The number of sumsets in a finite field

Noga Alon, Andrew Granville and Adrian Ubis

Abstract

We prove that there are $2^{p/2 + o(p)}$ distinct sumsets $A + B$ in $\mathbb{F}_p$ where $|A|, |B| \to \infty$ as $p \to \infty$.

1 Introduction

For any subsets $A$ and $B$ of a group $G$ we define the sumset

$$A + B := \{a + b : a \in A, b \in B\}$$

There are $2^n$ subsets of an $n$ element additive group $G$ and every one of them is a sumset, since $A = A + \{0\}$ for every $A \subset G$. However if we restrict our summands to be slightly larger then something surprising happens when $G = \mathbb{F}_p$, there are far fewer sumsets:

**Theorem 1.** Let $\psi(x)$ be any function for which $\psi(x) \to \infty$ and $\psi(x) \leq x/4$ as $x \to \infty$. There are exactly $2^{p/2 + o(p)}$ distinct sumsets in $\mathbb{F}_p$ with summands of size $\geq \psi(p)$; that is, exactly $2^{p/2 + o(p)}$ distinct sets of the form $A + B$ with $|A|, |B| \geq \psi(p)$ where $A, B \subset \mathbb{F}_p$.

This cannot be vastly improved due to the following result

**Theorem 2.** For any given prime $p$ and integer $k$ satisfying $k = o(p)$, there exists $A \subset \mathbb{F}_p$ with $|A| = k$ for which there are $\gg 2^{p/2 + o(p)}$ distinct sumsets of the form $A + B$ with $B \subset \mathbb{F}_p$.

These results do not give a good idea of the number of distinct sumsets of the form $A + B$, as $B$ varies through the subsets of $\mathbb{F}_p$ when $A$ has a given small size.
Theorem 3. For each fixed integer $k \geq 1$ there exists a constant $\mu_k \in [\sqrt{2}, 2]$ such that
\[
\max_{A \subseteq \mathbb{F}_p, |A|=k} \# \{ A + B : B \subseteq \mathbb{F}_p \} = \mu_k^{p+o(p)}.
\]
We have $\mu_1 = 2$, $\mu_2 := 1.754877666 \ldots$, the real root of $x^3 - 2x^2 + x - 1$ and, for each fixed integer $k \geq 3$, we have
\[
\sqrt{2} + \frac{1}{3k} \leq \mu_k \leq \sqrt{2} + O\left(\sqrt{\log \frac{k}{k}}\right).
\]
Moreover $\mu_k \leq (5^5/2^33^3)^{1/5} = 1.960131704 \ldots$ for all $k \geq 2$, so that if $|A| \geq 2$ then
\[
\# \{ A + B : B \subseteq \mathbb{F}_p \} \ll 1.9602^{p+o(p)}.
\]

We immediately deduce the following complement to Theorem 1:

Corollary 1. Fix integer $k \geq 1$. Let $\mu_k^* = \max_{\ell \geq k} \mu_\ell$. There are exactly $(\mu_k^*)^{p+o(p)}$ distinct sumsets in $\mathbb{F}_p$ with summands of size $\geq k$.

The existence of $\mu_k$ is deduced from the following result involving sumsets over the integers. Define $S(A,G)$ to be the number of distinct sumsets $A + B$ with $B \subseteq G$; above we have looked at $S(A,\mathbb{F}_p)$, but now we look at $S(A,\{1,2,\ldots,N\})$:

Proposition 1. For any finite set of non-negative integers $A$ with largest element $L$, there exists a constant $\mu_A$ such that $S(A,\{1,2,\ldots,N\}) = \mu_A^N e^{O(L)}$. Moreover
\[
\mu_k = \max_{A \subseteq \mathbb{Z}_{\geq 0}, |A|=k} \mu_A.
\]

By Theorem 3 (or by Theorems 1 and 2 taken together) we know that $\mu_k \to \sqrt{2}$ as $k \to \infty$. In fact we believe that it does so monotonically:

Conjecture 1. We have $\mu_1 = 2 > \mu_2 > \mu_3 > \ldots > \mu_k > \ldots > \sqrt{2}$.

If this is true then $\mu_k^* = \mu_k$, evidently.

