Minimum number of additive tuples in groups of prime order

Ostap Chervak Oleg Pikhurko∗ Katherine Staden†
Mathematics Institute and DIMAP
University of Warwick
Coventry, U.K.
oschervak@gmail.com, o.pikhurko@warwick.ac.uk

Mathematical Institute
University of Oxford
Oxford, U.K.
staden@maths.ox.ac.uk

Submitted: Oct 5, 2017; Accepted: Jan 30, 2019; Published: Feb 22, 2019
© The authors. Released under the CC BY-ND license (International 4.0).

Abstract

For a prime number \( p \) and a sequence of integers \( a_0, \ldots, a_k \in \{0, 1, \ldots, p\} \), let \( s(a_0, \ldots, a_k) \) be the minimum number of \((k+1)\)-tuples \((x_0, \ldots, x_k) \in A_0 \times \cdots \times A_k \) with \( x_0 = x_1 + \cdots + x_k \), over subsets \( A_0, \ldots, A_k \subseteq \mathbb{Z}_p \) of sizes \( a_0, \ldots, a_k \) respectively.

We observe that an elegant argument of Samotij and Sudakov can be extended to show that there exists an extremal configuration with all sets \( A_i \) being intervals of appropriate length. The same conclusion also holds for the related problem, posed by Bajnok, when \( a_0 = \cdots = a_k = a \) and \( A_0 = \cdots = A_k \), provided \( k \) is not equal 1 modulo \( p \). Finally, by applying basic Fourier analysis, we show for Bajnok’s problem that if \( p \geq 13 \) and \( a \in \{3, \ldots, p-3\} \) are fixed while \( k \equiv 1 \pmod{p} \) tends to infinity, then the extremal configuration alternates between at least two affine non-equivalent sets.

Mathematics Subject Classifications: 11B30, 05D99

1 Introduction

Let \( \Gamma \) be a given finite Abelian group, with the group operation written additively.

For \( A_0, \ldots, A_k \subseteq \Gamma \), let \( s(A_0, \ldots, A_k) \) be the number of \((k+1)\)-tuples \((x_0, \ldots, x_k) \in A_0 \times \cdots \times A_k \) with \( x_0 = x_1 + \cdots + x_k \). If \( A_0 = \cdots = A_k := A \), then we use the shorthand \( s_k(A) := S(A_0, \ldots, A_k) \). For example, \( s_2(A) \) is the number of Schur triples in \( A \), that is, ordered triples \((x_0, x_1, x_2) \in A^3 \) with \( x_0 = x_1 + x_2 \).

For integers \( n \geq m \geq 0 \), let \([m, n] := \{m, m+1, \ldots, n\} \) and \([n] := [0, n-1] \) = \( \{0, \ldots, n-1\} \). For a sequence \( a_0, a_k \in [\lceil \Gamma \rceil+1] = \{0, 1, \ldots, |\Gamma|\} \), let \( s(a_0, \ldots, a_k; \Gamma) \) be

∗Supported by ERC grant 306493 and EPSRC grant EP/K012045/1.
†Supported by ERC grant 306493
the minimum of \( s(A_0, \ldots, A_k) \) over subsets \( A_0, \ldots, A_k \subseteq \Gamma \) of sizes \( a_0, \ldots, a_k \) respectively. Additionally, for \( a \in [0, p] \), let \( s_k(a; \Gamma) \) be the minimum of \( s_k(A) \) over all \( a \)-sets \( A \subseteq \Gamma \).

The question of finding the maximal size of a sum-free subset of \( \Gamma \) (i.e. the maximum \( a \) such that \( s_2(a; \Gamma) = 0 \)) originated in a paper of Erdős [2] in 1965 and took 40 years before it was resolved in full generality by Green and Ruzsa [3]. Huczyńska, Mullen and Yucas [4], and later Samotij and Sudakov [7], introduced the problem of finding \( s_2(a; \Gamma) \). This function has a resemblance to some classical questions in extremal combinatorics, where one has to minimise the number of forbidden configurations, see [7, Section 1] for more details.

Huczyńska, Mullen and Yucas [4] were able to solve the \( s_2 \)-problem for \( \Gamma = \mathbb{Z}_p \), where \( p \) is prime and \( \mathbb{Z}_p \) is the cyclic group of order \( p \). Samotij and Sudakov [7] solved the \( s_2 \)-problem for various groups, including a different proof of the \( \mathbb{Z}_p \) case. Bajnok [1, Problem G.48] suggested the more general problem of considering \( s_k(a; \Gamma) \). Since even the \( s_2 \)-case is still wide open in full generality, Bajnok [1, Problem G.49] proposed, as a possible first step, to consider \( s_k(a; \mathbb{Z}_p) \), where \( p \) is prime and \( k \geq 3 \).

This paper concentrates on the latter question of Bajnok. Therefore, let \( p \) be a fixed prime and let, by default, the underlying group be \( \mathbb{Z}_p \), which we identify with the additive group of residues modulo \( p \) (also using the multiplicative structure on it when this is useful). In particular, we write \( s(a_0, \ldots, a_k) := s(a_0, \ldots, a_k; \mathbb{Z}_p) \) and \( s_k(a) := s_k(a; \mathbb{Z}_p) \).

Since the case \( p = 2 \) is trivial, let us assume that \( p \geq 3 \). By an \( m \text{-term arithmetic progression} \) (or \( m \text{-AP} \) for short) we mean a set of the form \( \{x, x + d, \ldots, x + (m - 1)d\} \) for some \( x, d \in \mathbb{Z}_p \) with \( d \neq 0 \). We call \( d \) the difference. For \( I \subseteq \mathbb{Z}_p \) and \( x, y \in \mathbb{Z}_p \), write \( x \cdot I + y := \{x \cdot z + y \mid z \in I\} \).

As we already mentioned, the case \( k = 2 \) has been completely resolved: Huczyńska, Mullen and Yucas determined \( s_2(a) \), and Samotij and Sudakov [7] showed that, when \( s_2(a) > 0 \), then the \( a \)-sets that achieve the minimum are exactly those of the form \( \xi \cdot I \) with \( \xi \in \mathbb{Z}_p \setminus \{0\} \), where \( I \) consists of the residues modulo \( p \) of \( a \) integers closest to \( \frac{p-1}{2} \in \mathbb{Z} \). Each such set is an arithmetic progression; its difference can be any non-zero value but the initial element has to be carefully chosen.

Here we propose a generalisation of Bajnok’s question, namely to investigate the function \( s(a_0, \ldots, a_k) \). First, by adopting the elegant argument of Samotij and Sudakov [7], we show that at least one extremal configuration consists of \( k + 1 \) arithmetic progressions with the same difference. Note that since

\[
s(A_0, \ldots, A_k) = s(\xi \cdot A_0 + \eta_0, \ldots, \xi \cdot A_k + \eta_k), \quad \text{for} \quad \xi \neq 0 \quad \text{and} \quad \eta_0 = \eta_1 = \cdots = \eta_k,
\]

finding such arithmetic progressions reduces to finding progressions with difference 1 (and starting element 0 for some \( k \) of the sets).

