Sum-Product Formulae

Andrew Granville and József Solymosi

No Institute Given

Summary. This is a survey on sum-product formulae and methods. We state old and new results. Our main objective was to introduce the basic techniques used to bound the size of the product and sum sets of finite subsets of a field.

1 Introduction

1.1 A Few Definitions

We define

$$A+B = \{g \in G : \text{ There exist } a \in A, b \in B \text{ such that } g = a+b\}; \text{ and}$$
$$A \cdot B = \{ab : a \in A, b \in B\}.$$

We let $r_{A+B}(n) := \#\{a \in A, b \in B : \ n = a+b\}$, $r_{AB}(n) := \#\{a \in A, b \in B : \ n = ab\}$, and note that $0 \le r_{A+B}(n) \le \min\{|A|, |B|\}$ since $r_{A+B}(n) = |A \cap (n-B)| \le |A|$ and $r_{A+B}(n) = |B \cap (n-A)| \le |B|$. We write $\hat{A}(t) = \sum_{a \in A} e(at)$ where $e(u) = e^{2i\pi u}$.

1.2 Multiplication Tables

We learnt to multiply by memorizing the multiplication tables; that is, we wrote down a table with the rows and columns indexed by the integers between 1 and N and the entries in the table were the row entry times the column entry.[1] Paul Erdős presumably learnt his multiplication tables rather more rapidly than the other students, and was left wondering: How many distinct integers are there in the N-by-N multiplication table? Note that if we take $A = \{1, 2, \dots, N\}$, then we are asking how big is $A \cdot A$? Or, more specifically, since the numbers in the N-by-N multiplication table are all $\le N^2$, what proportion of the integers up to N^2 actually appear in the table? That is,

$$\textit{Does } |A \cdot A|/N^2 \textit{ tend to a limit as } N \to \infty?$$

Erdős showed that the answer is, yes, and that the limit is 0. His proof comes straight from "*The Book*".[2] Erdős's proof is based on the celebrated result of Hardy and Ramanujan that "almost all" positive integers $n \le N$ have $\sim \log\log N$ (not necessarily distinct) prime factors (here "almost all" means for all but $o(N)$ values of $n \le N$): Hardy and Ramanujan's result

[1] A.G.: In my primary school we took $n = 12$ which was the basic multiple needed for understanding U.K. currency at that time.

[2] Erdős claimed that the Supreme Being kept a book of all the best proofs, and only occasionally would allow any mortal to glimpse at "The Book".

implies that "almost all" products ab with $a,b \le N$ have $\sim 2\log\log N$ prime factors, whereas "almost all" integers $\le N^2$ have $\sim \log\log(N^2) \sim \log\log N$ prime factors! The result follows from comparing these two statements.

1.3 The Motivating Conjectures

In fact one can show that $|A\cdot A|$ is large whenever A is an arithmetic progression or, more generally, when A is a generalized arithmetic progression of not-too-large dimension. [3]

This led Erdős and Szemerédi to the conjecture that for any $\varepsilon > 0$, there exists $c_\varepsilon > 0$ such that

$$|A+A| + |A\cdot A| \ge c_\varepsilon |A|^{2-\varepsilon}. \tag{1}$$

Even more, the second author conjectured that if $|A| = |B| = |C|$ then

$$|A+B| + |A\cdot C| \ge c_\varepsilon |A|^{2-\varepsilon}. \tag{2}$$

Perhaps the most general version is

$$\textit{Either } |A+B| \gg (|A||B|)^{1-\varepsilon} \textit{ or } |A\cdot C| \gg (|A||C|)^{1-\varepsilon}$$

with no restrictions on the sizes of A,B and C. The thinking in these conjectures is that if $A+B$ is small then A must be "structured", more precisely that it must look like a largish subset of a generalized arithmetic progression, and similarly if AC is small then $\log A$ must look like a largish subset of a generalized arithmetic progression, and that these two structures are incompatible.

2 Sum-Product for Real Numbers

2.1 Results Via Discrete Geometry

The second author proved (2) for $\varepsilon = 8/11$ [27] (see Theorem 1 below). We now prove (2) for $\varepsilon = 3/4$. We begin by stating the

Szemerédi–Trotter Theorem. *We are given a set $\mathscr{C}$ of m curves in $\mathbb{R}^2$ such that*

- *Each pair of curves meet in $\le \kappa_1$ points;*
- *Any pair of points lie on $\le \kappa_2$ curves.*

For any given set $\mathscr{P}$ of n points, there are $\le m + 4\kappa_2 n + 4\kappa_1\kappa_2^{1/3}(mn)^{2/3}$ pairs (π,γ) with point $\pi \in \mathscr{P}$ lying on curve $\gamma \in \mathscr{C}$.

Székely provided a gorgeous proof of this result, straight from *The Book*, via geometric and random graph theory. From this Elekes elegantly deduced the following:

---

[3] A *generalized arithmetic progression* is the image of a lattice, that is:

$$C := \{a_0 + a_1 n_1 + a_2 n_2 + \cdots + a_k n_k : \ 0 \le n_j \le N_j - 1 \text{ for } 1 \le j \le k\},$$

where $N_1, N_2, \ldots, N_k$ are integers ≥ 2. This generalized arithmetic progression is said to have *dimension* k and *volume* $N_1 N_2 \ldots N_k$; and is *proper* if its elements are distinct.

Theorem 1. *If $A,B,C \subset \mathbb{Z}$ then*

$$|A+B|+|A\cdot C| \geq \frac{1}{2}(|A|-1)^{3/4}(|B||C|)^{1/4}. \tag{3}$$

Proof. If $|A+B||A\cdot C| \geq (\frac{1}{2^4}|B||C|)^2$ then at least one of $|A+B|$ and $|A\cdot C|$ is $\geq \frac{1}{2^4}|B||C|$, and they are both $\geq |A|$, so that their product is $\geq \frac{1}{2^4}|A|^3|B||C|$. Hence $|A+B|+|A\cdot C| \geq \frac{1}{2}|A|^{3/4}(|B||C|)^{1/4}$, which implies the result. Hence we may assume that $|A+B||A\cdot C| < \frac{1}{2^8}(|B||C|)^2$.

Let $\mathscr{P}$ be the set of points $(A+B)\times(A\cdot C)$; and $\mathscr{C}$ the set of lines $y=c(x-b)$ where $b\in B$ and $c\in C$. In the Szemerédi–Trotter Theorem we have $\kappa_1=\kappa_2=1$ with

$$m=|B||C| \text{ and } n\leq N:=|A+B|\,|A\cdot C|,$$

since the set of points is $\cup_{a\in A}(a+B,aC)$. For fixed $b\in B$ and $c\in C$, all of the points $\{(a+b,ac):a\in A\}$ in $\mathscr{P}$ lie on the line $y=c(x-b)$, so that

$$\#\{(\pi,\gamma):\pi\in\mathscr{P} \text{ on } \gamma\in\mathscr{C}\}\geq |A|m.$$

Substituting this into the Szemerédi–Trotter Theorem we obtain

$$(|A|-1)m\leq 4n+4(mn)^{2/3}\leq 4N+4(mN)^{2/3}.$$

We assumed that $N<m^2/2^8$, so that $N<(mN)^{2/3}/2^{8/3}<(2^{2/3}-1)(mN)^{2/3}$, and hence $(|A|-1)m^{1/3}\leq 4(2N)^{2/3}$. This implies that $N>(|A|-1)^{3/2}(|B||C|)^{1/2}/16$. □

Corollary 2.1. *If $A\subset\mathbb{Z}$ then*

$$|A+A|+|A\cdot A|\geq\frac{1}{2}|A|^{5/4}.$$

Next we give an argument that improves this. It is still not the best result currently known in the direction of (1) but it uses the Szemerédi–Trotter Theorem only which has several advantages. The most important advantage is that incidence bounds between points and lines on a plane over any field K provide sum-product bounds in K. Even better, the point set where the incidence bounds are needed have a special Cartesian product structure. For example on the complex plane, $\mathbb{C}^2$, it is quite easy to give a Szemerédi-Trotter type bound (with the same exponents) for lines and points of a Cartesian product like $(A+B)\times(A\cdot C)$ above. The incidence bound for this special case appeared in [26]. Another example is Vinh's work [30] who used a Szemerédi-Trotter type bound to obtain a different proof of Garaev's sum-product estimate in finite fields (See Theorem 3 below).