One can ask even more precise questions, for example for the number of distinct sumsets $A + B$ where the sizes of $A$ and $B$ are given: Define
\[
S_{k,\ell}(G) = \# \{ A + B : A,B \subseteq G, |A| = k, |B| = \ell \}.
\]
for any integers $k,\ell \geq 1$. 

2
Theorem 4. We have the following bounds for $S_{k,\ell}(\mathbb{F}_p)$

(i) If $k + \ell = o(p)$ then $S_{k,\ell}(\mathbb{F}_p) = \left(\frac{p}{k+\ell}\right)^{1+o(1)} = 2^{o(p)}$.

(ii) If $k + \ell < p/2$ then $S_{k,\ell}(\mathbb{F}_p) \geq \left(\frac{p}{k+\ell}\right)^{p^{O(1)}}$.

(iii) If $k, \ell \geq p/(\log p)^{1/4}$ then $S_{k,\ell}(\mathbb{F}_p) \ll \left(\frac{x}{k+\ell}\right)^{1+o(1)}$ with $x$ such that $2^{p-x} \sim \left(\frac{x}{k+\ell}\right)$.

In particular $S_{k,\ell}(\mathbb{F}_p) \ll 2^{(1/2-\epsilon)p}$ unless $k + \ell = p/4 + O(\epsilon p)$, and therefore if $k, \ell \geq p/(\log p)^{1/4}$ with $k + \ell \sim p/4$ then

$$S_{k,\ell}(\mathbb{F}_p) = (\sqrt{2})^{p+o(p)}.$$

The structure of sumsets has a rich history, from Cauchy [Cau] onwards, and has been studied from several different perspectives. Most important are lower bounds on the size of the sumset, the lattice structure of $A$ and $B$ when the size of the sumset $A + B$ is very small (i.e. the Freiman-Ruzsa theorem), and the discovery of long arithmetic progressions in the sumset $A + B$ when it is fairly small.

2 Lower bounds

For a given integer $k$ let $A = \{0, [(p - k)/2] + 1, [(p - k)/2] + 2, \ldots, [(p - k)/2] + k - 1\}$. For any subset $B$ of $\{0, 1, 2, \ldots, [(p - k)/2]\}$, we see that $A + B \subset [0, p - 1]$ so that $B = (A + B) \cap \{0, 1, \ldots, [(p - k)/2]\}$, and thus the sets $A + B$ are all distinct. Hence there are at least $2^{[(p-k)/2]+1} \geq 2^{[(p-k)/2]}$ distinct sets $A + B$ as $B$ varies over the subsets of $\mathbb{F}_p$. This implies Theorem 2 hence the lower bound in Theorem 1 when $\psi(p) = o(p)$, and it also implies the lower bound on $\mu_k$ in Theorem 3.

Suppose that $u + v < p/2$. Let $A_1 \subset \{0, 1, \ldots, u\}$ with $|A_1| = k - 1$ and $A = A_1 \cup \{u + v\}$; and $B_1 \subset \{0, 1, \ldots, v - 1\}$ with $|B_1| = \ell - 1$ and $B = B_1 \cup \{u + 2v\}$. Then $B_1 + \{u + v\} = (A + B) \cap \{u + v, u + v + 1, \ldots, u + 2v - 1\}$, and $A_1 + \{u + 2v\} = (A + B) \cap \{u + 2v, u + v + 1, \ldots, 2u + 2v\}$, so that the sets $A + B$ are all distinct. Hence there are at least $\binom{u+1}{k-1}\binom{v}{\ell}$ distinct sets $A + B$ as $A$ and $B$ vary over the subsets of $\mathbb{F}_p$ of size $k$ and $\ell$ respectively.

We can maximize this lower bound by taking $u = \lceil p(k - 1)/2(k + \ell - 1) \rceil$ and
\[ v = \lfloor p\ell / 2(k + \ell - 1) \rfloor, \] and obtain \( p^{O(1)}(\frac{p}{k+\ell}) \) by Stirling’s formula. This implies the second part of Theorem 4.

Now, given \( k \leq p/4 \), select \( \ell = \lfloor p/4 \rfloor - k \) so that, by the above, there are \( p^{O(1)}(\frac{p}{k+\ell}) = 2^{p/2}p^{O(1)} \) distinct sumsets \( A + B \) as \( A \) and \( B \) vary over the subsets of \( \mathbb{F}_p \) of size \( k \) and \( \ell \) respectively; which implies the lower bound in Theorem 1.