**Theorem 1.** For arbitrary \( k \geq 1 \) and \( a_0, \ldots, a_k \in [0, p] \), there is \( t \in \mathbb{Z}_p \) such that

\[
s(a_0, \ldots, a_k) = s([a_0] + t, [a_1], \ldots, [a_k]).
\]

In particular, if \( a_0 = \cdots = a_k = a \), then one extremal configuration consists of \( A_1 = \cdots = A_k = [a] \) and \( A_0 = [t, t + a - 1] \) for some \( t \in \mathbb{Z}_p \). Given this, one can write
down some formulas for \( s(a_0, \ldots, a_k) \) in terms of \( a_0, \ldots, a_k \) involving summation (based on (3) or a version of (13)) but there does not seem to be a closed form in general.

If \( k \not\equiv 1 \pmod{p} \), then by taking \( \xi := 1, \eta_1 := \cdots := \eta_k := -t(k-1)^{-1} \), and \( \eta_0 := -kt(k-1)^{-1} \) in (1), we can get another extremal configuration where all sets are the same: \( A_0 + \eta_0 = \cdots = A_k + \eta_k \). Thus Theorem 1 directly implies the following corollary.

**Corollary 2.** For every \( k \geq 2 \) with \( k \not\equiv 1 \pmod{p} \) and \( a \in [0, p] \), there is \( t \in \mathbb{Z}_p \) such that \( s_k(a) = s_k([t, t + a - 1]) \).

Unfortunately, if \( k \geq 3 \), then there may be sets \( A \) different from APs that attain equality in Corollary 2 with \( s_k(|A|) > 0 \) (which is in contrast to the case \( k = 2 \)). For example, our (non-exhaustive) search showed that this happens already for \( p = 17 \), when

\[
s_3(14) = 2255 = s_3([-1, 12]) = s_3([6, 18] \cup \{3\}).
\]

Also, already the case \( k = 2 \) of the more general Theorem 1 exhibits extra solutions. Of course, by analysing the proof of Theorem 1 or Corollary 2 one can write a necessary and sufficient condition for the cases of equality. We do this in Section 2; in some cases this condition can be simplified.

However, by using basic Fourier analysis on \( \mathbb{Z}_p \), we can describe the extremal sets for Corollary 2 when \( k \not\equiv 1 \pmod{p} \) is sufficiently large.

**Theorem 3.** Let a prime \( p \geq 7 \) and an integer \( a \in [3, p-3] \) be fixed, and let \( k \not\equiv 1 \pmod{p} \) be sufficiently large. Then there exists \( t \in \mathbb{Z}_p \) for which the only \( s_k(a) \)-extremal sets are \( \xi \cdot [t, t + a - 1] \) for all non-zero \( \xi \in \mathbb{Z}_p \).

**Problem 4.** Find a ‘good’ description of all extremal families for Corollary 2 (or perhaps Theorem 1) for \( k \geq 3 \).

While Corollary 2 provides an example of an \( s_k(a) \)-extremal set for \( k \not\equiv 1 \pmod{p} \), the case \( k \equiv 1 \pmod{p} \) of the \( s_k(a) \)-problem turns out to be somewhat special. Here, translating a set \( A \) has no effect on the quantity \( s_k(A) \). More generally, let \( \mathcal{A} \) be the group of all invertible affine transformations of \( \mathbb{Z}_p \), that is, it consists of maps \( x \mapsto \xi \cdot x + \eta \), \( x \in \mathbb{Z}_p \), for \( \xi, \eta \in \mathbb{Z}_p \) with \( \xi \neq 0 \). Then

\[
s_k(\alpha(A)) = s_k(A), \quad \text{for every } k \equiv 1 \pmod{p} \text{ and } \alpha \in \mathcal{A}. \tag{2}
\]

Let us call two subsets \( A, B \subseteq \mathbb{Z}_p \) (affine) equivalent if there is \( \alpha \in \mathcal{A} \) with \( \alpha(A) = B \). By (2), we need to consider sets only up to this equivalence. Trivially, any two subsets of \( \mathbb{Z}_p \) of size \( a \) are equivalent if \( a \leq 2 \) or \( a \geq p - 2 \).

Again using Fourier analysis on \( \mathbb{Z}_p \), we show the following result.

**Theorem 5.** Let a prime \( p \geq 7 \) and an integer \( a \in [3, p-3] \) be fixed, and let \( k \equiv 1 \pmod{p} \) be sufficiently large. Then the following statements hold for the \( s_k(a) \)-problem.

1. If \( a \) and \( k \) are both even, then \([a]\) is the unique (up to affine equivalence) extremal set.
2. If at least one of \(a\) and \(k\) is odd, define \(I' := [a - 1] \cup \{a\} = \{0, \ldots, a - 2, a\}\). Then

(a) \(s_k(a) < s_k([a])\) for all large \(k\);
(b) \(I'\) is the unique extremal set for infinitely many \(k\);
(c) \(s_k(a) < s_k(I')\) for infinitely many \(k\), provided there are at least three non-equivalent \(a\)-subsets of \(\mathbb{Z}_p\).

It is not hard to see that there are at least three non-equivalent \(a\)-subsets of \(\mathbb{Z}_p\) if and only if \(p \geq 13\) and \(a \in [3, p - 3]\), or \(p \geq 11\) and \(a \in [4, p - 4]\). Thus Theorem 5 characterises pairs \((p, a)\) for which there exists an \(a\)-subset \(A\) which is \(s_k(a)\)-extremal for all large \(k \equiv 1 \pmod{p}\).

**Corollary 6.** Let \(p\) be a prime and \(a \in [0, p]\). There is an \(a\)-subset \(A \subseteq \mathbb{Z}_p\) with \(s_k(A) = s_k(a)\) for all large \(k \equiv 1 \pmod{p}\) if and only if \(a \leq 2\), or \(a \geq p - 2\), or \(p \in \{7, 11\}\) and \(a = 3\).

As is often the case in mathematics, a new result leads to further open problems.

**Problem 7.** Given \(a \in [3, p - 3]\), find a ‘good’ description of all \(a\)-subsets of \(\mathbb{Z}_p\) that are \(s_k(a)\)-extremal for at least one (resp. infinitely many) values of \(k \equiv 1 \pmod{p}\).

**Problem 8.** Is it true that for every \(a \in [3, p - 3]\) there is \(k_0\) such that for all \(k \geq k_0\) with \(k \equiv 1 \pmod{p}\), any two \(s_k(a)\)-extremal sets are affine equivalent?

## 2 Proof of Theorem 1

Here we prove Theorem 1 by adopting the proof of Samotij and Sudakov [7].