Theorem 2. *[27] If $A,B,C,D\subset\mathbb{Z}$ with $0\notin C$ then*

$$|A/C||A+B||C+D|\gg(|A||C|)^{3/2}(|B||D|)^{1/2}\min\left\{1,\frac{|A/C|}{|B||D|}\right\}^{1/4}.$$

Part of this follows from

Proposition 2.2. *If $A,B,C,D\subset\mathbb{Z}$ with $0\notin C$ and $|A/C|\leq|B||D|$ then*

$$|A/C|^3\,(|A+B||C+D|)^4>\frac{1}{2\cdot 10^{10}}(|A||C|)^6(|B||D|), \tag{4}$$

and

$$|AC|^3(|A+B||C+D|)^4\geq\frac{1}{10^9}\,\frac{(|A||C|)^6(|B||D|)}{\log^3(4\min\{|A|,|C|\})}$$

We note some consequences:

Corollary 2.3. *If $A,C \subset \mathbb{Z}$ with $0 \notin C$ then*

$$|A/C|^3\ |A+C|^8 > \frac{1}{2\cdot 10^{10}}(|A||C|)^7 \tag{5}$$

and

$$|AC|^3\ |A+C|^8 \ge \frac{1}{10^9}\ \frac{(|A||C|)^7}{\log^3(4\min\{|A|,|C|\})}.$$

Hence

$$|AA|^3\ |A+A|^8 \gg \frac{|A|^{14}}{\log^3(4|A|)};$$

in particular if $|A+A| \le \kappa|A|$ then $|AA| \gg \kappa^{-8/3}|A|^2/\log(4|A|)$.

Remark. If $A=\{1,\dots,N\}$ then $|AA| \asymp N^2/(\log N)^\delta(\log\log N)^{3/2}$ for some $\delta = 1-\frac{1+\log\log 2}{\log 2} = 0.08607\ldots$. Hence some power of log in the denominator in this last result is unavoidable.

Proof of Proposition 2.2. Let $V(k)$ denote the set of $m \in C/A$ for which $2^k \le r_{C/A}(m) < 2^{k+1}$, for $k = 0,1,2,\dots$ Consider $\mathscr{C} = \mathscr{C}_k$ the set of lines $y = mx+e$ with $m \in V(k)$ which contain at least one point $(x,y) \in B\times D$. Each $(x,y) \in B\times D$ lies on exactly $|V(k)|$ lines of $\mathscr{C}_k$ (as may be seen by taking $e = b-dm$ for each $m \in V(k)$), so that

$$|V(k)||B||D| \le |\mathscr{C}_k| + 4|B||D| + 4(|\mathscr{C}_k||B||D|)^{2/3}$$

by the Szemerédi–Trotter Theorem. Hence either $|V(k)| \le 14$ or $\frac{11}{15}|V(k)||B||D| \le |\mathscr{C}_k| + 4(|\mathscr{C}_k||B||D|)^{2/3}$ which implies that

$$|\mathscr{C}_k| \ge \min\{\frac{1}{4}|V(k)||B||D|, \frac{1}{27}|V(k)|^{3/2}(|B||D|)^{1/2}\}.$$

Now $|V(k)| \le |C/A| \le |B||D|$ so that $|\mathscr{C}_k| \ge \frac{1}{27}|V(k)|^{3/2}(|B||D|)^{1/2}$.

Now consider the set of points $(A+B)\times(C+D)$. If $y = mx+e$ is a line in $\mathscr{C}_k$ containing the point (b,d) then it also contains the points $(a+b, c+d)$ whenever $c/a = m$ with $a \in A, c \in C$. Hence each such line contains at least 2^k points from $(A+B)\times(C+D)$ and the Szemerédi–Trotter Theorem then yields $2^k|\mathscr{C}_k| \le |\mathscr{C}_k| + 4|A+B||C+D| + 4(|\mathscr{C}_k||A+B||C+D|)^{2/3}$. Hence

$$(2^k-1)|\mathscr{C}_k| \le \max\{80|A+B||C+D|, 4.2\ (|\mathscr{C}_k||A+B||C+D|)^{2/3}\}$$

which implies that

$$|\mathscr{C}_k| \le 80\max\left\{\frac{|A+B||C+D|}{(2^k-1)}, \frac{(|A+B||C+D|)^2}{(2^k-1)^3}\right\} = 80\frac{(|A+B||C+D|)^2}{(2^k-1)^3},$$

where this last inequality follows since $r_{C/A}(m) \le \min\{|A|,|C|\}$ which implies that

$$(2^k-1)^2 < r_{C/A}(m)^2 \le |A||C| \le |A+B||C+D|.$$

Combining the deductions at the end of the last two paragraphs gives

$$|V(k)| \le 6^2 10^{2/3}\frac{(|A+B||C+D|)^{4/3}}{(2^k-1)^2(|B||D|)^{1/3}}. \tag{6}$$

Therefore

$$\sum_{r_{C/A}(m)\geq 2^K} r_{C/A}(m) \leq \sum_{k\geq K}\sum_{m\in V(k)} 2^{k+1} \leq 335\sum_{k\geq K}\frac{2^k}{(2^k-1)^2}\cdot\frac{(|A+B||C+D|)^{4/3}}{(|B||D|)^{1/3}}$$
$$\leq 670\frac{2^K}{(2^K-1)^2}\cdot\frac{(|A+B||C+D|)^{4/3}}{(|B||D|)^{1/3}},$$

and this is $\leq \frac{1}{2}|A||C|$ provided $2^K \geq 671(|A+B||C+D|)^{4/3}/(|A||C|)(|B||D|)^{1/3}$. So select the smallest K for which this holds, so that

$$\frac{1}{2}|A||C| \leq \sum_{r_{C/A}(m)<2^K} r_{C/A}(m) < 2^K|C/A| < 1342|C/A|\,\frac{(|A+B||C+D|)^{4/3}}{(|A||C|)(|B||D|)^{1/3}},$$

and the first result follows.

Let us also note that, by the Cauchy–Schwarz inequality

$$(|A||C|)^2 = \left|\sum_n r_{AC}(n)\right|^2 \leq |AC|\sum_n r_{AC}(n)^2 = |AC|\sum_m r_{C/A}(m)^2 \leq 4|AC|\sum_k V(k)2^{2k}.$$

By (6), and the fact that $2^k \leq r_{C/A}(m) \leq \min\{|A|,|C|\}$ we deduce the second result. □

Proof of Theorem 2 when $|C/A| > |B||D|$. We may assume that

$$|A+B||C+D| \leq (|A||C|)^{3/2}/(118(|B||D|)^{1/2}),$$

else we obtain the result by multiplying through by $|A/C|^{3/4} > (|B||D|)^{3/4}$.