### 3 First upper bounds

The number of distinct sumsets \( A + B \) with \( |A| = k \) and \( |B| = \ell \), is no more than the number of pairs of such sets \( A, B \), which equals \( \binom{k}{\ell} \). This equals \( \binom{p/2}{k+\ell} 2^{O(k+\ell)} \), since \( 1 \leq \binom{k+\ell}{k} \leq 2^{k+\ell} \), which is \( \binom{p/2}{k+\ell}^{1+o(1)} \) provided \( k + \ell = o(p) \). This implies the upper bound in the first part of Theorem 4.

Throughout we denote \( r_{C-A}(n) \) to be the number of representations of \( n \) as \( c-a \) with \( a \in A \) and \( c \in C \).

**Proposition 2.** Let \( G \) be an abelian group of order \( n \) and let \( A \subset G \) be a subset of size \( k \). Then

\[
\#\{A + B : B \subset G\} \leq n \min_{2 \leq \ell \leq k} \sum_{j=0}^{n} \binom{n}{\lfloor j/\ell \rfloor} \min\{2^{n-j}, 2^{[j/(k-\ell+1)]}\}. \tag{3}
\]

**Proof.** Given a set \( B \) we order the elements of \( B \) by greed, selecting any \( b_1 \in B \), and then \( b_2 \in B \) so as to maximize \((A+\{b_2\}) \setminus (A+\{b_1\})\), then \( b_3 \in B \) so as to maximize \((A+\{b_3\}) \setminus (A+\{b_1, b_2\})\), etc. Let \( B_\ell \) be the set of \( b_i \) such that \( A+\{b_1, b_2, \ldots, b_i\} \) contains at least \( \ell \) more elements than \( A+\{b_1, b_2, \ldots, b_{i-1}\} \), and suppose that \( |B_\ell + A| = j \). By definition \( j = |B_\ell + A| \geq \ell |B_\ell| \), so that \( |B_\ell| \leq \lfloor j/\ell \rfloor \) and so there are no more than \( \sum_{i \leq \lfloor j/\ell \rfloor} \binom{n}{i} \) choices for \( B_\ell \). Note that \( j/\ell \leq n/2 \) and so \( \sum_{i \leq \lfloor j/\ell \rfloor} \binom{n}{i} \leq n \binom{n}{\lfloor j/\ell \rfloor} \). Next we have to determine the number of possibilities for \( A + B \) given \( B_\ell \) (and hence \( B_\ell + A \)).

Our first argument: Since \( B_\ell + A \subset B + A \subset G \), the number of such sets \( A + B \) is at most the total number of sets \( C \) for which \( B_\ell + A \subset C \subset G \), which equals \( 2^{n-j} \).

Our second argument: Let \( C = B_\ell + A \), and \( D \) be the set of \( d \in G \) for which \( r_{C-A}(d) \geq k+1-\ell \). If \( b \in B \setminus B_\ell \) then \( r_{C-A}(b) = |(b+A) \cap (B_\ell + A)| \geq k+1-\ell \), so that \( b \in D \). Hence \( (B \setminus B_\ell) \subset D \), and so there are \( \leq 2^{|D|} \) possible sets...
$B \setminus B_{\ell}$, and hence $B$, and hence $A + B$. Now

$$|D|(k + 1 - \ell) \leq \sum_{d \in G} r_{C - A}(d) = |A| |C| = kj,$$

so that $|D| \leq kj/(k + 1 - \ell)$, and the result follows.