Let \(A_1, \ldots, A_k\) be subsets of \(\mathbb{Z}_p\). Define \(\sigma(x; A_1, \ldots, A_k)\) as the number of \(k\)-tuples \((x_1, \ldots, x_k) \in A_1 \times \cdots \times A_k\) with \(x = x_1 + \cdots + x_k\). Also, for an integer \(r \geq 0\), let

\[
N_r(A_1, \ldots, A_k) := \{x \in \mathbb{Z}_p \mid \sigma(x; A_1, \ldots, A_k) \geq r\},
\]

\[
n_r(A_1, \ldots, A_k) := |N_r(A_1, \ldots, A_k)|.
\]

These notions are related to our problem because of the following easy identity:

\[
s(A_0, \ldots, A_k) = \sum_{r=1}^{\infty} |A_0 \cap N_r(A_1, \ldots, A_k)|. \tag{3}
\]

Let an interval mean an arithmetic progression with difference 1, i.e. a subset \(I\) of \(\mathbb{Z}_p\) of form \(\{x, x + 1, \ldots, x + y\}\). Its centre is \(x + y/2 \in \mathbb{Z}_p\); it is unique if \(I\) is proper (that is, \(0 < |I| < p\)). Note the following easy properties of the sets \(N_r\):

1. These sets are nested:

\[
N_0(A_1, \ldots, A_k) = \mathbb{Z}_p \supseteq N_1(A_1, \ldots, A_k) \supseteq N_2(A_1, \ldots, A_k) \supseteq \ldots \tag{4}
\]
2. If each $A_i$ is an interval with centre $c_i$, then $N_r(A_1, \ldots, A_k)$ is an interval with centre $c_1 + \cdots + c_k$.

We will also need the following result of Pollard [6, Theorem 1].

**Theorem 9.** Let $p$ be a prime, $k \geq 1$, and $A_1, \ldots, A_k$ be subsets of $\mathbb{Z}_p$ of sizes $a_1, \ldots, a_k$. Then for every integer $r \geq 1$, we have

$$\sum_{i=1}^{r} n_i(A_1, \ldots, A_k) \geq \sum_{i=1}^{r} n_i([a_1], \ldots, [a_k]).$$

**Proof of Theorem 1.** Let $A_0, \ldots, A_k$ be some extremal sets for the $s(a_0, \ldots, a_k)$-problem. We can assume that $0 < a_0 < p$, because $s(A_0, \ldots, A_k)$ is 0 if $a_0 = 0$ and $\prod_{i=1}^{k} a_i$ if $a_0 = p$, regardless of the choice of the sets $A_i$.

Since $n_0([a_1], \ldots, [a_k]) = p > p - a_0$ while $n_r([a_1], \ldots, [a_k]) = 0 < p - a_0$ when, for example, $r > \prod_{i=1}^{k-1} a_i$, there is a (unique) integer $r_0 \geq 0$ such that

$$n_r([a_1], \ldots, [a_k]) > p - a_0, \quad \text{all } r \in [0, r_0],$$

$$n_r([a_1], \ldots, [a_k]) \leq p - a_0, \quad \text{all integers } r \geq r_0 + 1. \tag{5, 6}$$

The nested intervals $N_1([a_1], \ldots, [a_k]) \supseteq N_2([a_1], \ldots, [a_k]) \supseteq \ldots$ have the same centre $c := ((a_1 - 1) + \cdots + (a_k - 1))/2$. Thus there is a translation $I := [a_0] + t$ of $[a_0]$, with $t$ independent of $r$, which has as small as possible intersection with each $N_r$-interval above given their sizes, that is,

$$|I \cap N_r([a_1], \ldots, [a_k])| = \max\{0, n_r([a_1], \ldots, [a_k]) + a_0 - p\}, \quad \text{for all } r \in \mathbb{N}. \tag{7}$$

This and Pollard’s theorem give the following chain of inequalities:

$$s(A_0, \ldots, A_k) \geq \sum_{i=1}^{\infty} |A_0 \cap N_i(A_1, \ldots, A_k)| \tag{3}
\geq \sum_{i=1}^{r_0} |A_0 \cap N_i(A_1, \ldots, A_k)|
\geq \sum_{i=1}^{r_0} (n_i(A_1, \ldots, A_k) + a_0 - p)
\geq \sum_{i=1}^{r_0} n_i([a_1], \ldots, [a_k]) + a_0 - p \tag{5-6}
\geq \sum_{i=1}^{\infty} \max\{0, n_i([a_1], \ldots, [a_k]) + a_0 - p\}
\geq \sum_{i=1}^{\infty} |I \cap N_i([a_1], \ldots, [a_k])| \tag{7}
\geq s(I, [a_1], \ldots, [a_k]),$$

giving the required. \qed
Let us write a necessary and sufficient condition for equality in Theorem 1 in the case \( a_0, \ldots, a_k \in [1, p - 1] \). Let \( r_0 \geq 0 \) be defined by (5)–(6). Then, by (4), a sequence \( A_0, \ldots, A_k \subseteq \mathbb{Z}_p \) of sets of sizes respectively \( a_0, \ldots, a_k \) is extremal if and only if

\[
\begin{align*}
A_0 \cap N_{r_0+1}(A_1, \ldots, A_k) &= \emptyset, \\
A_0 \cup N_{r_0}(A_1, \ldots, A_k) &= \mathbb{Z}_p, \\
\sum_{i=1}^{r_0} n_i(A_1, \ldots, A_k) &= \sum_{i=1}^{r_0} n_i([a_1], \ldots, [a_k]).
\end{align*}
\]

Let us now concentrate on the case \( k = 2 \), trying to simplify the above condition. We can assume that no \( a_i \) is equal to 0 or \( p \) (otherwise the choice of the other two sets has no effect on \( s(A_0, A_1, A_2) \) and every triple of sets of sizes \( a_0, a_1 \) and \( a_2 \) is extremal). Also, as in [7], let us exclude the case \( s(a_0, a_1, a_2) = 0 \), as then there are in general many extremal configurations. Note that \( s(a_0, a_1, a_2) = 0 \) if and only if \( r_0 = 0 \); also, by the Cauchy-Davenport theorem (the special case \( k = 2 \) and \( r = 1 \) of Theorem 9), this is equivalent to \( a_1 + a_2 - 1 \leq p - a_0 \). Assume by symmetry that \( a_1 \leq a_2 \). Note that (5) implies that \( r_0 \leq a_1 \).

The condition in (10) states that we have equality in Pollard’s theorem. A result of Nazarewicz, O’Brien, O’Neill and Staples [5, Theorem 3] characterises when this happens (for \( k = 2 \)), which in our notation is the following.