In the proof of Proposition 2.2 we note that if $V(k) > |B||D|$ then

$$V(k) \leq \frac{4|\mathscr{C}_k|}{|B||D|} \leq 320\frac{(|A+B||C+D|)^2}{(2^k-1)^3|B||D|},$$

so we obtain $\sum_{r_{C/A}(m)\leq 2^K} r_{C/A}(m) \geq \frac{1}{2}|A||C|$ provided

$$2^K \gg \max\left\{\frac{(|A+B||C+D|)^{4/3}}{(|A||C|)(|B||D|)^{1/3}},\ \frac{|A+B||C+D|}{(|A||B||C||D|)^{1/2}}\right\} \asymp \frac{|A+B||C+D|}{(|A||B||C||D|)^{1/2}},$$

the last equality following from our assumption, and the result follows as in the proof of Proposition 2.2. □

2.2 Some Easier Ideas

The *multiplicative energy* of two finite sets A,B is defined as

$$E_\times(A,B) = \#\{a_1,a_2\in A,\ b_1,b_2\in B:\ a_1b_1 = a_2b_2\} = \sum_{a,b\in B}|aA\cap bA|.$$

By the Cauchy–Schwarz inequality we have $E_\times(A,B)^2 \leq E_\times(A,A)E_\times(B,B)$. We also can write

$$E_\times(A,B) = \sum_m r_{AB}(m)^2 = \sum_n r_{A/B}(n)^2 = \sum_n r_{A/A}(n)r_{B/B}(n),$$

and hence, by the Cauchy–Schwarz inequality

$$(|A||B|)^2 = \left(\sum_m r_{AB}(m)\right)^2 \le |AB|\sum_m r_{AB}(m)^2 = |AB|E_\times(A,B). \tag{7}$$

Similarly $(|A||B|)^2 \le |A/B|E_\times(A,B)$. Finally, if $A_1 \cap A_2 = \emptyset$ then $r_{(A_1\cup A_2)B}(n) = r_{A_1B}(n) + r_{A_2B}(n)$ and so, by the Cauchy–Schwarz inequality,

$$\begin{aligned} E_\times(A_1\cup A_2,B) &= \sum_m r_{(A_1\cup A_2)B}(m)^2 \\ &\le 2\sum_m (r_{A_1B}(n)^2 + r_{A_2B}(n)^2) = 2(E_\times(A_1,B)+E_\times(A_2,B)). \end{aligned} \tag{8}$$

Proposition 2.4. (Solymosi, [28]) *If A and B are finite sets of real numbers, not containing $\{0\}$ then*

$$E_\times(A,B) \le 12|A+A||B+B|\log(3\min\{|A|,|B|\}), \tag{9}$$

and hence, by (7)*,*

$$|A+A||B+B|\min\{|A/B|,|AB|\} \ge (|A||B|)^2/(12\log(3\min\{|A|,|B|\}).$$

Remarks. Note that there are examples with $0 \in A\cup B$ where this bound cannot hold. For example, if $0 \in B$ then $E_\times(A,B) \ge \#\{a,a' \in A : a0 = 0 = a'0\} = |A|^2$ whereas the bound in Proposition 2.4 is smaller than $|A|^2$ if A and B are both arithmetic progressions with $|B| \ll |A|/\log|A|$.

Note also that this bound is, more-or-less, best possible in any example with $|A+A| \ll |A|$ and $|B+B| \ll |B|$ since, trivially, $E_\times(A,B) \ge |A||B|$.

Proof. We begin by proving this result when A and B are both finite sets of positive real numbers. Let $m := \min\{|A|,|B|\}$. If $m = 1$ then $E_\times(A,B) = \max\{|A|,|B|\}$ and the result is easy, so we may assume $m \ge 2$.

Let $R_{B/A}(\ell) = \{(a,b) \in A\times B : b = \ell a\}$ which has size $r_{B/A}(\ell)$, and note that $r_{B/A}(\ell) \le m$. Let $L_k := \{\ell :\ 2^k \le r_{B/A}(\ell) < 2^{k+1}\}$ and $K = [\log m/\log 2] + 1$. Then

$$\sum_{k=0}^{K-1}\sum_{\ell\in L_k} r_{B/A}(\ell)^2 = \sum_\ell r_{B/A}(\ell)^2 = E_\times(A,B),$$

and

$$\sum_{\substack{k=0\\|L_k|=1}}^{K-1}\sum_{\ell\in L_k} r_{B/A}(\ell)^2 \le \sum_{k=0}^{K-1} 2^{2k+2} < \frac{2^{2K+2}}{3} \le \frac{16m^2}{3}\,.$$

Hence there exists k with $|L_k| \ge 2$ for which

$$\sum_{\ell\in L_k} r_{B/A}(\ell)^2 \ge \frac{1}{K}\left(E_\times(A,B) - \frac{16m^2}{3}\right).$$

Let $L_k = \{\ell_1 < \ell_2 < \ldots < \ell_r\}$ with $r \ge 2$; we claim that the elements of

$$\bigcup_{i=1}^{r-1}(R_{B/A}(\ell_i) + R_{B/A}(\ell_{i+1})) \subset A\times B + A\times B$$

are distinct. For if $i < j$ with $a_i + a_{i+1} = a_j + a_{j+1}$ then

$$\ell_i a_i + \ell_{i+1}a_{i+1} < \ell_{i+1}(a_i + a_{i+1}) \le \ell_j(a_j + a_{j+1}) < \ell_j a_j + \ell_{j+1}a_{j+1};$$

and if $(a+a', \ell_i a + \ell_{i+1}a') = (x,y)$ then a and a' are determined, and so unique. Now, as $A\times B + A\times B = (A+A)\times(B+B)$, we deduce that

$$\begin{aligned} |A+A||B+B| = |A\times B + A\times B| &\ge \sum_{i=1}^{r-1} r_{B/A}(\ell_i) r_{B/A}(\ell_{i+1}) \\ &\ge (r-1)2^{2k} \ge \frac{r}{2}\cdot 2^{2k} \ge \frac{1}{8}\sum_{i=1}^{r} r_{B/A}(\ell_i)^2 \\ &\ge \frac{1}{8K}\left(E_\times(A,B) - \frac{16m^2}{3}\right). \end{aligned}$$

From this we deduce that

$$E_\times(A,B) \le \frac{8\log 2m}{\log 2}\,|A+A||B+B| + \frac{16}{3}\min\{|A|,|B|\}^2,$$

and then (9) follows for $m \ge 2$ after a little calculation, using the fact that $|C+C| \ge 2|C|-1$.

Now if A only has positive real numbers, and $0 \notin B$ then write $B = B_+ \cup B_-$ where $B_\pm = \{b \in B : \pm b > 0\}$. Now $E_\times(A,B_-) = E_\times(A,-B_-)$ by definition so, by (8) and then the case in which we have already proved (9), we have

$$\begin{aligned} E_\times(A,B) &\le 2E_\times(A,B_+) + 2E_\times(A,-B_-) \\ &\le 12|A+A|(|B_+ + B_+| + |B_- + B_-|)\log(3\min\{|A|,|B|\}) \end{aligned}$$

which implies (9), since $B_+ + B_+$ and $B_- + B_-$ are evidently disjoint subsets of $B+B$ (as their elements are of different signs).

Finally, if $0 \notin A \cup B$, then (9) follows similarly from this last result by partitioning A as $A_+ \cup A_-$. □

Remark. We can deduce bounds on $E_\times(A,B)$, when $0 \in A \cup B$, from Proposition 2.4, using the following

If $0 \notin A$ but $0 \in B = B_0 \cup \{0\}$ then, by definition, $E_\times(A,B) = E_\times(A,B_0) + |A|^2$ and $B+B = (B_0+B_0)\cup B$.

If $0 \in A = A_0 \cup \{0\}$ and $0 \in B = B_0 \cup \{0\}$ then $E_\times(A,B) = E_\times(A_0,B_0) + (|A|+|B|-1)^2$ with $A+A = (A_0+A_0)\cup A$ and $B+B = (B_0+B_0)\cup B$.

