The number of sumsets in a finite field

Noga Alon * Andrew Granville † Adrián 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 the situation changes dramatically (at least when $G = \mathbb{F}_p$): there are far fewer sumsets, as the following result shows.

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

Green and Ruzsa [GrRu] proved that there are only $2^{p/3+o(p)}$ distinct sumsets $A+A$ in $\mathbb{F}_p$. The count in Theorem 1 cannot be decreased by restricting the size of one of the sets.

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 at least $2^{p/2+o(p)}$ distinct sumsets of the form $A+B$ with $B \subset \mathbb{F}_p$.*

*Tel Aviv University, Tel Aviv 69978, Israel and IAS, Princeton, NJ, 08540, USA. Research supported by the Israel Science Foundation, by a USA-Israel BSF grant, and by the Ambrose Monell Foundation. Email: nogaa@tau.ac.il

†Université de Montréal, Montréal QC H3C 3J7, Canada. L'auteur est partiellement soutenu par une bourse de la Conseil de recherches en sciences naturelles et en génie du Canada. Email: andrew@DMS.UMontreal.CA

‡Universidad de La Rioja, Logroño 26004, Spain. Supported by a Spanish MEC grant and by a Comunidad de Madrid-Univ. Autónoma Madrid grant. Email: adrian.ubis@gmail.com

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*

$$\max_{A \subset \mathbb{F}_p,\ |A|=k} \#\{A+B:\ B \subset \mathbb{F}_p\} = \mu_k^{p+o(p)}. \tag{1}$$

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

$$\sqrt{2} + \frac{1}{3^k} \leq \mu_k \leq \sqrt{2} + O\left(\sqrt{\frac{\log k}{k}}\right). \tag{2}$$

Moreover $\mu_k \leq (5^5/2^2 3^3)^{1/5} = 1.960131704\ldots$ for all $k \geq 2$, so that if $|A| \geq 2$ then

$$\#\{A+B:\ B \subset \mathbb{F}_p\} \leq (1.9602)^{p+o(p)}.$$

Remark: With a more involved method the constant 1.9602 in the last bound can be improved to 1.9184 (see [Ubi]).

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 μ_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 $B \subset G$; 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 μ_A such that $S(A, \{1, 2, \ldots, N\}) = \mu_A^{N+O(L)}$. Moreover*

$$\mu_k = \sup_{\substack{A \subset \mathbb{Z}_{\geq 0} \\ |A|=k}} \mu_A\ .$$

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

$$S_{k,\ell}(G) = \#\{A+B:\ A, B \subset G, |A| = k, |B| = \ell\}.$$

for any integers $k, \ell > 1$. By Theorem 1 we know that if $k, \ell \to \infty$ as $p \to \infty$ then $S_{k,\ell}(\mathbb{F}_p) \leq 2^{p/2+o(p)}$. We wish to determine for which values of k and ℓ we have that $S_{k,\ell}(\mathbb{F}_p) \geq 2^{p/2+o(p)}$. The Cauchy-Davenport Theorem [Cau] says that for any $A, B \subset \mathbb{F}_p$ we have $|A+B| \geq \min(p, |A| + |B| - 1)$, hence $S_{k,\ell}(\mathbb{F}_p) = p^{O(1)}$ whenever $k + \ell > p - O(1)$. Let us see what we can say otherwise

Theorem 4. *Let $\phi = \frac{1+\sqrt{5}}{2}$ and let $\psi(x)$ be any function for which $\psi(x) \to \infty$ as $x \to \infty$.*

(i) If $k+\ell \le \sqrt{p}$ then $S_{k,\ell}(\mathbb{F}_p) \gg \binom{[p/2]}{k+\ell-2}/\sqrt{\min\{k,\ell\}}$

If $k+\ell \le p/2\phi$ then $S_{k,\ell}(\mathbb{F}_p) \ge p^{O(1)}\binom{[p/2]}{k+\ell}$

If $\phi p/3 + O(1) > k+\ell > p/2\phi$ then $S_{k,\ell}(\mathbb{F}_p) \gg \phi^{p-k-\ell}/p$.

If $p \ge k+\ell \ge \phi p/3 + O(1)$ then $S_{k,\ell}(\mathbb{F}_p) \gg p\phi^{p-k-\ell}/(p+1-k-\ell)$.