**Theorem 10.** For \( k = 2 \) and \( 1 \leq r_0 \leq a_1 \leq a_2 < p \), we have equality in (10) if and only if at least one of the following conditions holds:

1. \( r_0 = a_1 \),
2. \( a_1 + a_2 \geq p + r_0 \),
3. \( a_1 = a_2 = r_0 + 1 \) and \( A_2 = g - A_1 \) for some \( g \in \mathbb{Z}_p \),
4. \( A_1 \) and \( A_2 \) are arithmetic progressions with the same difference.

Let us try to write more explicitly each of these four cases, when combined with (8) and (9).

First, consider the case \( r_0 = a_1 \). We have \( N_{a_1}([a_1], [a_2]) = [a_1 - 1, a_2 - 1] \) and thus \( n_{a_1}([a_1], [a_2]) = a_2 - a_1 + 1 > p - a_0 \), that is, \( a_2 - a_1 \geq p - a_0 \). The condition (8) holds automatically since \( N_i(A_1, A_2) = \emptyset \) whenever \( i > |A_1| \). The other condition (9) may be satisfied even when none of the sets \( A_i \) is an arithmetic progression (for example, take \( p = 13 \), \( A_1 = \{0, 1, 3\} \), \( A_2 = \{0, 2, 3, 5, 6, 7, 9, 10\} \) and let \( A_0 \) be the complement of \( N_3(A_1, A_2) = \{3, 6, 10\} \)). We do not see any better characterisation here, apart from stating that (9) holds.

Next, suppose that \( a_1 + a_2 \geq p + r_0 \). Then, for any two sets \( A_1 \) and \( A_2 \) of sizes \( a_1 \) and \( a_2 \), we have \( N_{r_0}(A_1, A_2) = \mathbb{Z}_p \); thus (9) holds automatically. Similarly to the previous case, there does not seem to be a nice characterisation of (8). For example, (8) may hold...
even when none of the sets $A_i$ is an AP: e.g. let $p = 11$, $A_1 = A_2 = \{0, 1, 2, 3, 4, 5, 7\}$, and let $A_0 = \{0, 2, 10\}$ be the complement of $N_{\frac{p}{2}}(A_1, A_2) = \{1, 3, 4, 5, 6, 7, 8, 9\}$ (here $r_0 = 3$).

Next, suppose that we are in the third case. The primality of $p$ implies that $g \in \mathbb{Z}_p$ satisfying $A_2 = g - A_1$ is unique and thus $N_{\frac{r_0}{2}+1}(A_1, A_2) = \{g\}$. Therefore (8) is equivalent to $A_0 \not\ni g$. Also, note that if $I_1$ and $I_2$ are intervals of size $r_0 + 1$, then $n_{r_0}(I_1, I_2) = 3$. By the definition of $r_0$, we have $p - 2 \leq a_0 \leq p - 1$. Thus we can choose any integer $r_0 \in [1, p - 2]$ and $(r_0 + 1)$-sets $A_2 = g - A_1$, and then let $A_0$ be obtained from $\mathbb{Z}_p$ by removing $g$ and at most one further element of $N_{r_0}(A_1, A_2)$. Here, $A_0$ is always an AP (as a subset of $\mathbb{Z}_p$ of size $a_0 \geq p - 2$) but $A_1$ and $A_2$ need not be.

Finally, let us show that if $A_1$ and $A_2$ are arithmetic progressions with the same difference $d$ and we are not in Case 1 nor 2 of Theorem 10, then $A_0$ is also an arithmetic progression whose difference is $d$. By (1), it is enough to prove this when $A_1 = [a_1]$ and $A_2 = [a_2]$ (and $d = 1$). Since $a_1 + a_2 \leq p - 1 + r_0$ and $r_0 + 1 \leq a_1 \leq a_2$, we have that

$$N_{\frac{r_0}{2}}(A_1, A_2) = [r_0 - 1, a_1 + a_2 - r_0 - 1]$$

$$N_{\frac{r_0}{2}+1}(A_1, A_2) = [r_0, a_1 + a_2 - r_0 - 2]$$

have sizes respectively $a_1 + a_2 - 2r_0 + 1 < p$ and $a_1 + a_2 - 2r_0 - 1 > 0$. We see that $N_{\frac{r_0}{2}+1}(A_1, A_2)$ is obtained from the proper interval $N_{r_0}(A_1, A_2)$ by removing its two endpoints. Thus $A_0$, which is sandwiched between the complements of these two intervals by (8)–(9), must be an interval too. (And, conversely, every such triple of intervals is extremal.)

### 3 The proof of Theorems 3 and 5

Let us recall the basic definitions and facts of Fourier analysis on $\mathbb{Z}_p$. For a more detailed treatment of this case, see e.g. [8, Chapter 2]. Write $\omega := e^{2\pi i/p}$ for the $p$th root of unity. Given a function $f : \mathbb{Z}_p \to \mathbb{C}$, we define its Fourier transform to be the function $\hat{f} : \mathbb{Z}_p \to \mathbb{C}$ given by

$$\hat{f}(\gamma) := \sum_{x=0}^{p-1} f(x) \omega^{-x\gamma}, \quad \text{for } \gamma \in \mathbb{Z}_p.$$

Parseval’s identity states that

$$\sum_{x=0}^{p-1} f(x) \overline{g(x)} = \frac{1}{p} \sum_{\gamma=0}^{p-1} \overline{\hat{f}(\gamma)} \hat{g}(\gamma). \quad (11)$$

The convolution of two functions $f, g : \mathbb{Z}_p \to \mathbb{C}$ is given by

$$(f * g)(x) := \sum_{y=0}^{p-1} f(y) g(x - y).$$

It is not hard to show that the Fourier transform of a convolution equals the product of Fourier transforms, i.e.

$$\hat{f_1} \ast \ldots \ast \hat{f_k} = \hat{f_1} \cdot \ldots \cdot \hat{f_k}. \quad (12)$$
Thus, let $\gamma$ be a parameter of interest, $s_k(A)$, can be written as a simple function of the Fourier coefficients of $1_A$. Indeed, let $A \subseteq \mathbb{Z}_p$ and $x \in \mathbb{Z}_p$. Then the number of tuples $(a_1, \ldots, a_k) \in A^k$ such that $a_1 + \ldots + a_k = x$ (which is $\sigma(x; A, \ldots, A)$ in the notation of Section 2) is precisely $1_A^k(x)$. The function $s_k(A)$ counts such a tuple if and only if its sum $x$ also lies in $A$. Thus,

$$s_k(A) = \sum_{x=0}^{p-1} 1_A^k(x) 1_A(x) = \sum_{\gamma=0}^{p-1} 1_A^k(\gamma) \overline{1_A(\gamma)} = \sum_{\gamma=0}^{p-1} \left( 1_A(\gamma) \right)^k \overline{1_A(\gamma)}. \quad (13)$$

Since every set $A \subseteq \mathbb{Z}_p$ of size $a$ has the same trivial Fourier coefficient (namely $\overline{1_A(0)} = a$), let us re-write (13) as

$$ps_k(A) - a^{k+1} = \sum_{\gamma=1}^{p-1} \left( 1_A(\gamma) \right)^k \overline{1_A(\gamma)} = : F(A). \quad (14)$$

Thus we need to minimise $F(A)$ (which is a real number for any $A$) over $a$-subsets $A \subseteq \mathbb{Z}_p$. To do this when $k$ is sufficiently large, we will consider the largest in absolute value non-trivial Fourier coefficient $\overline{1_A(\gamma)}$ of an $a$-subset $A$. Indeed, the term $(1_A(\gamma))^k \overline{1_A(\gamma)}$ will dominate $F(A)$, so if it has strictly negative real part, then $F(A) < F(B)$ for all $a$-subsets $B \subseteq \mathbb{Z}_p$ with $\max_{\delta \neq 0} |1_B(\delta)| < |1_A(\gamma)|$.