Corollary 2.5. *If A is any finite set of real numbers then*

$$E_\times(A,B) \le 12|A+A|^2\log(3|A|), \tag{10}$$

and hence, by (7)*,*

$$|A+A|^2\min\{|A/A|,|AA|\} \ge |A|^4/(12\log(3|A|).$$

Proof. If $0 \notin A$ then our bound follows from setting $B = A$ in (9). If $0 \in A$ then we use the information in the previous remark, together with (9), to obtain

$$\begin{aligned} E_\times(A,A) &= E_\times(A_0,A_0) + (2|A|-1)^2 \le 12|A_0+A_0|^2\log(3|A|) + (2|A|-1)^2 \\ &= 12(|A+A|-|A|)^2\log(3|A|) + (2|A|-1)^2 \\ &= 12|A+A|^2\log(3|A|) + 12\log(3|A|)(|A|^2 - 2|A||A+A|) + (2|A|-1)^2 \\ &\le 12|A+A|^2\log(3|A|) + 12\log(3|A|)(2|A| - 3|A|^2) + (2|A|-1)^2 \end{aligned}$$

as $|A+A| \ge 2|A|-1$, which yields (10). □

A similar bound for complex numbers was obtained by Konyagin and Rudnev in [21]. Very recently Konyagin and Shkredov announced an improvement on the sum-product bound. They proved in [22] that

$$|A+A|+|A\cdot A| \gg |A|^{4/3+c}$$

where $\frac{1}{20598} > c > 0$ is an absolute constant.

2.3 Small Product Sets

From Corollary 2.3 it follows that if the sumset is very small then the product set is almost quadratic. The opposite statement is surprisingly hard to prove. It was Chang's observation [6] that one can use a powerful tool, the Subspace Theorem, to obtain such bound. For the history and more details about the Subspace Theorem we refer to the excellent survey paper of Yuri Bilu [3].

An important variant of the Subspace Theorem was proved by Evertse, Schlickewei and Schmidt [8]. We present the version with the best known bound due to Amoroso and Viada [1].

Theorem 2.6. *Let K be a field of characteristic 0, Γ a subgroup of K^* of rank r, and $a_1, a_2, \dots, a_n \in K^*$. Then the number of solutions of the equation*

$$a_1z_1 + a_2z_2 + \cdots + a_nz_n = 1 \tag{11}$$

with $z_i \in \Gamma$ and no subsum on the left hand side vanishing is at most

$$A(n,r) \le (8n)^{4n^4(n+nr+1)}.$$

We are going to use the following result of Freiman (Lemma 1.14 in [10]).

Proposition 1. *Let $A \subset \mathbb{C}$. If $|AA| \le C|A|$ then A is a subset of a multiplicative subgroup of $\mathbb{C}^*$ of rank at most r, where r is a constant depending on C.*

Theorem 2.7. *Let $A \subset \mathbb{C}$ with $|A| = n$. Suppose $|AA| \le Cn$. Then there is a constant C' depending only on C such that*

$$|A+A| \ge \frac{n^2}{2} + C'n.$$

Proof. We consider solutions of $x_1 + x_2 = x_3 + x_4$ with $x_i \in A$. A solution of this equation corresponds to two pairs of elements from A that give the same element in $A+A$. Let us suppose that $x_1 + x_2 \neq 0$ (there are at most $|A| = n$ solutions of the equation $x_1 + x_2 = 0$ with $x_1, x_2 \in A$.)

First we consider the solutions with $x_4 = 0$. Then by rearranging we get

$$\frac{x_1}{x_3} + \frac{x_2}{x_3} = 1. \tag{12}$$

By Proposition 1 and Theorem 2.6 there are at most $s_1(C)$ solutions of $y_1 + y_2 = 1$ with no subsum vanishing. Each of these gives at most n solutions of (12) since there are n choices for x_3. There are only two solutions of $y_1 + y_2 = 1$ with a vanishing subsum, namely $y_1 = 0$ or $y_2 = 0$, and each of these gives n solutions of (12). So we have a total of $(s_1(C)+2)n$ solutions of (12).

For $x_4 \neq 0$ we get

$$\frac{x_1}{x_4}+\frac{x_2}{x_4}-\frac{x_3}{x_4}=1. \tag{13}$$

Again by Proposition 1 and Theorem 2.6, the number of solutions of this with no vanishing subsum is at most $s_2(C)n$. If we have a vanishing subsum then $x_1=-x_2$ which is a case we excluded earlier or $x_1=x_3$ and then $x_2=x_4$, or $x_2=x_3$ and then $x_1=x_4$. So we get at most $2n^2$ solutions of (13) with a vanishing subsum (these are the $x_1+x_2=x_2+x_1$ identities.)

So, in total, we have at most $2n^2+s(C)n$ solutions of $x_1+x_2=x_3+x_4$ with $x_i\in A$. Suppose $|A+A|=k$ and $A+A=\{\alpha_1,\dots,\alpha_k\}$. We may assume that $\alpha_1=0$. Recall that we ignore sums $a_i+a_j=0$. Let

$$P_i=\{(a,b)\in A\times A : a+b=\alpha_i\},\quad 2\le i\le k.$$

Then

$$\sum_{i=2}^{k}|P_i|\ge n^2-n=n(n-1).$$

Also, a solution of $x_1+x_2=x_3+x_4$ corresponds to picking two values from P_i where $x_1+x_2=\alpha_i$. Thus

$$2n^2+s(C)n\ge\sum_{i=2}^{k}|P_i|^2\ge\frac{1}{k-1}\left(\sum_{i=2}^{k}|P_i|\right)^2\ge\frac{n^2(n-1)^2}{k-1}$$

by the Cauchy-Schwarz inequality. The bound for $k=|A+A|$ follows.

2.4 Upper Bounds in the Sum-Product Inequality

One obvious way to obtain upper bounds is to select A to be a largish subset of $\{1,\dots,x\}$ with lots of multiplicative structure. For example we could let A be the set of integers $\le x$ all of whose prime factors are $\le y$, so that $|A|=\Psi(x,y)$, $|AA|\le\Psi(x^2,y)$ and $|A+A|\le 2x$. Roughly $\Psi(x,y)=x((e+o(1))/u\log u)^u$ when $x=y^u$; so that $|AA|/|A|^2=(1/2+o(1))^{2u}$ and $|A+A|/|A|^2=(u\log u/(e+o(1)))^{2u}/x$. We select u so that these are roughly equal, that is $u=\frac{\log x}{2\log\log x}\left(1+\frac{1+o(1)}{\log\log x}\right)$, and thus $y\asymp(\log x)^2$. Therefore $|A|=x^{1/2}2^{(1+o(1))u}$. Hence we have an infinite family of examples in which, if $|A|=N$ then

$$\max\{|A+A|,|AA|\}\le N^{2-\frac{\log 4+o(1)}{\log\log N}}.$$

We can obtain this result without using any "machinery": Let A be the set of $N:=\binom{\pi(y)+u}{u}$ integers composed of no more than u not necessarily distinct primes factors $\le y$. Then $A\subset[1,y^u]$ so that $|A+A|\le 2y^u$, whereas $|AA|=\binom{\pi(y)+2u}{2u}$. We select $u=[ey^{1/2}/2\log y]$ so that, by Stirling's formula and the prime number theorem,

$$N=\left(\frac{e(y/\log y)}{u}\right)^u e^{O(u/\log y)}=\left(2y^{1/2}\right)^u e^{O(u/\log y)}=(u(\log u)^{O(1)})^u,$$

and therefore $u\sim\log N/\log\log N$. Now, by similar calculations we find that

$$|AA|\text{ and }|A+A|=|A|^2/2^{(2+o(1))u}e^{O(u/\log y)}=N^{2-\frac{\log 4+o(1)}{\log\log N}}.$$

3 Sum-Product Inequalities over Finite Fields

The Szemerédi–Trotter Theorem does not hold over $\mathbb{F}_q$ which renders all of the above results moot in this setting. However such results in finite fields are the most applicable, so we will now pursue this. The first thing to note is that we must modify (1) when the set A is large, hence Garaev conjectured that if $A \subset \mathbb{F}_p$ then

$$|A+A| + |A\cdot A| \gg \min\{|A|^2/\sqrt{p}, \sqrt{p|A|}\}/|A|^{o(1)}. \tag{14}$$

Upper Bounds in $\mathbb{F}_p$ We begin by showing that the lower bound in Garaev's conjecture cannot, in general, be increased:

Proposition 3.1. *For any given integers I,J,N with $1 \le N \le I,J \le p$ and $N \le \lceil IJ/p \rceil$, there exist $A \subset B, C \subset \mathbb{F}_p$, with $|A| = N, |B| = J, |C| = I$ such that*

$$|A+B| < 2|B| \text{ and } |A\cdot C| < 2|C|.$$

In particular, for any given $N, 1 \le N \le p$, there exists $A \subset \mathbb{F}_p$ with $|A| = N$ such that

$$\max\{|A+A|, |A\cdot A|\} \le \min\{|A|^2, 2\sqrt{p|A|}+1\}.$$

Remark. If we have $\max\{|A+A|, |A\cdot A|\} \gg |A|^{2-o(1)}$ for all sets $A \subset \mathbb{F}_p$ of size N, then the second part of Proposition 3.1 implies that $N \ll p^{1/3-o(1)}$.