In summary, if $k+\ell \le p$ then $S_{k,\ell}(\mathbb{F}_p) \ge p^{O(1)} \max_h \binom{[(p-h)/2]}{k+\ell-h}$

(ii) For any integers with $k,\ell \ge \psi(p)$ and $p-k-\ell \gg p$, we have $S_{k,\ell}(\mathbb{F}_p) \ll \binom{x}{k+\ell}^{1+o(1)}$ with x such that $2^{p-x} \sim \binom{x}{k+\ell}$.

In particular, if $k,\ell \ge \psi(p)$ then

$$S_{k,\ell}(\mathbb{F}_p) = 2^{p/2+o(p)} \tag{3}$$

if and only if $k+\ell \sim p/4$.

Note that Theorem 4(ii) cannot hold for $k+\ell$ very close to p by the last estimate in Theorem 4(i). In Theorem 4 the upper and lower bounds are different in general; we guess that our lower bounds are likely to be nearer to the true size of $S_{k,l}(\mathbb{F}_p)$.

Following the results in this paper, one is led naturally to several open problems: Give an asymptotic formula for the number of sumsets $A+B$ with $|A|, |B| \ge k$, particularly as $k \to \infty$, as well as for the number of sumsets $A+A$. What sets have $> 2^{cp}$ representations as $A+B$? Find an efficient algorithm to determine whether a given set is a sumset $A+B$ with both A, B large.

2 Lower bounds

For a given integer k let

$$A = \{0, [(p-k)/2]+1, [(p-k)/2]+2, \dots, [(p-k)/2]+k-1\}.$$

For *any* subset B of $\{0,1,2,\dots,[(p-k)/2]\}$, we see that $A+B \subset [0,p-1]$ and

$$B = (A+B) \cap \{0,1,\dots 2,[(p-k)/2]\},$$

and thus the sets $A+B$ are all distinct. Hence there are at least $2^{[(p-k)/2]+1} \ge 2^{(p-k)/2}$ 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 $\mu_k \ge \sqrt{2}$ in Theorem 3.

$$\text{Let} \quad A = \{0\} \cup [x+1,\dots,x+k-u-1] \cup (x+k-u-1+A_1)$$

$$\text{and} \quad B = B_1 \cup [x+1,\dots,x+\ell-v-1] \cup \{x+\ell-v-1+y\}$$

where $A_1 \subset [1,y]$ with $|A_1| = u$, and $B_1 \subset [1,x]$ with $|B_1| = v$, where $u < k, y$, and $v < \ell, x$. Therefore $B_1 = (A+B) \cap [1,x]$ and $N + A_1 = (A+B) \cap (N + [1,y])$ for $N = 2x+y+k+\ell-u-v-2$. Also $A + B \subset [0, p-1]$ provided $2x + 2y + k + \ell - u - v - 2 < p$. Therefore $S_{k,\ell} \geq \binom{y}{u}\binom{x}{v}$.

If $k+\ell \leq p/2\phi$ then we select $u = k-1$, $v = \ell-1, y = [p(k-1)/2(k+\ell-2)]$, $x = (p-1)/2-y$. This gives $S_{k,\ell} \geq p^{O(1)}\binom{[p/2]}{k+\ell}$ by Stirling's formula; $S_{k,\ell} \gg \binom{[p/2]}{k+\ell-2}/\sqrt{\min\{k,\ell\}}$ if $k + \ell \leq \sqrt{p}$.

If $k + \ell > p/2\phi$ then we select $u = [k(p+1-k-\ell)/\sqrt{5}(k+\ell)]$, $v = [\ell(p+1-k-\ell)/\sqrt{5}(k+\ell)], y = [\phi u]$, $x = [\phi v]$ to obtain $S_{k,\ell} \gg \phi^{p-k-\ell}/(p+1-k-\ell)$ by Stirling's formula. If $k + \ell \geq \phi p/3 + O(1)$ then we change the above construction slightly: If instead we take $B_1 \subset [0, x-1]$ then there is a unique block of $\geq k+\ell-u-v-3$ consecutive integers in $A+B$ starting with $2x + 2$. Now we can also consider the sums $(r + A) + B$, for any $r \pmod p$; notice that we can identify the value of r from $A + B$, since the longest block of consecutive integers in $A + B$ starts with $2x + 2 + r$. Hence $S_{k,\ell} \gg p\phi^{p-k-\ell}/(p + 1 - k - \ell)$.