Given $a \in [p-1]$, let

$I := [a] = \{0, \ldots, a - 1\}$ and $I' := [a-1] \cup \{a\} = \{a, \ldots, a - 2, a\}$.

In order to prove Theorems 3 and 5, we will make some preliminary observations about these special sets. The set of $a$-subsets which are affine equivalent to $I$ is precisely the set of $a$-APs.

Next we will show that

$$F(I) = 2 \sum_{\gamma=1}^{(p-1)/2} (-1)^{\gamma(a-1)(k-1)} \left| \overline{1_I(\gamma)} \right|^{k+1} \quad \text{if } k \equiv 1 \pmod{p}. \quad (15)$$

Note that $(-1)^{\gamma(a-1)(k-1)}$ equals $(-1)^\gamma$ if both $a, k$ are even and 1 otherwise. To see (15), let $\gamma \in \{1, \ldots, \frac{p-1}{2}\}$ and write $\overline{1_I(\gamma)} = r e^{\theta i}$ for some $r > 0$ and $0 \leq \theta < 2\pi$. Then $\theta$ is the midpoint of $0, -2\pi\gamma/p, \ldots, -2(a-1)\gamma p/p$, i.e. $\theta = -\pi(a-1)\gamma/p$. Choose $s \in \mathbb{N}$ such that $k = sp + 1$. Then

$$\left( \overline{1_I(\gamma)} \right)^k \overline{1_I(\gamma)} = \left( r e^{-\pi(a-1)\gamma/p} \right)^k r e^{\pi i(a-1)\gamma/p} = r^{k+1} e^{-\pi i(a-1)\gamma s}, \quad (16)$$
and $e^{-\pi i(a-1)s}$ equals 1 if $(a-1)s$ is even, and $-1$ if $(a-1)s$ is odd. Note that, since $p$ is an odd prime, $(a-1)s$ is odd if and only if $a$ and $k$ are both even. So (16) is real, and the fact that $\widehat{f}(p-\gamma) = \widehat{f}(-\gamma)$ implies that the corresponding term for $p-\gamma$ is the same as for $\gamma$. This gives (15). A very similar calculation to (16) shows that

$$F(I + t) = \sum_{\gamma=1}^{p-1} e^{-\pi i (2t+a-1)(k-1)\gamma/p} |\widehat{f}(\gamma)|^{k+1} \text{ for all } k \geq 3. \quad (17)$$

Given $r > 0$ and $0 \leq \theta < 2\pi$, we write $\arg(re^{\theta i}) := \theta$.

**Proposition 11.** Suppose that $p \geq 7$ is prime and $a \in [3, p-3]$. Then $\arg(\widehat{f}(1))$ is not an integer multiple of $\pi/p$.

**Proof.** Since $\widehat{f}(\gamma) = -\widehat{f}(\gamma)$ for all $A \subseteq \mathbb{Z}_p$ and non-zero $\gamma \in \mathbb{Z}_p$, we may assume without loss of generality that $a \leq p - a$. Since $p$ is odd, we have $a \leq (p-1)/2$.

Suppose first that $a$ is odd. Let $m := (a-1)/2$. Then $m \in [1, \frac{p-3}{2}]$. Observe that translating any $A \subseteq \mathbb{Z}_p$ changes the arguments of its Fourier coefficients by an integer multiple of $2\pi/p$. So, for convenience of angle calculations, here we may redefine $I := [-m, m]$ and $I' := \{m - 1\} \cup [-m + 1, m]$. Also let $I^- := [-m + 1, m - 1]$, which is non-empty. The argument of $\widehat{f}(1)$ is 0. Further, $\widehat{f}(1) = \widehat{f}(-1) + \omega^{m+1} + \omega^{-m}$. Since $\omega^{m+1}, \omega^{-m}$ lie on the unit circle, the argument of $\omega^{m+1} + \omega^{-m}$ is either $\pi/p$ or $\pi + \pi/p$. But the bounds on $m$ imply that it has positive real part, so $\arg(\omega^{m+1} + \omega^{-m}) = \pi/p$. By looking at the non-degenerate parallelogram in the complex plane with vertices 0, $\widehat{f}(-1), \omega^{m+1} + \omega^{-m}, \widehat{f}(1)$, we see that the argument of $\widehat{f}(1)$ lies strictly between that of $\widehat{f}(-1)$ and $\omega^{m+1} + \omega^{-m}$, i.e. strictly between 0 and $\pi/p$, giving the required.

Suppose now that $a$ is even and let $m := (a-2)/2 \in [1, \frac{p-5}{4}]$. Again without loss of generality we may redefine $I := [-m, m + 1]$ and $I' := \{m - 1\} \cup [-m + 1, m + 1]$. Let also $I^- := [-m + 1, m]$, which is non-empty. The argument of $\widehat{f}(-1)$ is $-\pi/p$. Further, $\widehat{f}(1) = \widehat{f}(-1) + \omega^{m+1} + \omega^{-(m+1)}$. The argument of $\omega^{m+1} + \omega^{-(m+1)}$ is 0, so as before the argument of $\widehat{f}(1)$ is strictly between $-\pi/p$ and 0, as required.  

---

The Electronic Journal of Combinatorics 26(1) (2019), #P1.30
We say that an $a$-subset $A$ is a punctured interval if $A = I' + t$ or $A = -I' + t$ for some $t \in \mathbb{Z}_p$. That is, $A$ can be obtained from an interval of length $a + 1$ by removing a penultimate point.

**Lemma 12.** Let $p \geq 7$ be prime and let $a \in \{3, \ldots, p - 3\}$. Then the sets $I, I' \subseteq \mathbb{Z}_p$ are not affine equivalent. Thus no punctured interval is affine equivalent to an interval.