Proof. Let $C := \{g^1, \dots, g^I\}$ where g is a primitive root mod p, and $A_x := C \cap B_x$ for each $x \in \mathbb{F}_p$ where $B_x := x + \{1, \dots, J\}$. Now

$$\sum_x |A_x| = \sum_{i=1}^{I} \sum_{j=1}^{J} \#\{x \in \mathbb{F}_p : \ x = g^i - j\} = IJ,$$

so that there exists x with $|A_x| \ge IJ/p$. Let A be any subset of A_x of size N, and $B = B_x$. Therefore $A+A \subset A+B \subset B+B = \{2x+2, \dots, 2x+2J\}$, $A\cdot A \subset C\cdot C \subset \{g^2, \dots, g^{2I}\}$ so that $|A+A| \le |A+B| \le |B+B| < 2J$ and $|A\cdot A| \le |A\cdot C| \le |C\cdot C| < 2I$, which completes the proof of the first part. Now, taking $I = J = \lceil \sqrt{pN} \rceil$ we find that $|A+A|, |A\cdot A| \le 2\lceil \sqrt{pN} \rceil - 1$, which implies the second part. □

A Little Cauchying Let us make note of a couple of inequalities, for characteristic functions of sets: By Cauchy we obtain

$$\left(\sum_j \left|A\left(\frac{j}{p}\right)\right| \left|\hat{B}\left(\frac{-j}{p}\right)\right|\right)^2 \le \sum_j \left|A\left(\frac{j}{p}\right)\right|^2 \sum_j \left|B\left(\frac{-j}{p}\right)\right|^2 = p|A|\cdot p|B|. \tag{15}$$

By Cauchy we obtain

$$\begin{aligned}\left|\sum_{a\in A}\sum_{b\in B} e\left(\frac{kab}{p}\right)\right|^2 &\le \sum_{a\in A} 1 \cdot \sum_{a\in A}\left|\sum_{b\in B} e\left(\frac{kab}{p}\right)\right|^2 \\ &\le |A|\cdot \sum_{a\in \mathbb{F}_p}\left|\sum_{b\in B} e\left(\frac{kab}{p}\right)\right|^2 = |A|\cdot p|B|,\end{aligned} \tag{16}$$

by Parseval.

Lower Bounds in $\mathbb{F}_p$

Theorem 3. (Garaev) *If $A,B,C \subset \mathbb{F}_p$ with $0 \notin C$ then*

$$|A+B| \cdot |A \cdot C| \geq \frac{|A|}{4} \cdot \min\left\{ \frac{|A||B||C|}{p}, 2p \right\}.$$

Remark. Taking $I = J = [\sqrt{pN}]$ in Proposition 3.1, we obtain examples with $|A+B|\,|A \cdot C| \leq 4p|A|$. Therefore Theorem 3 is best possible, up to a factor of 8, when $|A||B||C| \geq 2p^2$. In particular for $|A| = |B| = |C| \geq 2p^{2/3}$.

Theorem 3 and its proof remain valid, with suitable modifications, in $\mathbb{F}_p \times \mathbb{F}_p$ (changing both occurrences of p to $q = p^2$ in the lower bound). If we select a set $D \subset \mathbb{F}_p$ such that $|D+D|, |DD| \asymp \min\{|D|^2, p\}$ then taking $A = B = C = D \times \mathbb{F}_p$ we have $|A+B|, |AC| \asymp p\min\{|D|^2, p\} = \min\{|A|^2/p, p^2\}$ so that $|A+A||AA| \asymp \min\{|A|^4/q, q^2\}$. Therefore Theorem 3 is best possible up to a constant factor when $q^{1/2} \leq |A| \leq q^{3/4}$, in this setting.

First by letting $C \to 1/C$ above, and then by taking $A = B = C$ we deduce:

Corollary 3.2. *If $A,B,C \subset \mathbb{F}_p$ with $0 \notin C$ then*

$$|A+B| \cdot |A/C| \geq \frac{|A|}{4} \cdot \min\left\{ \frac{|A||B||C|}{p}, 2p \right\}.$$

If $A \subset \mathbb{F}_p$ with $0 \notin A$ then

$$|A+A| \cdot |AA|,\ |A+A| \cdot |A/A| \geq \frac{|A|}{4} \cdot \min\left\{ \frac{|A|^3}{p}, 2p \right\}.$$

If A is a multiplicative subgroup of $\mathbb{F}_p^*$ then

$$|A+A| \geq \min\left\{ \frac{|A|^3}{4p}, p/2 \right\}.$$

Proof of Theorem 3. For any $a \in A, b \in B, c \in C$ we have a distinct solution to

$$u/c + b = v \tag{17}$$

with $u \in AC, c \in C, b \in B,\ v \in V = A+B$, where $u = ac$ and $v = a+b$. Hence $|A||B||C|$ is no more than the total number of solutions of (3.4), which equals

$$\begin{aligned}
&\sum_{u \in AC} \sum_{c \in C} \sum_{b \in B} \sum_{v \in A+B} \frac{1}{p} \sum_{j=0}^{p-1} e\left(\frac{j(u/c+b-v)}{p} \right) \\
&= \frac{1}{p} \sum_{j=0}^{p-1} \hat{B}\left(\frac{j}{p}\right) \hat{V}\left(\frac{-j}{p}\right) \sum_{u \in AC} \sum_{c \in C} e\left(\frac{ju/c}{p}\right) \\
&\leq \frac{|B||C||A+B||AC|}{p} + \frac{1}{p} \max_{k \neq 0} \left| \sum_{u \in AC} \sum_{c \in C} e\left(\frac{ku/c}{p}\right) \right| \sum_{j \neq 0} \left|\hat{B}\left(\frac{j}{p}\right)\right| \left|\hat{V}\left(\frac{-j}{p}\right)\right| \\
&\leq \frac{|B||C||A+B||AC|}{p} + \sqrt{p|B||C||A+B||AC|}
\end{aligned}$$

by (15) with A replaced by V, and by (16) with A replaced by AC and B replaced by $1/C$, and the result follows. □

Theorem 4. *Suppose that $A,B,C,D \subset \mathbb{F}_p$, and C does not contain* 0.
If $|A+B||C+D||A/C|^2|B||D| \le p^4$ then

$$|A+B||C+D| \ge \frac{(|A||C|)^2|B||D|}{4p^2}.$$

If $|A+B||C+D||A/C|^2|B||D| > p^4$ then

$$|A+B||C+D||A/C| \ge \frac{p}{2}|A||C|.$$

Corollary 3.3. *If $0 \notin A \subset \mathbb{F}_p$ and $|A+A||A/A||A| \le p^2$, then $|A+A| \ge |A|^3/2p$.*
If $|A+A| \le \kappa|A|$ and $|A| \ge p^{2/3}$ then $|AA|, |A/A| \ge p/2\kappa$ (by Corollary 3.2).
If $|A+A| \le \kappa|A|$ and $(2\kappa p)^{1/2} < |A| \le p^{2/3}$ then $|A/A| > p/2\kappa^2$.