These last three paragraphs together imply the first part of Theorem 4.

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

3 First upper bounds

In this section we shall use a combinatorial argument to bound the number of sumsets $A + B$ whenever A is small, in which case we can consider A fixed. Throughout we let $r_{C+A}(n)$ (and $r_{C-A}(n)$) denote the number of representations of n as $c + a$ (respectively, $c - a$) 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 $k \geq 2$. Then*

$$\#\{A + B : \ B \subset G\} \leq n \ \min_{2 \leq \ell \leq k} \ \sum_{j=0}^{n} \binom{n}{[j/\ell]} \ \min\{2^{n-j}, 2^{[jk/(k-\ell+1)]}\}. \tag{4}$$

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_ℓ be the set of b_i such that $A + \{b_1, b_2, \dots, b_i\}$ contains at least ℓ more elements than $A + \{b_1, b_2, \dots, b_{i-1}\}$, and suppose that $|B_\ell + A| = j$. By definition $j = |B_\ell + A| \geq \ell |B_\ell|$, so that $|B_\ell| \leq [j/\ell]$ and so there are no more than $\sum_{i \leq [j/\ell]} \binom{n}{i}$ choices for B_ℓ. Note that $j/\ell \leq n/2$, for $\ell \geq 2$, and so $\sum_{i \leq [j/\ell]} \binom{n}{i} \leq n\binom{n}{[j/\ell]}$. Next we have to determine the number of possibilities for $A + B$ given B_ℓ (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 H for which $B_\ell + A \subset H \subset G$, which equals 2^{n-j}.

Our second argument: Let $C = B_\ell + A$, and let 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 at most

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

Now $\binom{n}{[j/\ell]} 2^{[jk/(k-\ell+1)]}$ is a non-decreasing function of j, as $\ell \geq 2$, and so the above is not greater than

$$n^2 \min_{2 \leq \ell \leq k} \max_{\frac{(k-\ell+1)}{(2k-\ell+1)} n \leq j \leq n} \binom{n}{[j/\ell]} 2^{n-j}.$$

The $(j+\ell)$th term equals the jth term times $(n-[j/\ell])/2^\ell([j/\ell]+1)$. This is smaller than 1 if and only if $n < (2^\ell+1)[j/\ell] + 2^\ell$. Now

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

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

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

using Stirling's formula. A brief Maple calculation yields that $\nu_k > 2$ for all $k \leq 7$ and $\nu_8 = 1.982301294$, $\nu_9 = 1.961945316$, $\nu_{10} = 1.942349376, \ldots$, with $\nu_k < 1.91$ for $k \geq 12$, and ν_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 ℓ so that $\ell^2 \sim k \log k / \log 2$, one gets that

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

which implies the upper bound in (2) of Theorem 3, as well as the upper bound implicit in Theorem 1 when $\min\{|A|, |B|\} = o(p)$.

4 Upper bounds on $S_{k,\ell}(\mathbb{F}_p)$ using combinatorics

The value of x in Theorem 4(ii) must always lie in the range $[p/2, p]$ since $\binom{x}{k+\ell} \le 2^x$. Therefore if $k+\ell = o(p)$ then the number of sumsets $A+B$ is smaller than the number of possibilities for A and B so that

$$S_{k,l}(\mathbb{F}_p) \le \binom{p}{k}\binom{p}{\ell} = \binom{p}{k+\ell}2^{O(k+\ell)} = \binom{x}{k+\ell}2^{O(k+\ell)} = \binom{x}{k+\ell}^{1+o(1)}.$$

The Cauchy-Davenport Theorem states that $|A+B| \ge \min\{|A|+|B|-1, p\}$, so that

$$S_{k,l}(\mathbb{F}_p) \le \sum_{j=k+\ell-1}^{p} \binom{p}{j} \ll \binom{p}{k+\ell-1}$$

for $k+\ell > (1/2+\epsilon)p$. For the last part of Theorem 4, note that this is $< 2^{8p/17}$ for $k+\ell \ge 9p/10$.