**Simplifying the upper bound:** The upper bound in Proposition 2 is evidently

$$\leq n^2 \min_{2 \leq \ell \leq k} \max_{0 \leq j \leq n} \left( \frac{n}{j/\ell} \right) \min\{2^{n-j}, 2^{jk/(k-\ell+1)}\}.$$

Now $\left( \frac{n}{j/\ell} \right) 2^{jk/(k-\ell+1)}$ is a non-decreasing function of $j$, as $\ell \geq 2$, and so the above is

$$\leq n^2 \min_{2 \leq \ell \leq k} \max_{\lfloor j/\ell \rfloor n \leq j \leq n} \left( \frac{n}{j/\ell} \right) 2^{n-j}.$$

The $(j + \ell)$th term equals the $j$th term times $(n - \lfloor j/\ell \rfloor)/2^\ell (\lfloor j/\ell \rfloor + 1)$. This is $< 1$ if and only if $n < (2^\ell + 1)\lfloor j/\ell \rfloor + 2^\ell$. Now

$$\frac{(2^\ell + 1)\lfloor j/\ell \rfloor + 2^\ell}{\ell} \geq \frac{2^\ell + 1}{\ell} \cdot \frac{(k - \ell + 1)}{(2k - \ell + 1)} n,$$

and this is $\geq n$ unless $\ell = k \leq 4$. Hence one minimizes by taking $j = \lfloor k/\ell \rfloor n + O(1)$ at a cost of a factor of at most $n$. Therefore our bound becomes $\ll n^{O(1)} \nu_k^n$ where $\nu_k := \min_{2 \leq \ell \leq k} \nu_{k,\ell}$ and

$$\nu_{k,\ell} := \left( \frac{2^k (\ell(2k - \ell + 1))^{2k-\ell+1}}{(k - \ell + 1)^{k-\ell+1} (\ell(2k - \ell + 1) - (k - \ell + 1))^{2k-\ell+1-\ell+1}} \right)^{\frac{1}{2k-\ell+1}},$$

using Stirling’s formula. A brief Maple calculation yields that $\nu_k > 2$ for all $k \leq 7$ and $\nu_8 = 1.982301294$, $\nu_9 = 1.961945316$, $\nu_{10} = 1.942349376$, ..., with $\nu_k < 1.91$ for $k \geq 12$, and $\nu_k$ decreasing rapidly and monotonically (e.g. $\nu_k < 1.9$ for $k \geq 13$, $\nu_k < 1.8$ for $k \geq 23$, $\nu_k < 1.7$ for $k \geq 45$, and $\nu_k < 1.6$ for $k \geq 117$). In general taking $\ell$ so that $\ell^2 \sim k \log k/\log 2$, one gets that

$$\nu_k = \sqrt{2} \exp \left( \frac{1}{2} + o(1) \right) \sqrt{\frac{\log 2 \cdot \log k}{k}},$$

which implies the upper bound in (2) of Theorem 3 and the upper bound implicit in Theorem 1 provided we add in the hypothesis that $\min\{|A|, |B|\}$ is $o(p)$. 

5
4 Sumsets from big sets

We modify, simplify and generalize Green and Ruzsa’s argument [GrRu], which they used to bound the number of sumsets $A + A$ in $\mathbb{F}_p$: For a given set $S$, define $d_S := \{ds : s \in S\}$. Let $G = \mathbb{Z}/m\mathbb{Z}$. For any $A \subset G$ define $\hat{A}(x) = \sum_{a \in A} e(ax/m)$. For a given positive integer $L < m$ let $H$ be the set of integers in the interval $[-(L - 1), L - 1]$. For a given integer $d$ with $1 < dL < m$ we partition the integers in $[1, m]$ as best we can into arithmetic progressions with difference $d$ and length $L$. That is for $1 \leq i \leq d$ we have the progressions $I_{i,k} := \{i + jd : kL \leq j \leq \min\{(k + 1)L - 1, ((m - i)/d)\}$ for $0 \leq k \leq [(m - i)/Ld]$. We then let $A_{L,d}$ be the union of the $I_{i,k}$ that contain an element of $A$ (so that $A \subset A_{L,d}$). Note that there are $\leq \lceil m/L \rceil + d$ such intervals $I_{i,k}$.

Our goal is to prove the following analogy to Proposition 3 in [GrRu]:

Proposition 3. If $A, B \subset \mathbb{Z}/m\mathbb{Z}$, with $\alpha = |A|/m$ and $\beta = |B|/m$ and

$$m > (4L)^{1+32\alpha\beta L^4\epsilon_2^2\epsilon_3^{-1}},$$

then there exists an integer $d$, with $1 \leq d \leq m/4L$, such that $A + B$ contains all those values of $n$ for which $r_{A_{L,d}+B_{L,d}}(n) > \epsilon_2 m$, with no more than $\epsilon_3 m$ exceptions.