Remark. The first part is stronger than Garaev's $|AA||A+A| \ge |A|^4/2p$ in this range.

Proof of Theorem 4. Let us look at solutions to $u-b = m(v-d)$ with $b \in B, d \in D, u \in U = A+B, v \in V = C+D, m \in M = A/C$. For each $(a,b,c,d) \in A \times B \times C \times D$ we have the (distinct) solution $(b,d,u,v,m) = (b,d,a+b,c+d,a/c)$, so there are at least $|A||B||C||D|$ solutions. On the other hand we can give an exact count via the exponential sum

$$\sum_{b,d,u,v,m} \frac{1}{p} \sum_{j=0}^{p-1} e\left(\frac{j(u-b-m(v-d))}{p}\right) = \frac{|B||D||U||V||M| + \text{Error}}{p},$$

where

$$|\text{Error}| \le \max_{i \ne 0} \left| \sum_{m \in M} \hat{D}\left(\frac{im}{p}\right) \hat{V}\left(\frac{-im}{p}\right) \right| \cdot \left| \sum_{j=1}^{p-1} \hat{U}\left(\frac{j}{p}\right) \hat{B}\left(\frac{-j}{p}\right) \right|.$$

By Cauchy–Schwarz this gives

$$|\text{Error}|^2 \le \sum_{m \in M} \left|\hat{D}\left(\frac{im}{p}\right)\right|^2 \cdot \sum_{m \in M} \left|\hat{V}\left(\frac{-im}{p}\right)\right|^2 \cdot \sum_{j=1}^{p-1} \left|\hat{U}\left(\frac{j}{p}\right)\right|^2 \cdot \sum_{j=1}^{p-1} \left|\hat{B}\left(\frac{-j}{p}\right)\right|^2$$

which is $\le p|D|p|V|p|U|p|B|$ by (15). Hence we have proved

$$|A||B||C||D| \le \frac{|B||D||U||V||M|}{p} + p\sqrt{|B||D||U||V|},$$

and the result follows. □

Theorem 5 ([17]). *Suppose that $A,B,C,D \subset \mathbb{F}_p$, and A,C do not contain* 0.
If $|A+B||C+D||B||D| \ll p^3$ then

$$|AC|^2|A+B||C+D| \gg (|A||C|)^2|B||D|/p.$$

If $|A+B||C+D||B||D| \gg p^3$ then

$$|AC||A+B||C+D| \gg p|A||C|.$$

Remark. We claim that these bounds can be obtained trivially if $|A||C|(|B||D|)^2 \ll p^3$: The second case cannot hold since

$$|A||C|(|B||D|)^2 = |A||B||C||D||B||D| \ge |A+B||C+D||B||D| \gg p^3,$$

but then $|AC|^2|A+B||C+D| \ge |A||C|\,|A|^{2/3}|B|^{1/3}\,|C|^{2/3}|D|^{1/3} = (|A||C|)^2|B||D|/p \cdot (p^3/(|A||C|(|B||D|)^2)^{1/3} \gg (|A||C|)^2|B||D|/p$.

Proof. There exists $m \in AC$ such that $r_{AC}(m) \geq |A||C|/|AC|$. Now in the set

$$\{(u,v) \in (A+B)\times(C+D),\ (b,d) \in B\times D :\ (u-b)(v-d) = m\}$$

we evidently have the distinct points $((a+b,c+d),(b,d))$ for every $b \in B, d \in D$ and $a \in A, c \in C$ with $ac = m$; a total of $|B||D|r_{AC}(m) \geq |A||B||C||D|/|AC|$ points. To get an exact count, write $U = A+B$, $V = C+D$, to obtain

$$\sum_{b,d,u,v} \sum_{\substack{r,s \\ rs=m}} \frac{1}{p}\sum_i e\left(\frac{i(u-b-r)}{p}\right)\cdot\frac{1}{p}\sum_j e\left(\frac{j(v-d-s)}{p}\right)$$
$$= \frac{1}{p^2}\sum_{i,j}\hat{U}\left(\frac{i}{p}\right)\hat{B}\left(\frac{-i}{p}\right)\hat{V}\left(\frac{j}{p}\right)\hat{D}\left(\frac{-j}{p}\right)\sum_{\substack{r,s \\ rs=m}} e\left(\frac{-(ir+js)}{p}\right).$$

The $i = j = 0$ term yields $\frac{p-1}{p^2}|U||B||V||D|$, since there are exactly $p-1$ solutions to $rs = m$. If $j = 0$ but $i \neq 0$ then our final sum equals -1, so that the sum over $i \neq 0$ is $\frac{1}{p^2}|V||D|(|U||B| - p|U \cap B|)$. Similarly with $i = 0$. Finally if $i \neq 0$ and $j \neq 0$ then the final term is $\leq 2\sqrt{p}$ in absolute value by a well-known result on Kloosterman sums, and the total contribution is therefore

$$\leq \frac{2}{p^{3/2}}\sum_{i\neq 0}\left|\hat{U}\left(\frac{i}{p}\right)\hat{B}\left(\frac{-i}{p}\right)\right|\sum_{j\neq 0}\left|\hat{V}\left(\frac{j}{p}\right)\hat{D}\left(\frac{-j}{p}\right)\right|,$$

and by Cauchying the square of this is

$$\leq \frac{4}{p^3}\sum_i\left|\hat{U}\left(\frac{i}{p}\right)\right|^2\sum_i\left|\hat{B}\left(\frac{-i}{p}\right)\right|^2\sum_j\left|\hat{V}\left(\frac{j}{p}\right)\right|^2\sum_j\left|\hat{D}\left(\frac{-j}{p}\right)\right|^2 = 4p|U||B||D||V|.$$

Putting this altogether we obtain

$$\frac{|A||B||C||D|}{|AC|} \leq \frac{p+1}{p^2}|U||B||V||D| + 2\sqrt{p|U||B||D||V|},$$

which implies the result. □

Corollary 3.4. *Suppose that* $0 \notin A \subset \mathbb{F}_p$. *If* $|A+A||A| \ll p^{3/2}$ *then*

$$|AA||A+A| \gg |A|^3/\sqrt{p}.$$

If $|A+A||A| \gg p^{3/2}$ *then*

$$|AA||A+A|^2 \gg p|A|^2.$$

This is only non-trivial if $|A| \gg p^{1/2}$.

4 Ruzsa–Plunnecke Type Inequalities

We begin with a key result of Ruzsa:

Proposition 4.1. *If* $X, A_1, \ldots, A_k \subset \mathbb{F}_p$ *then there exists a non-empty* $Y \subset X$ *such that*

$$\frac{|Y+A_1+\ldots+A_k|}{|Y|} \leq \prod_{i=1}^k \frac{|X+A_i|}{|X|}.$$

Corollary 4.2. *If $A,B,C \subset \mathbb{F}_p$ then*

$$|A \pm B| \le \frac{|A+C||B+C|}{|C|}.$$

Proof. We can define an injective map $\phi : (A-B)\times C \to (A+C)\times(B+C)$ as follows, so that the inequality $|A-B| \le |A+C||B+C|/|C|$ holds: If $\lambda \in A-B$ fix $a_\lambda \in A$, $b_\lambda \in B$ such that $\lambda = a_\lambda - b_\lambda$ and then define $\phi(\lambda,c) = (a_\lambda + c, b_\lambda + c)$. The map is injective since if $u = a_\lambda + c$ and $v = b_\lambda + c$ then $\lambda = u - v$ and then $c = u - a_\lambda$.