Now we consider the case $\ell, p-k-\ell \gg p$ with $k < o(p)$ and $k \to \infty$. For each fixed A of cardinality k and B of cardinality ℓ, we proceed as in Proposition 2 (taking ℓ there as m here, and choosing $m = o(k)$ with $m \to \infty$): Hence there exists a subset $B_m \subset B$ with $|A+B_m| = j$ and $|B_m| \le j/m \le p/m$, and a subset D, determined by A and B_m, with $|D| \le \frac{kj}{k+1-m} \le j(1+O(m/k))$ and $B \setminus B_m \subset D$. Now $A+B = (A+B_m) \cup (A+(B \setminus B_m))$ so the number of possibilities for $(A+B)\setminus(A+B_m)$ is bounded above by the number of subsets of $\mathbb{F}_p \setminus (A+B_m)$, which is 2^{p-j}, and also by the number of subsets of D with cardinality in the range $[\ell - [j/m], \ell]$, which is at most

$$\sum_{i=\ell-[j/m]}^{\ell} \binom{|D|}{i} \le 2^{o(p)}\binom{j}{\ell+k},$$

since $|D| \le j + o(p)$ and $i = \ell + k + o(p)$. Hence the number of possible sumsets $A+B$ is bounded by $\binom{p}{k} \le 2^{o(p)}$, the number of possibilities for A, times $\sum_{i \le [p/m]}\binom{p}{i} \le 2^{o(p)}$, the number of possibilities for B_m, times $2^{o(p)}\min\{\binom{j}{\ell+k}, 2^{p-j}\}$, the number of possibilities for $(A+B)\setminus(A+B_m)$. This gives us the upper bound

$$S_{k,\ell}(\mathbb{F}_p) \le 2^{o(p)}\min\{\binom{j}{\ell+k}, 2^{p-j}\} = 2^{(1+o(1))(p-x)} \tag{5}$$

where x is chosen as in Theorem 4(ii), noting that $p-x \gg p$ as $\ell+k \gg p$.

5 Sumsets from big sets

this section we adapt to our problem the method of q‘granular sets”of Green and Ruzsa [GrRu]. e adaptation is not difficult, t for accuracy and completeness we present a sketch of the thod, including several simple modifications of the original guments, that arify the

ideas and slightly improve the bounds. ffalse this section we adapt to our problem Green and Ruzsa's method of q'granular sets"[GrRu]. his is straightforward, d we could even quote directly from that paper, but we think it is rthwhile for accuracy and completeness to present a sketch of the thod. Moreover we make a couple of small changes to their arguments at, we think, clarify the ideas involved (and even improve the bounds little). i For a given set S, define $dS := \{ds : s \in S\}$. 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 as we can into arithmetic progressions with difference d and length L. That is for $1 \le i \le d$ we have the progressions

$$I_{i,k} := \{i + jd : \ kL \le j \le \min\{(k+1)L - 1, [(m-i)/d]\}\}$$

for $0 \le k \le [(m-i)/Ld]$. 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 $\le [m/L] + d$ 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*

$$m > (4L)^{1+16\alpha\beta L^4 \epsilon_2^{-2}\epsilon_3^{-1}}, \quad \text{with } L \ge 3, \tag{6}$$

then there exists an integer d, with $1 \le d \le m/4L$, 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.

In this paragraph we follow the proof of Proposition 3 in [GrRu] (with the obvious modifications):

Lemma 1. *If $A \subset \mathbb{Z}/m\mathbb{Z}$ then there exists $1 \le d \le m/4L$ such that*

$$|\hat{A}(x)|^2 \left| 1 - \left(\frac{\hat{H}(dx)}{2L-1} \right)^2 \right|^2 \le \frac{\log 4L}{\log(m/4L)} \ |A|m, \quad \text{with } L \ge 3,$$

for all $x \in \mathbb{Z}/m\mathbb{Z}$.