Proof. In this paragraph we follow the proof of Proposition 3 in [GrRu] (with the obvious modifications): There exists $1 \leq d \leq m/4L$ such that

$$|\hat{A}(x)|^2 \left| 1 - \hat{H}(dx)/(2L - 1) \right|^2 \leq \frac{\log 4L}{\log(m/4L)} |A|m,$$

and thence, by Parseval’s formula,

$$\sum_n \left| r_{A+B}(n) - \frac{r_{A+dH+B+dH}(n)}{(2L - 1)^2} \right|^2 \leq \frac{1}{m} \sum_x |\hat{A}(x)|^2 |\hat{B}(x)|^2 \left| 1 - \left( \frac{\hat{H}(dx)}{2L - 1} \right)^2 \right|^2 \leq \frac{2\log 4L}{\log(m/4L)} |A||B|m.$$

Now if $g \in A_{L,d}$ then there exists $j \in H$ such that $g + dj \in A$, by definition, and hence $r_{A+dH}(g) \geq 1$. Therefore $r_{A+dH}(g) \geq r_{A_{L,d},d}(g)$ for all $g \in G$, so that

$$r_{A+dH+B+dH}(n) \geq r_{A_{L,d}+B_{L,d}}(n)$$
for all $n$. Therefore if $N$ is the set of $n \notin A + B$ such that $r_{A_{L,d} + B_{L,d}}(n) > \epsilon_2 m$, then $r_{A + dH + B + dH}(n) > \epsilon_2 m$ and the above yields $|N| \leq \epsilon_3 m$ by restricting the sum of squares to a sum over $n \in N$, in our range for $m$. 

Next we prove a combinatorial lemma based on Proposition 5 of [GrRu]:

**Proposition 4.** For any subsets $C, D$ of $\mathbb{F}_p$, and any $m \leq r \leq \min(|C|, |D|)$, there are at least $\min(|C| + |D|, p) - r - (m - 1)p/r$ values of $n$ (mod $p$) such that $r_{C + D}(n) \geq m$.

**Proof.** Pollard’s generalization of the Cauchy-Davenport Theorem [Pol] states that

$$\sum_n \max\{r, r_{C+D}(n)\} \geq r(\min(p, |C| + |D|) - r).$$

The left hand side here is $\leq (m - 1)(p - N_m) + rN_m$ where $N_m$ is the number of $n$ (mod $p$) such that $r_{C+D}(n) \geq m$. The result follows since $p - N_m \leq p$. 

**Proof of upper bounds on $S_{k, \ell}(\mathbb{F}_p)$:** Suppose that $L$ is given and $d \leq p/4L$, and that $M$ and $N$ are unions of some of the arithmetic progressions $I_{i,j}$. Note that there are $\leq 2^{p/L+d}$ such sets $M$ (given $d$), and hence a total of $\epsilon^{O(p/L)}$ possibilities for $d, M$ and $N$.

We now bound the number of distinct sumsets $A + B$ for which $A_{L,d} = M$ and $B_{L,d} = N$ in two different ways:

First, since $A \subset M$ and $B \subset N$ there can be no more than $\binom{|M|}{k} \binom{|N|}{\ell} \leq 2^{\binom{|M|+|N|}{k+\ell}}$ such pairs.

Second, select $2\epsilon_1 p \leq \min(|M|, |N|)$ and $2\epsilon_3 p \leq \max(|M|, |N|)$. Let $Q$ be the values of $n$ (mod $p$) such that $r_{M+N}(n) \geq \epsilon_1 p$. Taking $r = \epsilon_1 p$ and $m = \epsilon_1^2 p$ in Proposition 4, we have $|Q| \geq R := \min(|M| + |N|, p) - 2\epsilon_1 p$. By Proposition 3, $A + B$ is given by $Q$ less at most $\epsilon_3 p$ elements, union some subset of $\mathbb{F}_p \setminus Q$. Hence the number of distinct sumsets $A + B$ is

$$\leq 2^{p-|Q|} \sum_{i=0}^{\epsilon_3 p} \binom{|Q|}{i} \leq p2^{p-|Q|} \binom{|Q|}{\epsilon_3 p} \leq p2^{\max(p-|M|-|N|,0)+2\epsilon_1 p} \binom{|M|+|N|}{\epsilon_3 p}$$

as $|Q| \geq R > 2\epsilon_3 p$.