**Proof.** Suppose on the contrary that there is $\alpha \in A$ with $\alpha(I') = I$. Let a reflection mean an affine map $R_c$ with $c \in \mathbb{Z}_p$ that maps $x$ to $-x + c$. Clearly, $I = [a]$ is invariant under the reflection $R := R_{a - 1}$. Thus $I'$ is invariant under the map $R' := \alpha^{-1} \circ R \circ \alpha$. As is easy to see, $R'$ is also some reflection and thus preserves the cyclic distances in $\mathbb{Z}_p$. So $R'$ has to fix $a$, the unique element of $I'$ with both distance-1 neighbours lying outside of $I'$. Furthermore, $R'$ has to fix $a - 2$, the unique element of $I'$ at distance 2 from $a$. However, no reflection can fix two distinct elements of $\mathbb{Z}_p$, a contradiction. \hfill \square

We remark that the previous lemma can also be deduced from Proposition 11. Indeed, for any $A \subseteq \mathbb{Z}_p$, the multiset of Fourier coefficients of $A$ is the same as that of $x \cdot A$ for $x \in \mathbb{Z}_p \setminus \{0\}$, and translating a subset changes the argument of Fourier coefficients by an integer multiple of $2\pi/p$. Thus for every subset which is affine equivalent to $I$, the argument of each of its Fourier coefficients is an integer multiple of $\pi/p$.

Let

$$\rho(A) := \max_{\gamma \in \mathbb{Z}_p \setminus \{0\}} |\hat{1}_A(\gamma)| \quad \text{and} \quad R(a) := \left\{ \rho(A) : A \in \left(\mathbb{Z}_p\over a\right) \right\} = \{m_1(a) > m_2(a) > \ldots\}.$$

Given $j \geq 1$, we say that $A$ attains $m_j(a)$, and specifically that $A$ attains $m_j(a)$ at $\gamma$ if $m_j(a) = \rho(A) = |\hat{1}_A(\gamma)|$. Notice that, since $\hat{1}_A(-\gamma) = \overline{\hat{1}_A(\gamma)}$, the set $A$ attains $m_j(a)$ at $\gamma$ if and only if $A$ attains $m_j(a)$ at $-\gamma$ (and $\gamma, -\gamma \neq 0$ are distinct values).

As we show in the next lemma, the $a$-subsets which attain $m_1(a)$ are precisely the affine images of $I$ (i.e. arithmetic progressions), and the $a$-subsets which attain $m_2(a)$ are the affine images of the punctured interval $I'$.

**Lemma 13.** Let $p \geq 7$ be prime and let $a \in [3, p - 3]$. Then $|R(a)| \geq 2$ and

(i) $A \in \left(\mathbb{Z}_p\over a\right)$ attains $m_1(a)$ if and only if $A$ is affine equivalent to $I$, and every interval attains $m_1(a)$ at 1 and $-1$ only;

(ii) $B \in \left(\mathbb{Z}_p\over a\right)$ attains $m_2(a)$ if and only if $B$ is affine equivalent to $I'$, and every punctured interval attains $m_2(a)$ at 1 and $-1$ only.

**Proof.** Given $D \in \left(\mathbb{Z}_p\over a\right)$, we claim that there is some $D_{\text{pri}} \in \left(\mathbb{Z}_p\over a\right)$ with the following properties:

- $D_{\text{pri}}$ is affine equivalent to $D$;
- $\rho(D) = |\hat{1}_{D_{\text{pri}}}(1)|$; and
Call such a $D_{\text{pri}}$ a primary image of $D$. Indeed, suppose that $\rho(D) = |\widehat{1}_D(\gamma)|$ for some non-zero $\gamma \in \mathbb{Z}_p$, and let $\widehat{1}_D(\gamma) = r' e^{\theta' i}$ for some $r' > 0$ and $0 \leq \theta' < 2\pi$. (Note that we have $r' > 0$ since $p$ is prime.) Choose $\ell \in \{0, \ldots, p - 1\}$ and $-\pi/p < \phi < \pi/p$ such that $\theta' = 2\pi\ell/p + \phi$. Let $D_{\text{pri}} := \gamma \cdot D + \ell$. Then

$$|\widehat{1}_{D_{\text{pri}}}(1)| = \left| \sum_{x \in D} \omega^{-\gamma x - \ell} \right| = |\omega^{-\ell} \widehat{1}_D(\gamma)| = |\widehat{1}_D(\gamma)| = \rho(D),$$

and

$$\arg\left( \widehat{1}_{D_{\text{pri}}}(1) \right) = \arg(e^{\theta' i} \omega^{-\ell}) = 2\pi\ell/p + \phi - 2\pi\ell/p = \phi,$$

as required.

Let $D \subseteq \mathbb{Z}_p$ have size $a$ and write $\widehat{1}_D(1) = r e^{\theta i}$. Assume by the above that $-\pi/p < \theta \leq \pi/p$. For all $j \in \mathbb{Z}_p$, let

$$h(j) := \Re(e^{-j\theta} e^{-\theta i}) = \cos\left( \frac{2\pi j}{p} + \theta \right),$$

where $\Re(z)$ denotes the real part of $z \in \mathbb{C}$. Given any $a$-subset $E$ of $\mathbb{Z}_p$, we have

$$H_D(E) := \sum_{j \in E} h(j) = \Re\left( e^{-\theta i} \sum_{j \in E} \omega^{-j} \right) = \Re\left( e^{-\theta i} \widehat{1}_E(1) \right) \leq |\widehat{1}_E(1)|. \quad (18)$$

Then

$$H_D(D) = \sum_{j \in D} h(j) = \Re(e^{-\theta i} \widehat{1}_D(1)) = r = |\widehat{1}_D(1)|. \quad (19)$$

Note that $H_D(E)$ is the (signed) length of the orthogonal projection of $\widehat{1}_E(1) \in \mathbb{C}$ on the 1-dimensional line $\{xe^{\theta} \mid x \in \mathbb{R}\}$. As stated in (18) and (19), $H_D(E) \leq |\widehat{1}_E(1)|$ and this is equality for $E = D$. (Both of these facts are geometrically obvious.) If $|\widehat{1}_D(1)| = m_1(a)$ is maximum, then no $H_D(E)$ for an $a$-set $E$ can exceed $m_1(a) = H_D(D)$. Informally speaking, the main idea of the proof is that if we fix the direction $e^{i\theta}$, then the projection length is maximised if we take $a$ distinct elements $j \in \mathbb{Z}_p$ with the $a$ largest values of $h(j)$, that is, if we take some interval (with the runner-up being a punctured interval).

Let us provide a formal statement and proof of this now.

**Claim 14.** Let $I_a$ be the set of length-$a$ intervals in $\mathbb{Z}_p$.

(i) Let $M_1(D) \subseteq \left( \frac{\mathbb{Z}_p}{a} \right)$ consist of $a$-sets $E \subseteq \mathbb{Z}_p$ such that $H_D(E) \geq H_D(C)$ for all $C \in \left( \frac{\mathbb{Z}_p}{a} \right)$. Then $M_1(D) \subseteq I_a$.