Proof. (Sketch) Fix δ so that the right side above equals $(\delta m)^2$, and hence $\delta \ge 2/m$. Let R be the set of $r \in \mathbb{Z}/m\mathbb{Z}$ such that $|\hat{A}(r)| \ge \delta m$; the result follows immediately for any $x \notin R$. By Parseval's formula $\sum_x |\hat{A}(x)|^2 = m|A|$ we have the bound $|R| \le \delta^{-2}|A|/m$. Moreover, by the arithmetic-geometric mean inequality and Parseval's formula, we have

$$\prod_{r\in R} |\hat{A}(r)|^2 \le \left(\frac{1}{|R|} \sum_{r\in R} |\hat{A}(r)|^2 \right)^{|R|} \le \left(\frac{m|A|}{|R|} \right)^{|R|}. \tag{7}$$

Consider the vectors $v_i \in [0,1)^{|R|}$ with rth coordinate $ri/m \pmod 1$ for each $r \in R$. If we partition the unit interval for the rth coordinate into intervals of roughly equal length, all not

greater than $(\delta m)^{1/2}/(4L-1)|\hat{A}(r)|^{1/2}$, then, by the pigeonhole principle, two such vectors, with $0 \le i < j \le m/4L$, lie in the same intervals since

$$\prod_{r\in R}\left(1+(4L-1)\left|\frac{\hat{A}(r)}{\delta m}\right|^{1/2}\right) \le \prod_{r\in R} 4L\left|\frac{\hat{A}(r)}{\delta m}\right|^{1/2} \le \left(4L\left(\frac{|A|/m}{\delta^2|R|}\right)^{1/4}\right)^{|R|} \le \frac{m}{4L},$$

using (7), the previous bound for $|R|$ and the definition of δ. Therefore for $d = j - i$ we have

$$\left\|\frac{dr}{m}\right\| \le \frac{1}{4L-1}\left(\frac{\delta m}{|\hat{A}(r)|}\right)^{1/2}$$

for all $r \in R$, where $\|t\|$ is the shortest distance from t to an integer. Now $\mathrm{Re}(1-e(t)) \le 2\pi^2\|t\|^2$ and $\|jt\| \le |j|\|t\|$, and the result follows from $1 + \hat{H}(dr)/(2L-1) \le 2$ and

$$1 - \frac{\hat{H}(dr)}{2L-1} = \frac{1}{2L-1}\sum_{j=-(L-1)}^{L-1}\left(1-e\left(\frac{jdr}{m}\right)\right) \le \frac{2\pi^2L^2}{3}\left\|\frac{dr}{m}\right\|^2 \le \frac{\delta m}{2|\hat{A}(r)|}.$$

□

Proof of Proposition 3. By Parseval's formula, and then Lemma 1 we have

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

$$\le \frac{\log 4L}{\log(m/4L)}\ |A|\sum_x |\hat{B}(x)|^2 = \frac{\log 4L}{\log(m/4L)}\ |A||B|m \le \frac{\epsilon_2^2\epsilon_3 m^3}{16L^4} \tag{8}$$

in this range for m. (Here $r_{A+dH+B+dH}(n)$ denotes the number of representations of n as $a+di+b+dj$ with $a \in A,\ b \in B$ and $i, j \in H$.) Now if $g \in A_{L,d}$ then there exists $j \in H$ such that $g + dj \in A$, by definition, and hence $r_{A+dH}(g) \ge 1$. Hence $r_{A+dH}(g) \ge r_{A_{L,d}}(g)$ for all $g \in G$, so that $r_{A+dH+B+dH}(n) \ge r_{A_{L,d}+B_{L,d}}(n)$ 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 $r_{A+dH+B+dH}(n) > \epsilon_2 m$, $r_{A+B}(n) = 0$ and (8) yields $|N| \le \epsilon_3 m$. □

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 $m \le r \le \min(|C|,|D|)$, there are at least $\min(|C|+|D|,p) - r - (m-1)p/r$ values of $n \pmod p$ such that $r_{C+D}(n) \ge m$.*

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

$$\sum_n \min\{r, r_{C+D}(n)\} \ge r\min(p, |C|+|D|-r) \ge r[\min(p,|C|+|D|) - r].$$

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

Proof of upper bounds on $S_{k,\ell}(\mathbb{F}_p)$ using Fourier analysis:

Suppose that L is given and $d \le p/4L$, and that M and N are unions of some of the arithmetic progressions $I_{i,j}$. Note that there are $\le 2^{p/L+d}$ such sets M (given d), and hence a total of $e^{O(p/L)}$ 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 $\binom{|M|}{k}\binom{|N|}{\ell} \le \binom{|M|+|N|}{k+\ell} \le 2^{|M|+|N|}$ such pairs.