If $|M| + |N| \leq p/2$ then the number of sumsets is $\leq 2^{p/2}$ by the first argument. Let $L = [(\log p)^{1/10}]$ and $\epsilon_1 = \epsilon_3 = 1/2L$. If $|A|, |B| > p/L$ then $|M| \geq |A| > 2\epsilon_1 p$ and $|N| \geq |B| > 2\epsilon_1 p$. so the second argument
is applicable; therefore if $|M| + |N| > p/2$ then the number of sumsets is
\[ \leq 2^{p/2}L^{O(p/L)}. \]
Hence the total number of sumsets $A + B$ with $|A|, |B| > p/L$
is $\ll 2^{p/2}L^{O(p/L)}$ which implies the upper bound in Theorem 1 (taken together
with the argument, for $\min\{|A|, |B|\} = o(p)$, given at the end of section 3).

Without loss of generality we may assume $k \geq \ell$, and we now suppose
that $k, \ell \geq p/(\log p)^{1/4}$. We select $\epsilon_1 = \ell/2p \log \log p$, $\epsilon_2 = k/2p \log \log p$ and $L = [(\log p)^{1/20}]$, so that the second argument above is applicable. Taking

\[ x = |M| + |N| \]
we have that

\[ S_{k,\ell}(\mathbb{F}_p) \leq \max_{0 \leq x \leq 2p} \min \left\{ \left( \frac{x}{k + \ell} \right)^{2^\max(p-x,0)} , (1/\epsilon_2)^{O(\epsilon_3^p)} = 2^{(1+o(1))(p-x')} \right\} \]

where $x'$ is chosen so that $2^{p-x'} \sim \binom{x'}{k+\ell}$. Note that $x' \gtrsim p/2$ unless $k + \ell \sim p/4$, in which case $x' \sim p/2$.

5 Sumsets in finite fields and the integers

Let $A \subseteq \mathbb{F}_p$ be of given size $k \geq 2$, and let $d = [p^{1-1/k}]$. Consider the sets

\[ iA, \text{ the least residues of } ia, a \in A, \text{ for } 0 \leq i \leq p - 1. \]
Two, say $iA$ and $jA$ with $i \not\equiv j \pmod{p}$, must have those least residues between the same
two multiples of $p^{1-1/k}$ for each $a \in A$ (since there are $< (p/p^{1-1/k})^k = p$
possibilities), and so the least residues of $\ell a, a \in A$, with $\ell = i - j$ are all $\leq d$
in absolute value. Hence the elements of $d + \ell A$ are all integers in $[0, 2d]$; and $S(A, \mathbb{F}_p) = S(d + \ell A, \mathbb{F}_p)$ as may be seen by mapping $A + B \to (d + \ell A) + (\ell B)$. Hence we may assume, without loss of generality, that $A$ a set of integers in

\[ [0, L] \]
where $L \leq 2p^{1-1/k}$.

The case $k = 2$ is of particular interest since then $S(A, \mathbb{F}_p) = S(\{0, 1\}, \mathbb{F}_p)$
by taking $\ell = 1/(b - a)$, $d = -a\ell$ when $A = \{a, b\}$.

We now compare $S(A, \mathbb{F}_p)$ with $S(A, \{1, 2, \ldots, p\})$. When we reduce $A +$ $B$, where $A \subseteq \{0, \ldots, L\}$ and $B \subseteq \{0, \ldots, p - 1\}$ are sets of integers, modulo
$p$, the reduction only effects the residues in $\{0, \ldots, L - 1\}$ (mod $p$). Hence

\[ S(A, \{1, 2, \ldots, p\}) \mid 2^{-L} \leq S(A, \mathbb{F}_p) \leq S(A, \{1, 2, \ldots, p\}). \]