For the other case take $k = 2, A_1 = A, A_2 = B, X = C$ in Proposition 4.1 to obtain that there exists non-empty $Y \subset C$ such that

$$|A+B| \le |Y+A+B| \le \frac{|A+C|}{|C|}\cdot\frac{|B+C|}{|C|}\cdot|Y| \le \frac{|A+C|}{|C|}\cdot\frac{|B+C|}{|C|}\cdot|C|.$$

□

Corollary 4.3. *If $X, A_1, \dots, A_k \subset \mathbb{F}_p$ then there exists $Z \subset X$ such that $|Z| \ge \frac{1}{2}|X|$ and*

$$\frac{|Z+A_1+\ldots+A_k|}{|Z|} \le 2^k \prod_{i=1}^{k} \frac{|X+A_i|}{|X|}.$$

Proof. By Proposition 4.1 we know that there exists a set $Z \subset X$ for which the inequality holds, so let Z' be the largest subset of X for which this inequality is satisfied and suppose that $|Z'| \le \frac{1}{2}|X|$. Apply Corollary 4.2 with $X' = X \setminus Z'$ in place of X. Noting that $|X'| > |X|/2$, and each $|X'+A_i| \le |X+A_i|$ we deduce that there exists a non-empty $Y \subset X'$ such that $|Y+A_1+\ldots+A_k| < 2^k|Y|\prod_{i=1}^k(|X+A_i|/|X|)$. Now let $Z = Z' \cup Y$ so that $|Z+A_1+\ldots+A_k| \le |Z'+A_1+\ldots+A_k| + |Y+A_1+\ldots+A_k|$, which is $\le 2^k \prod_{i=1}^k(|X+A_i|/|X|)$ times $|Y|+|Z'| = |Z|$, and thus our inequality is satisfied by Z which is larger than Z', contradicting the hypothesis. □

Corollary 4.4. *For any $a,b \in \mathbb{F}_p^*$ and $A,B \subset \mathbb{F}_p$ we have*

$$|aA \pm bB| \le \frac{|A+A||B+B|}{|aA \cap bB|},$$

and

$$|aA \pm bB| \le \frac{|A+A||B+B|}{\max_{n\in\mathbb{F}_p} r_{aA+bB}(n)}.$$

Proof. In Corollary 4.2 replace A by aA, B by bB, and take $C = (x+aA)\cap bB$ for some $x \in \mathbb{F}_p$. We note that $aA + C \subset x + aA + aA$ which has the same size as $A+A$ and, similarly, $bB + C \subset bB + bB$ which has the same size as $B+B$. The first result follows taking $x = 0$. Now $|C| = r_{bB-aA}(x)$, so writing $x = -n$ and changing b to $-b$, we get our second result. □

Corollary 4.5. *For any $a,b \in \mathbb{F}_p^*$ and $A,B \subset \mathbb{F}_p$ we have*

$$|aA \pm bB| \le \frac{|A+B|^2}{|bA \cap aB|},$$

and

$$|aA \pm bB| \le \frac{|A+B|^2}{\max_{n\in\mathbb{F}_p} r_{aB+bA}(n)}.$$

Proof. In Corollary 4.2 now replace A by aA, B by bB, and take $C = (x+bA)\cap aB$ for some $x \in \mathbb{F}_p$. We note that $aA+C \subset aA+aB$ which has the same size as $A+B$ and, similarly, $bB+C \subset x+bB+bA$ which also has the same size as $A+B$. The first result follows taking $x=0$. Now $|C| = r_{aB-bA}(x)$, so changing b to $-b$, we get our second result □

5 Lower Bounds on the Size of $A+tB$

Lemma 5.1. *If $A,B \subset \mathbb{F}_p$ with $|A||B| > p$ then $\frac{A-A}{B-B} = \mathbb{F}_p \cup \{\infty\}$.*

Proof. As $|A||B| > p$ and each of $|A|,|B| \le p$ hence $|A|,|B| > 1$ that is $|A|,|B| \ge 2$. Hence $0 = (a-a)/(b_1-b_2)$ and $\infty = (a_1-a_2)/(b-b)$. If $t \ne 0$ then there are $|A||B| > p$ numbers $a+tb$ so that two must be congruent mod p. Taking their difference implies the result. □

Remark. One might expect that if $A,B \subset \mathbb{F}_p$ with $|A||B| > p$ then $AB+AB = \mathbb{F}_p$. However if $A = B = \{m \pmod p : (m/p) = 1\}$ where p is a prime $\equiv 3 \pmod 4$ then evidently $0 \notin AB+AB$ (and here $|A|,|B| = (p-1)/2$). Hence the best we can hope for is that if $|A||B| > p$ then $AB+AB+AB = \mathbb{F}_p$, and perhaps $AB+AB = \mathbb{F}_p^*$.

Glibichuk [13] proved that if $|A||B| \ge 2p$ then $8AB = \mathbb{F}_p$ (so that if $|A||B| \ge p$ then $8AB = \mathbb{F}_p$, since then $|A+A||B| \ge 2p$ so that $16AB \supseteq 8(A+A)B = \mathbb{F}_p$, unless A is an arithmetic progression, which can be handled).

Let $T = \frac{A-A}{B-B} \setminus \{0,\infty\}$. We are interested in the size of $A+tB$ when $|A|,|B| > 1$. Evidently $|A+tB| \le |A|\,|B|$ with equality if and only if $t \notin T \cup \{0\}$.

Let $R(t) = R_{A,B}(t)$ denote the number of solutions $a,c \in A,\ b,d \in B$ to $a+tb = c+td$. We always have the the "diagonal solutions" where $a=c$ and $b=d$ to $a+tb = c+td$, so that $R(t) \ge |A||B|$. Equality holds, that is $R(t) = |A||B|$, if and only if $t \notin T \cup \{0\}$. Hence

$$|A+tB| = |A|\,|B| \iff t \notin T \cup \{0\} \iff R(t) = |A||B|. \tag{18}$$

There is a link between $|A+tB|$ that holds no matter what, which is given by the Cauchy–Schwarz inequality: Let $r_t(n) = \#\{a \in A,\ b \in B :\ n = a+tb\}$ so that

$$(|A|\,|B|)^2 = \left(\sum_{n \in A+tB} r_t(n)\right)^2 \le |A+tB|\, R_{A,B}(t), \tag{19}$$

since $R_{A,B}(t) = \sum_n r_t(n)^2$.

Proposition 5.2. *For any finite sets A,B,S with $0 \notin S$, there exists $t \in S$ for which*

$$|A+tB| > \frac{1}{2}\,\min\{|S|, |A|\,|B|\}.$$

If $A,B \in \mathbb{F}_p$ then we also have

$$|A+tB| > \frac{1}{2}\,\min\left\{p, \frac{|A|\,|B|\,|S|}{p}\right\}.$$

Proof. Note that $|A+0B| = |A|$ and $R(0) = |A|\,|B|^2$. Since $R(t) \ge |A||B|$ for all t hence

$$\sum_{t\in S}(R(t)-|A||B|) \le \sum_{t\ne 0}(R(t)-|A||B|) = \#\{a,c\in A,\ b,d\in B:\ b\ne d \text{ and } a\ne c\}$$
$$= (|A|^2-|A|)(|B|^2-|B|). \tag{20}$$

Therefore there exists $t\in S$ with

$$R(t) \le \frac{(|A|^2-|A|)(|B|^2-|B|)}{|S|} + |A|\,|B| < 2|A||B| \max\left\{\frac{|A||B|}{|S|},1\right\},$$

whence, by (19),

$$|A|\,|B| \le 2|A+tB|\ \max\left\{\frac{|A||B|}{|S|},1\right\}$$

and so the first result follows.

When we are working mod p, we have

$$R(t) = \sum_n r_t(n)^2 = \frac{1}{p}\sum_{j=0}^{p-1}|\hat{A}(j/p)|^2|\hat{B}(jt/p)|^2 \ge \frac{(|A|\,|B|)^2}{p},$$

taking the $j=0$ term, since every term is non-negative. If $|A||B| > p$ then $R(t) > |A|\,|B|$ and so $T = \mathbb{F}_p\setminus\{0\}$ by (18) giving another proof of the lemma above.