Second, select $2\epsilon_1 p \le \min(|M|,|N|)$ and $2\epsilon_3 p \le \max(|M|,|N|)$. Let Q be the values of $n \pmod p$ such that $r_{M+N}(n) \ge \epsilon_1^2 p$. Taking $r = \epsilon_1 p$ and $m = \epsilon_1^2 p$ in Proposition 4, we have $|Q| \ge R := \min(|M|+|N|, p) - 2\epsilon_1 p$. 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

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

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

If $|M|+|N| \le p/2$ then the number of sumsets is $\le 2^{p/2}$ by the first argument. Let $L = [(\log p)^{1/10}]$ and $\epsilon_1 = \epsilon_3 = 1/2L$. If $|A|, |B| > p/L$ then $|M| \ge |A| > 2\epsilon_1 p$ and $|N| \ge |B| > 2\epsilon_1 p$, so the second argument is applicable; therefore if $|M|+|N| > p/2$ then the number of sumsets is $\le 2^{p/2}L^{O(p/L)}$. Hence the total number of sumsets $A+B$ with $|A|, |B| > p/L$ is at most $2^{p/2}L^{O(p/L)}$ 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).

Assume that $\ell \ge k \ge p/(\log p)^{1/4}$ with $p-k-\ell \gg p$. We select $\epsilon_1 = k/2p\log\log p$, $\epsilon_3 = \ell/2p\log\log p$ and $L = [(\log p)^{1/20}]$, so that the second argument above is applicable. Taking $x = |M|+|N|$ we have that

$$S_{k,\ell}(\mathbb{F}_p) \le \max_{0\le x'\le 2p} \min\left\{\binom{x'}{k+\ell}, 2^{\max(p-x',0)}\right\} (1/\epsilon_3)^{O(\epsilon_3 p)} = 2^{(1+o(1))(p-x)}2^{o(p)}$$

as in (5). This completes the proof of Theorem 4(ii), combined with the results of the previous section.

Finally, (3) follows noting that $x \gtrsim p/2$ unless $k+\ell \sim p/4$, in which case $x \sim p/2$.

6 Sumsets in finite fields and the integers

Let $A \subset \mathbb{F}_p$ be of given size $k \ge 2$, and let $d = [p^{1-1/k}]$. Consider the sets iA, the least residues of $ia, a \in A$, for $0 \le i \le p-1$. 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 $\le d$ 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 is a set of integers in $[0, L]$ where $L \le 2p^{1-1/k}$.

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 $S(A, \{1, 2, \dots, p\})$. When we reduce $A + B$, where $A \subset \{0, \dots, L\}$ and $B \subset \{0, \dots, p - 1\}$ are sets of integers, modulo p, the reduction only affects the residues in $\{0, \dots, L - 1\} \pmod p$. Hence

$$S(A, \{1, 2, \dots, p\})2^{-L} \le S(A, \mathbb{F}_p) \le S(A, \{1, 2, \dots, p\}). \tag{9}$$

Now suppose $A \subset \{0, \dots, L\}$ is a set of integers. Suppose that $Mr \le N < M(r + 1)$ for positive integers M, r, N. We see that

$$S(A, \{1, 2, \dots, N\}) \le S(A, \{1, 2, \dots, M(r + 1)\}) \le S(A, \{1, 2, \dots, M\})^{r+1},$$

the last inequality coming since the sumsets $A + B$ with $B \subset \{1, 2, \dots, M(r + 1)\}$ are the union of the sumsets $A + B_i$ with $B_i \subset \{Mi + 1, 2, \dots, M(i + 1)\}$ for $i = 0, 1, 2, \dots, r$. In particular for $m_A(N) := S(A, \{1, 2, \dots, N\})^{1/N}$ we have $m_A(N) \le m_A(M)^{1+1/r}$. This implies that $\limsup_N m_A(N) \le m_A(M)$ for any fixed M, and then $\limsup_N m_A(N) = \liminf_N m_A(N)$ so the limit, say μ_A, exists and satisfies

$$S(A, \{1, 2, \dots, M\}) \ge \mu_A^M. \tag{10}$$

In the other direction we note that if $B = \cup_i B_i$ where $B_i \subset \{(M + L)i + 1, (M + L)i + 2, \dots, (M + L)i + M\}$ then distinct $\{A + B_i\}_{0 \le i \le r-1}$ give rise to distinct $A + B$. Hence $S(A, \{1, 2, \dots, M\})^r \le S(A, \{1, 2, \dots, r(M + L)\})$ and letting $r \to \infty$ we have

$$S(A, \{1, 2, \dots, M\}) \le \mu_A^{M+L}. \tag{11}$$

Finally, by the inequalities (9), (10) and (11) we arrive at

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

This proves Proposition 1, as well as the first part of Theorem 3.