Now suppose $A \subseteq \{0, \ldots, L\}$ is a set of integers. Suppose that $Mr \leq N <$
$M(r + 1)$ for positive integers $M, r, N$. We see that $S(A, \{1, 2, \ldots, N\}) \leq S(A, \{1, 2, \ldots, M(r + 1)\}) \leq S(A, \{1, 2, \ldots, M\})^{r+1}$, the last inequality coming since the sumsets $A + B$ with $B \subseteq \{1, 2, \ldots, M(r + 1)\}$ are the union of the
sumsets $A + B_i$ with $B_i \subset \{Mi + 1, 2, \ldots, M(i + 1)\}$ for $i = 0, 1, 2, \ldots, r$. In the other direction we note that if $B = \bigcup B_i$ where $B_i \subset \{Mi + 1, 2, \ldots, M(i + 1) - L\}$ then distinct $\{A + B_i\}_{0 \leq i \leq r - 1}$ give rise to distinct $A + B$. Hence $S(A, \{1, 2, \ldots, N\}) \geq S(A, \{1, 2, \ldots, Mr\}) \geq S(A, \{1, 2, \ldots, M - L\})^r \geq (S(A, \{1, 2, \ldots, M\})2^{-L})^r$. In particular for $m_A(N) := S(A, \{1, 2, \ldots, N\})^{1/N}$ we have

$$m_A(M)^{1+1/r} \geq m_A(N) \geq m_A(M)^{1-1/r}2^{-L/M}.$$ 

This implies that if $M' \geq M$ then $m_A(M') - m_A(M) \ll L/M$ (by taking $N = (M')^2$) and so the $m_A(M)$ form a Cauchy sequence and hence tend to a limit, $\mu_A$. Taking $N$ very large above we see that $\mu_A \leq m_A(M) \leq \mu_A2^{-L/M}$, that is

$$S(A, \{1, 2, \ldots, N\}) = \mu_A^N e^{O(L)}$$

uniformly, and hence

$$S(A, \mathbb{F}_p) = \mu_A^p e^{O(L)} = \mu_A^p e^{O(p^{1-1/k})} = \mu_A^{p+o(p)}$$

via the first displayed equation in this section. This proves Proposition 1, as well as the first part of Theorem 3.

### 5.1 Precise bounds when $k = 2$

By the previous section we know that $\mu_2 = \mu_{\{0,1\}}$. Now $S$ is a subset of the form $\{0, 1\} + B$ if and only if, when one writes the sequence of 0’s and 1’s given by $s_n = 1$ if $n \in S$, otherwise $s_n = 0$ if $n \not\in S$, there are no isolated 1’s.

Let $C_n$ be the number of sequences of 0’s and 1’s of length $n$ such that there are no isolated 1’s, so that $S(\{0, 1\}, \{1, 2, \ldots, N\}) = C_{N+1}$. We can determine $C_{n+1}$ by induction: If the $n + 1$th element added is a 0 then it can be added to any element of $C_n$. If the $n + 1$th element added is a 1 then the $n$th digit must be a 1, and then we either have an element of $C_{n-1}$ or we the next two digits are 1 and 0 followed by any element of $C_{n-3}$. Hence $C_{n+1} = C_n + C_{n-1} + C_{n-3}$ with $C_1 = 1, C_2 = 2, C_3 = 4, C_4 = 7$. In fact it is easily checked, by induction, that $C_{n+1} = 2C_n - C_{n-1} + C_{n-2}$ (which is explained by the fact that $x^4 - x^3 - x^2 - 1 = (x + 1)(x^3 - 2x^2 + x - 1)$, where the higher degree polynomials are characteristic polynomials for the recurrence sequence), and hence $C_n \sim c\mu_A^2$ for some constant $c > 0$, implying a strong form of the first part of Theorem 3 for $k = 2$. 
5.2 Precise bounds when $k = 3$

We know that $\mu_3 = \mu_{\{0,a,b\}}$ for some small positive coprime $a$ and $b$. Maple experimentation leads us to guess that $\mu_3 = \mu_{\{0,1,4\}} = 1.6863\ldots$, a root of an irreducible polynomial of degree 21. We can reduce the hunt for the value of $\mu_3$ to a finite but tedious search, by developing some of the ideas above, but would prefer to find a more illuminating method. (The key observation is that $\mu_{\{0,a,b\}} \to \mu_*$ as $a + b \to \infty$, where we define

$$
\mu_* = \lim_{p \to \infty} \# \{B + \{(0,0),(1,0),(0,1)\} : B \subset \mathbb{F}_p \times \mathbb{F}_p\}^{1/p^2},
$$

which we can prove exists, and is $< \mu_2$.)

