a \). As before, set \( I = I_1 \cup \cdots \cup I_{2t} \) with \( I_i = [(i - 1)a, ia) \). Note that \( T \cap I = \emptyset \), so that \( \mathbf{T} \cap I = S \cap I \). Let us first suppose

\[ \ell_A(S) + \ell(A) > p \quad (27) \]

which implies \( \mathbf{T} \cap I_i \neq \emptyset \) for every \( i = 1, \ldots, 2t \), so that by Lemma 16 and denoting \( \mathbf{T}' = \mathbf{T} \cap I = S \cap I \),

\[ |S + A| \geq |S \cup (S + A) \cap I| \geq |S| + (t - \frac{1}{2})a \left( \rho A - \frac{\mathbf{T}'}{(2t - 1)a} \right). \quad (28) \]
By definition of $t$ and by (16) we have

$$(2t - 1)a > p - \ell(T) - 3a \geq p - |T| - 5a. \quad (29)$$

Now Lemma 18 (ii) and (15) give the following upper bound on $a$

$$a \leq |2A| \leq \lambda(4 + 10\epsilon)|A| \leq \lambda(4 + 10\epsilon)4m \leq \lambda(4 + 10\epsilon)2\epsilon p$$

so that we can write $-5a \geq -\lambda f(\epsilon)p$ with $f(\epsilon) = 10(4 + 10\epsilon)\epsilon$. Writing $|T| = \lambda|S| < \lambda p/2$, (29) becomes

$$(2t - 1)a > p(1 - \lambda(\frac{1}{2} + f(\epsilon))) \quad (30)$$

Next we write $|T'| \leq |T| = |S| - |T| = (1 - \lambda)|S|$, so that $|S| \leq p/2$ gives

$$|T'| \leq \frac{p}{2}(1 - \lambda) \quad (31)$$

Finally we bound $\rho A$ from below. Apply again Lemma 18 (ii) and (15) to get

$$\ell(A) \leq (1 - \delta/2)|2A| \leq (1 - \delta/2)\lambda(4 + 10\epsilon)(|A| - 1),$$

so that we have

$$\rho A \geq \frac{1}{\lambda(1 - \delta/2)(4 + 10\epsilon)} \quad (32)$$

Applying (30), (31) and (32) to (28) now gives

$$|S + A| > |S| + \frac{P}{2}\left[\frac{1 - \lambda(\frac{1}{2} + f(\epsilon))}{\lambda(1 - \delta/2)(4 + 10\epsilon)} - \frac{1}{2}(1 - \lambda)\right].$$

Together with (3), writing $|A| \leq 4m$ and $m \leq \epsilon p/2$, we obtain

$$\frac{1 - \lambda(\frac{1}{2} + f(\epsilon))}{\lambda(1 - \delta/2)(4 + 10\epsilon)} - \frac{1}{2}(1 - \lambda) - 5\epsilon < 0. \quad (33)$$

Now there exists $\epsilon_3 > 5.8 \times 10^{-3} > 0$ such that for every $\epsilon \leq \epsilon_3$, the lefthandside of (33) is strictly positive for every value of $\lambda \in [0, 1]$. In that case (33) can not hold and we obtain a contradiction with the hypothesis (27). Therefore Theorem 7 implies the result. \[\square\]

**Numerical values:** As it has been shown in the proofs Theorem 3 holds with $\epsilon = 10^{-4}$. As for the value of $p_0$, we use $m \geq 6$ in Section 4, so in order to cover smaller values of $m$, the prime $p$ should satisfy the condition in Lemma 11 that $\log_4 p \geq 6m + 11 \geq 47$ which is equivalent to $p \geq 2^{94}$. We have tried to strike a balance between readability and obtaining the best possible constants. These values of $\epsilon$ and $p_0$ are not the best possible, but they give a reasonable account of what can be achieved through the methods of this paper.
References

[1] Y. Bilu, V.F. Lev and I.Z. Ruzsa, Rectification Principles in Additive Number Theory, *Discrete Comput. Geom.* 19 (1998) 343-353.
[2] G. A. Freiman, The addition of finite sets I, *Izv. Vyss. Ucebn. Zaved. Matematika* 3 (13) (1959), 202–213.
[3] G. A. Freiman, Foundations of a structural theory of set addition, *Transl. Math. Monographs* 37, Amer. Math. Soc., Providence, RI, 1973.
[4] B. Green and I.Z. Ruzsa, Sets with small sumset and rectification, *Bull. London Math. Soc.* 38 (2006) 43–52.
[5] Y. O. Hamidoune, On the connectivity of Cayley digraphs, *Europ. J. Combinatorics*, 5 (1984), 309–312.
[6] Y. O. Hamidoune, An isoperimetric method in Additive Theory, *J. Algebra* 179 (1996), 622–630.
[7] Y. O. Hamidoune, On subsets with a small sum in abelian groups, *Europ. J. of Combinatorics* 18 (1997), 541–556.
[8] Y. O. Hamidoune, Some results in Additive Number Theory I: The critical pair Theory, *Acta Arithmetica* 96 (2000) 97–119.
[9] Y. O. Hamidoune and Ø. Rødseth, An inverse theorem modulo p, *Acta Arithmetica* 92 (2000)251–262.
[10] Y. O. Hamidoune, O. Serra and G. Zémor, On the critical pair Theory, *Acta Arithmetica* 121.2 (2006) pp. 99–115.
[11] Y. O. Hamidoune, O. Serra and G. Zémor, On the critical pair theory in abelian groups : Beyond Chowla’s Theorem, to appear in *Combinatorica*. http://arxiv.org/abs/math.NT/0603478
[12] S. Lev and P. Y. Smeliansky, On addition of two sets of integers, *Acta Arith.* 70 (1995), 85–91.
[13] M.B. Nathanson, *Additive Number Theory. Inverse problems and the geometry of sumsets*, Grad. Texts in Math. 165, Springer, 1996.
[14] Ø. Rødseth, On Freiman’s 2.4-Theorem, *K.Norske Vidensk.Selsk.Skr.* 4 (2006) 11–18.
[15] I.Z. Ruzsa, An application of graph theory to additive number theory, *Scientia, Ser. A* 3 (1989), 97 109.
[16] O. Serra and G. Zémor, On a generalization of a theorem by Vosper. *Integers* 0, A10, 10 p., electronic only (2000).
[17] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge University Press, 2006.
[18] G. Vosper, The critical pairs of subsets of a group of prime order, *J. London Math. Soc.* 31 (1956), 200–205.