(ii) Let $M_2(D) \subseteq \left( \frac{\mathbb{Z}_p}{a} \right)$ be the set of $E \notin I_a$ for which $H_D(E) \geq H_D(C)$ for all $C \in \left( \frac{\mathbb{Z}_p}{a} \right) \setminus I_a$. Then every $E \in M_2(A)$ is a punctured interval.
Proof. Suppose that $0 < \theta < \pi/p$. Then $h(0) > h(1) > h(-1) > h(2) > h(-2) > \ldots > h\left(\frac{p-1}{2}\right)$. In other words, $h(j_{\ell}) > h(j_k)$ if and only if $\ell < k$, where $j_m := (-1)^m - [m/2]$. Letting $J_{a-1} := \{j_0, \ldots, j_{a-2}\}$, we see that

$$H_D(J_{a-1} \cup \{j_a\}) > H_D(J_{a-1} \cup \{j_a\}), H_D(J_{a-2} \cup \{j_{a-1}, j_a\}) > H_D(J)$$

for all other $a$-subsets $J$. But $J_{a-1} \cup \{j_{a-1}\}$ and $J_{a-1} \cup \{j_a\}$ are both intervals, and $J_{a-1} \cup \{j_{a+1}\}$ and $J_{a-2} \cup \{j_{a-1}, j_a\}$ are both punctured intervals. So in this case $M_1(D) := \{J_{a-1} \cup \{j_{a-1}\}\}$ and $M_2(D) \subseteq \{J_{a-1} \cup \{j_{a+1}\}, J_{a-2} \cup \{j_{a-1}, j_a\}\}$, as required.

The case when $-\pi/p < \theta < 0$ is almost identical except now $j_{\ell} := (-1)^\ell \ell/2$ for all $0 \leq \ell \leq p - 1$. If $0 = \theta$ then $h(0) > h(1) = h(-1) > h(2) > h(-2) > \ldots > h\left(\frac{p-1}{2}\right) = h\left(\frac{-p-1}{2}\right)$. If $\theta = -\pi/p$ then $h(0) = h(-1) > h(1) = h(-2) > \ldots = h\left(\frac{-p-1}{2}\right) > h\left(\frac{p-1}{2}\right)$. □

We can now prove part (i) of the lemma. Suppose $A \in \left(\mathbb{Z}_p\right)^a$ attains $m_1(a)$ at $\gamma \in \mathbb{Z}_p \setminus \{0\}$. Then the primary image $D$ of $A$ satisfies $\lvert \widehat{\mathbf{I}}_D(1) \rvert = m_1(a) = \lvert \widehat{\mathbf{I}}_A(\gamma) \rvert$. So, for any $E \in M_1(D)$,

$$\lvert \widehat{\mathbf{I}}_A(\gamma) \rvert = \lvert \widehat{\mathbf{I}}_D(1) \rvert \overset{(19)}{=} H_D(D) \leq H_D(E) \overset{(18)}{=} \lvert \widehat{\mathbf{I}}_E(1) \rvert,$$

with equality in the first inequality if and only if $D \in M_1(D)$. Thus, by Claim 14(i), $D$ is an interval, and so $A$ is affine equivalent to an interval, as required. Further, if $A$ is an interval then $D$ is an interval if and only if $\gamma = \pm 1$. This completes the proof of (i).

For (ii), note that $m_2(a)$ exists since by Lemma 12, there is a subset (namely $I'$) which is not affine equivalent to $I$. By (i), it does not attain $m_1(a)$, so $\rho(I') \leq m_2(a)$. Suppose now that $B$ is an $a$-subset of $\mathbb{Z}_p$ which attains $m_2(a)$ at $\gamma \in \mathbb{Z}_p \setminus \{0\}$. Let $D$ be the primary image of $B$. Then $D$ is not an interval. This together with Claim 14(i) implies that $H_D(D) < H_D(E)$ for any $E \in M_1(D)$. Thus, for any $C \in M_2(D)$, we have

$$m_2(a) = \lvert \widehat{\mathbf{I}}_B(\gamma) \rvert = \lvert \widehat{\mathbf{I}}_D(1) \rvert = H_D(D) \leq H_D(C) \leq \lvert \widehat{\mathbf{I}}_C(1) \rvert,$$

with equality in the first inequality if and only if $D \in M_2(D)$. Since $C$ is a punctured interval, it is not affine equivalent to an interval. So the first part of the lemma implies that $\lvert \widehat{\mathbf{I}}_C(1) \rvert \leq m_2(a)$. Thus we have equality everywhere and so $D \in M_3(D)$. Therefore $B$ is the affine image of a punctured interval, as required. Further, if $B$ is a punctured interval, then $D$ is a punctured interval if and only if $\gamma = \pm 1$. This completes the proof of (ii).

We will now prove Theorem 3.

Proof of Theorem 3. Recall that $p \geq 7$, $a \in [3, p - 3]$ and $k > k_0(a, p)$ is sufficiently large with $k \equiv 1 \pmod{p}$. Let $I = [a]$. Given $t \in \mathbb{Z}_p$, write $\rho_t := (\widehat{\mathbf{I}}_{t+1}(1))^k \widehat{\mathbf{I}}_{t+1}(1)$ as $r_\ell e^{\theta_\ell}$, where $\theta_\ell \in [0, 2\pi)$ and $r_\ell > 0$. Then (17) says that $\theta_\ell$ equals $-\pi(2t + a - 1)(k - 1)/p$ modulo $2\pi$. Increasing $t$ by 1 rotates $\rho_t$ by $-2\pi(k - 1)/p$. Using the fact that $k - 1$ is invertible modulo $p$, we have the following. If $(a - 1)(k - 1)$ is even, then the set of $\theta_\ell$ for $t \in \mathbb{Z}_p$ is precisely $0, 2\pi/p, \ldots, (2p - 2)\pi/p$, so there is a unique $t$ (resp. a unique $t'$)
in $\mathbb{Z}_p$ for which $\theta_t = \pi + \pi/p$ (resp. $\theta_t = \pi - \pi/p$). Furthermore, $t' = -(a - 1) - t$ and $I + t' = -(I + t)$; thus $I + t$ and $I + t'$ have the same set of dilations. If $(a - 1)(k - 1)$ is odd, then the set of $\theta_t$ for $t \in \mathbb{Z}_p$ is precisely $\pi/p, 3\pi/p, \ldots, (2p - 1)\pi/p$, so there is a unique $t \in \mathbb{Z}_p$ for which $\theta_t = \pi$. We call $t$ (and $t'$, if it exists) optimal.

Let $t$ be optimal. To prove the theorem, we will show that $F(\xi \cdot (I + t)) < F(A)$ (and so $s_k(\xi \cdot (I + t)) < s_k(A)$) for any $a$-subset $A \subseteq \mathbb{Z}_p$ which is not a dilation of $I + t$.