5.3 Lower bound on $\mu_k$

Let $A_k = \{1, 3, \ldots, 3^{k-1}\}$, and write and $B \subset \{1, 2, \ldots, 3n\}$ as $B = 3B_0 \cup (3B_1 - 1)$ with $B_0, B_1 \subset \{1, 2, \ldots, n\}$. Since $A_{k+1} = 1 \cup 3A_k$ we have

$$(B + A_{k+1}) \setminus \mathbb{Z} = (3(B_1 + A_k) - 1) \cup (3B_0 + 1),$$

which shows that $S(A_{k+1}, \{1, 2, \ldots, 3n\}) \geq S(A_{k+1}, \{1, 2, \ldots, n\}) 2^n$, and so

$$
\mu_{A_{k+1}} \geq 2^{\frac{1}{3}} \mu_A^{\frac{1}{3}}.
$$

Since $\mu_{A_1} = 2$, an induction argument implies $\mu_k \geq \mu_{A_k} \geq 2^{1/2 + 3^{-k}/2}$, which gives the lower bound for $\mu_k$ in (2).

6 A non-trivial bound for fixed $k \geq 2$

Let $A$ be any set of given size $k \geq 2$. For any two distinct elements $a, b \in A$ we can map $x \to (x - a)/(b - a)$ so that $0, 1 \in A$, and this will not effect the count of the number of sumsets containing $A$.

The number of sumsets $C = A + B$ with $B \subset \mathbb{F}_p$ is obviously bounded above by

$$
\# \left\{ B : |B| \leq \frac{2p}{5} \right\} + \# \left\{ C : |C| \geq \frac{3p}{5} \right\} + \# \left\{ C : \exists B : \frac{2p}{5} < |B| < |C| < \frac{3p}{5} \text{ and } B + \{0,1\} \subset C \right\}.
$$
The first two terms have size $\leq 2p^{(p/2p)/5}$, the third requires some work: We observe that such $C$ must have at least $(1/2 - \delta)p$ pairs of consecutive elements; so if $c$ is the smallest integer $\geq 1$ that belongs to $c$ then we suppose that $C = \cup_{k=1}^m (c + I_k)$ and $\overline{C} = \cup_{k=1}^m (c + J_k)$ where $I_1, J_1, I_2, J_2, \ldots, I_m, J_m$ is a partition of $\{0, \ldots, p - 1\}$ into non-empty set of integers from intervals taken in order. Any such set partition will do provided, for $i_k = \mid I_k \mid$ and $j_k = \mid J_k \mid$, we have each $i_k, j_k \geq 1$,

$$\frac{3p}{5} \geq \sum_{k=1}^m i_k \geq m + \frac{2p}{5},$$

since $|C| = \sum_{k=1}^m i_k$ and $\sum_{k=1}^m (i_k - 1) \geq |B|$, and $\sum_{k=1}^m i_k + \sum_{k=1}^m j_k = p$. Now there are $\leq p$ possible values for $c$, and the number of possible sets of values of $i_k$ such that $\sum_{k=1}^m i_k = x$ is $\binom{x-1}{m-1}$, and of $j_k$ is $\binom{p-x-1}{m-1}$. Therefore the number of possible such $C$ is

$$\leq p \sum_{m \leq p/5} \sum_{2p/5+m \leq x \leq 3p/5} \binom{x-1}{m-1} \binom{p-x-1}{m-1}.$$

$$\leq p^2 \sum_{m \leq p/5} \binom{p-2}{2m-2} \leq p^3 \binom{p-2}{2p/5 - 2} \ll p^3 \binom{p}{2p/5}.$$

Hence the number of sumsets $A + B$ is $\ll p^3 \binom{p}{2p/5} = c^{p+o(p)}$ where $c = (5^5/2^33^3)^{1/5} = 1.960131704\ldots$. This implies the bound $\mu_k \leq c$ for all $k \geq 2$ of Theorem 3; and we deduce the last part of Theorem 3 immediately from this taken together with Theorem 1.

References