6.1 Precise bounds when $k = 2$

By the previous section we know that $\mu_2 = \mu_{\{0,1\}}$. Now S is a sumset 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 $n \notin S$, 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. One can check that $C_{n+1} = 2C_n - C_{n-1} + C_{n-2}$ so that $C_n \sim c\mu_2^n$ for some constant $c > 0$, where μ_2 is as in Theorem 3, implying a strong form of the first part of Theorem 3 for $k = 2$.

By a more precise analysis we could even estimate the number of sets $C = \{0, 1\} + B$ with either C or B of given size.

6.2 Precise bounds when $k = 3$

It is not hard to generalize the procedure for the case $|A| = 2$ to any $A \subset \mathbb{Z}$ finite, namely to prove that μ_A is the root of a polynomial with integer coefficients (and degree smaller than 2^{2L+1} when $A \subset \{0, 1, \dots, L\}$), and also that $\mu_{cA+d} = \mu_A$ for any $c, d \in \mathbb{Z}$, $c \neq 0$.

Therefore, in the special case of three elements is enough to deal with $A = \{0, a, b\}$ for a, b coprime positive integers. By a suitable bijection, one can show that $\mu_{\{0,a,b\}} \to \mu_*$ as $a + b \to \infty$ (with $(a, b) = 1$), 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 one can prove exists, and is $< \mu_2$. Therefore either $\mu_3 = \mu_{\{0,a,b\}}$ for some a and b or $\mu_3 = \mu_*$. 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. All this is detailed in Chapter 3 of the third author's PhD. thesis [Ubi].

6.3 Lower bounds on μ_k

That $\mu_k \geq \sqrt{2}$ follows by choosing $A = 1 \cup 2A'$ with $A' \subset \mathbb{Z}$ any finite set. Let $A_k = \{1, 3, \dots, 3^{k-1}\}$, and write $B \subset \{1, 2, \dots, 3n\}$ as $B = 3B_0 \cup (3B_1 - 1)$ with $B_0, B_1 \subset \{1, 2, \dots, n\}$. Since $A_{k+1} = 1 \cup 3A_k$ we have

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

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

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

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

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

Let A be any set of given size $k \geq 2$ in $\mathbb{F}_p$. 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 $\leq 2p\binom{p}{[2p/5]}$, the third requires some work: We observe that such C must have at least $2p/5$ pairs of consecutive elements; so if c is the smallest integer ≥ 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 $i_k = |I_k|$ and $j_k = |J_k|$, 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

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

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

(Note that $\binom{a}{b}\binom{c}{d} \leq \binom{a+c}{b+d}$ follows from defining $\binom{a}{b}$ to be the number of ways of choosing b elements from a.) Hence the number of sumsets $A + B$ is $\ll p^3\binom{p}{[2p/5]} = c^{p+o(p)}$ where $c = (5^5/2^2 3^3)^{1/5} = 1.960131704\ldots$ 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

[Cau] A. Cauchy. Recherches sur les nombres. *J. École Polytech*, 9 99–116, 1813.

[GrRu] B. Green, I. Z. Ruzsa. Counting sumsets and sum-free sets modulo a prime. *Studia Sci. Math. Hungar.*, 41 285–293, 2004.

[Pol] J. M. Pollard. A generalisation of the theorem of Cauchy and Davenport. *J. London Math. Soc. (2)*, 8 460–462, 1974.

[Ubi] A. Ubis. *Cuestiones de la Aritmética y del Análisis Armónico.* PhD. thesis, Universidad Autónoma de Madrid, Madrid, 2006 (English translation available at the web address `http://www.uam.es/gruposinv/ntatuam/downloads/phdubis.pdf`).