We will first show that $F(I + t) < F(A)$ for any $a$-subset $A$ which is not affine equivalent to an interval. By Lemma 13(i), we have that $|\overline{I}_{I+t}(\pm 1)| = m_1(a)$ and $\rho(A) \leq m_2(a)$. Let $m'_2(a)$ be the maximum of $\overline{I}_J(\gamma)$ over all length-$a$ intervals $J$ and $\gamma \in [2, p - 2]$. Lemma 13(i) implies that $m'_2(a) < m_1(a)$. Thus

$$|F(I + t) - 2(m_1(a))^{k+1} \cos(\theta_t) - F(A)| \leq (p - 1)(m_2(a))^{k+1} + (p - 3)(m'_2(a))^{k+1}. \quad (20)$$

Now $\cos(\theta_t) \leq \cos(\pi - \pi/p) < -0.9$ since $p \geq 7$. This together with the fact that $k \geq k_0(a, p)$ and Lemma 13 imply that the absolute value of $2(m_1(a))^{k+1} \cos(\theta_t) < 0$ is greater than the right-hand size of (20). Thus $F(I + t) < F(A)$, as required.

The remaining case is when $A = \xi \cdot (I + v)$ for some non-optimal $v \in \mathbb{Z}_p$ and non-zero $\xi \in \mathbb{Z}_p$. Since $s_k(A) = s_k(I + v)$, we may assume that $\xi = 1$. Note that $\cos(\theta_t) \leq \cos(\pi - \pi/p) < \cos(\pi - 2\pi/p) \leq \cos(\theta_v)$. Thus

$$F(I + t) - F(I + v) \leq 2(m_1(a))^{k+1}(\cos(\theta_t) - \cos(\theta_v)) + (2p - 4)(m'_2(a))^{k+1}$$

$$\leq 2(m_1(a))^{k+1}(\cos(\pi - \pi/p) - \cos(\pi - 2\pi/p)) + (2p - 4)(m'_2(a))^{k+1}$$

$$< 0$$

where the last inequality uses the fact that $k$ is sufficiently large. Thus $F(I + t) < F(I + v)$, as required.

Finally, using similar techniques, we prove Theorem 5.

Proof of Theorem 5. Recall that $p \geq 7$, $a \in [3, p - 3]$ and $k > k_0(a, p)$ is sufficiently large with $k \equiv 1 \pmod{p}$. Let $I := [a]$ and $I' = [a - 1] \cup \{a\}$.

Suppose first that $a$ and $k$ are both even. Let $A \subseteq \mathbb{Z}_p$ be an arbitrary $a$-set not affine equivalent to the interval $I$. By Lemma 13, $I$ attains $m_1(a)$ (exactly at $x = \pm 1$), while $\rho(A) < m_1(a)$. Also, $m'_2(a) < m_1(a)$, where $m'_2(a) := \max_{\gamma \in [2, p - 2]} |\overline{I}_I(\gamma)|$. Thus

$$F(I) - F(A) \leq \sum_{\gamma = 1}^{p-1} (-1)^\gamma |\overline{I}_I(\gamma)|^{k+1} + \sum_{\gamma = 1}^{p-1} |\overline{I}_A(\gamma)|^{k+1}$$

$$\leq -2(m_1(a))^{k+1} + (2p - 4)(\max\{m_2(a), m'_2(a)\})^{k+1} < 0,$$

where the last inequality uses the fact that $k$ is sufficiently large. So $s_k(a) = s_k(I)$. Using Lemma 13, the same argument shows that, for all $B \in \left(\mathbb{Z}_p \atop a\right)$, we have $s_k(B) = s_k(a)$ if and only if $B$ is an affine image of $I$. This completes the proof of Part 1 of the theorem.
Suppose now that at least one of $a, k$ is odd. Let $A$ be an $a$-set not equivalent to $I$. Again by Lemma 13, we have

$$F(I) - F(A) \geq \sum_{\gamma=1}^{p-1} \left| \mathcal{I}_I(\gamma) \right|^{k+1} - \sum_{\gamma=1}^{p-1} \left| \mathcal{I}_A(\gamma) \right|^{k+1} \geq 2(m_1(a))^{k+1} - (p-1)(m_2(a))^{k+1} > 0.$$  

So the interval $I$ and its affine images have in fact the largest number of additive $(k+1)$-tuples among all $a$-subsets of $\mathbb{Z}_p$. In particular, $s_k(a) < s_k(I)$.

Suppose that there is some $A \in \binom{\mathbb{Z}_p}{a}$ which is not affine equivalent to $I$ or $I'$. (If there is no such $A$, then the unique extremal sets are affine images of $I'$ for all $k > k_0(a,p)$, giving the required.) Write $\rho := re^{\theta i} = I^\rho(1)$. Then by Lemma 13(ii), we have $r = m_2(a)$, and $\rho(A) \leq m_3(a)$. Given $k \geq 2$, let $s \in \mathbb{N}$ be such that $k = sp + 1$. Then

$$\left| F(I') - 2m_2(a)^{k+1} \cos(sp\theta) - F(A) \right| \leq (p-1)m_3(a)^{k+1} + (p-3)(m_2'(a))^{k+1}.$$  \hspace{1cm} (21)

Proposition 11 implies that there is an even integer $\ell \in \mathbb{N}$ for which $c := p\theta - \ell \pi \in (-\pi, \pi) \setminus \{0\}$. Let $\varepsilon := \frac{1}{2} \min\{|c|, \pi - |c|\} > 0$. Given an integer $t$, say that $s \in \mathbb{N}$ is $t$-good if $sc \in ((t - \frac{1}{2})\pi + \varepsilon, (t + \frac{1}{2})\pi - \varepsilon)$. This real interval has length $\pi - 2\varepsilon > |c| > 0$, so must contain at least one integer multiple of $c$. In other words, for all $t \in \mathbb{Z} \setminus \{0\}$ with the same sign as $c$, there exists a $t$-good integer $s > 0$. As $sp\theta \equiv sc \pmod{2\pi}$, the sign of $\cos(sp\theta)$ is $(-1)^t$. Moreover, Lemma 13 implies that $m_2(a) > m_3(a), m_2'(a)$. Thus, when $k = sp + 1 > k_0(a,p)$, the absolute value of $2m_2(a)^{k+1} \cos(sp\theta)$ is greater than the right-hand side of (21). Thus, for large $|t|$, we have $F(A) > F(I')$ if $t$ is even and $F(A) < F(I')$ if $t$ is odd, implying the theorem by (14). \hspace{1cm} \Box

Acknowledgements

We are grateful to an anonymous referee for their careful reading of the manuscript and for many helpful comments.

References


Corrigendum added March 12 2019

After the publication of this paper, we learned that Theorem 1 follows from a result of Lev in [1] (Theorem 1) on solutions to the linear equation $x_1 + \cdots + x_k = 0$ in $\mathbb{Z}_p$.

References