Now, rearranging (20) we obtain

$$\sum_{t\in S}\left(R(t)-\frac{(|A|\,|B|)^2}{p}\right) \le \sum_{t\ne 0}\left(R(t)-\frac{(|A|\,|B|)^2}{p}\right) = p|A||B|\left(1-\frac{|A|}{p}\right)\left(1-\frac{|B|}{p}\right),$$

so there exists $t\in S$ with

$$|A|\,|B| \le 2|A+tB|\ \max\left\{\frac{p}{|S|},\frac{|A|\,|B|}{p}\right\}$$

by (19), and the result follows. □

Corollary 5.3. *Suppose that $|A|,|B|\ge 2$. If $\frac{A-A}{B-B} = \mathbb{F}_p$ then there exists $a_1,a_2\in A, b_1,b_2\in B$ such that*

$$|(a_1-a_2)B+(b_1-b_2)A| > \frac{1}{2}\ \min\{p,|A|\,|B|\}.$$

If $\frac{A-A}{B-B}\ne\mathbb{F}_p$ then there exists $a_1,a_2\in A, b_1,b_2\in B$ such that

$$|(a_1-a_2+b_1-b_2)B+(b_1-b_2)A| = |A|\,|B|.$$

In other words, the elements $(a_1-a_2+b_1-b_2)b+(b_1-b_2)a$ with $a\in A, b\in B$, are distinct.

Proof. For any $t\in T$, there exist $a_1,a_2\in A,\ b_1,b_2\in B,$ for which $(a_2-a_1)+(b_1-b_2)t = 0$. The first case follows immediately from Proposition 5.2 by multiplying through by b_1-b_2.

Suppose that $T\ne\mathbb{F}_p^*$. Now T contains a non-zero element, say t, since A has at least two elements. Moreover we may assume $1\le t < p/2$ since $T = -T$ (as may be seen by swapping a_1 and a_2). Hence there exists $t\in T$ such that $t+1\notin T$. Therefore $|A+(t+1)B| = |A||B|$ by (18) and the result follows by multiplying through by b_1-b_2. □

Lemma 5.4. *Let $I(A,B) := (B-B)A+(A-A)B$.*
(i) *If $t \in \frac{A-A}{B-B}$ then $|I(A,B)| \geq |A+tB|$.*
(ii) *$AB-AB \subset I(A,B)$, so that $|I(A,B)| \geq \min\{p, 2|AB|-1\}$.*

Proof. There exist $a_1,a_2 \in A,\ b_1,b_2 \in B$ for which $(a_2-a_1)+(b_1-b_2)t = 0$. Each element of $S = (b_1-b_2)(A+tB)$ can be written as $(b_1-b_2)a+(b_1-b_2)tb = (b_1-b_2)a+(a_1-a_2)b \in I(A,B)$ and (i) follows. Also if $a_1,a_2 \in A,\ b_1,b_2 \in B$ then $a_1b_1-a_2b_2 = (b_1-b_2)a_1+(a_1-a_2)b_2 \in I(A,B)$, that is $AB-AB \subset I(A,B)$, and (ii) follows from the Cauchy–Davenport theorem. □

Corollary 5.5. *If $|A||B| > p$ then there exists $t \in T$ for which $|A+tB| > p/2$. Hence $|I(A,B)| > p/2$. We can rephrase this as: There exists $b_1,b_2 \in B, a_1,a_2 \in A$ such that $|(b_1-b_2)A+(a_1-a_2)B| > p/2$.*

Proof. By Lemma 5.1 we know that $\frac{A-A}{B-B} = \mathbb{F}_p$. Taking $S = T$ in Proposition 5.2 we deduce that there exists $t \in T$ with $|A+tB| > p/2$. The result then follows from Lemma 5.4. □

Proposition 5.6. *Let $R_k(B)$ be the set of $n \in \mathbb{F}_p$ for which $r_{B/B}(n) \geq k$ for $1 \leq k \leq |B|$; note that $1 \in R_k(B)$. Let $G_k(B)$ be the multiplicative group generated by $R_k(B)$, and then $H_k(B) = G_k(B)\frac{A-A}{B-B}$. There exists $t \in T$ for which $|A+tB| \gg \min\{k|A|, |H_k|\}$.*

Proof. If $H_k(B) = \frac{A-A}{B-B}$ then the result follows from proposition 5.2 (since $|B| \geq k$). Otherwise $\frac{A-A}{B-B} \subsetneq H_k(B)$ so there exists $g \in G_k(B)$ and $t_0 \in T$ such that $gt_0 \notin T$. Now any $g \in G_k(B)$ can be written as $g = n_1n_2\dots n_\ell$ where each $n_j \in R_k(B)$. Define $t_j = n_jt_{j-1}$ for each j, so that $t_0 \in T$ and $t_\ell = gt_0 \notin T$: hence there exists $t = t_{j-1} \in T$ and $n = n_j \in R_k(B)$ such that $nt = t_j \notin T$. But then $|A+ntB| = |A||B|$ by (18); that is the elements of $A+ntB$ are all distinct. Now $r_{B/B}(n) \geq k$ by the definition of $R_k(B)$, and so there are at least k values of $b \in B$ for which nb is also in B, and hence $A+tB$ contains at least $|A|k$ distinct elements. □

Lemma 5.7. *Let $B = B_1 \cup B_2$ be a partition of B where $b_1/b_2 \notin G_k$ for any $b_1 \in B_1, b_2 \in B_2$. Then $|B_1||B_2| \leq (k-1)|B_1B_2|$.*

Proof. If $s \notin B_1/B_2$ then $r_{B_1/B_2}(s) = 0$. If $s \in B_1/B_2$ then $s \notin G_k$ by hypothesis, so that $s \notin R_k$ and hence $r_{B_1/B_2}(s) \leq r_{B/B}(s) < k$. Therefore

$$(|B_1||B_2|)^2 = \left(\sum_s r_{B_1B_2}(s)\right)^2 \leq |B_1B_2|\sum_s r_{B_1B_2}(s)^2 = |B_1B_2|\sum_s r_{B_1/B_2}(s)^2$$
$$\leq |B_1B_2|(k-1)\sum_s r_{B_1/B_2}(s) = |B_1B_2|(k-1)|B_1||B_2|.$$

□

Lemma 5.8. *Let $k = |B|^2/100|BB|$. There exists $h \neq 0$ for which $|B \cap hG_k(B)| \geq \frac{49}{50}|B|$.*

Proof. Let H be the set of cosets of G_k in $\mathbb{F}_p^*$. For any partition $H = H_1 \cup H_2$ let $B_j := \cup_{h \in H_j}(B \cap hG_k)|$ for $j = 1,2$ so that $B_1 \cup B_2$ is a partition of B; note that $b_1/b_2 \in (h_1/h_2)G_k$ for some $h_1 \in H_1, h_2 \in H_2$ so that $h_1 \neq h_2$ and thus $b_1/b_2 \notin G_k$. Now $|B_1|(|B|-|B_1|) = |B_1||B_2| < k|B_1B_2| < k|BB| = \frac{|B|^2}{100}$ by lemma 5.7, and so either $|B_1|$ or $|B_2|$ is $> \frac{49}{50}|B|$.

Now let H_1 be a maximal subset of H such that $|B_1| < |B|/50$. Therefore for any $h \in H_2$ we must have $|B_1 \cup (B \cap hG_k)| \geq |B|/50$ and hence $> \frac{49}{50}|B|$ by the previous paragraph, so that $|B \cap hG_k| \geq \frac{24}{25}|B|$. We deduce H_2 has no more than one element, and thus exactly one element (since $|B_2| > 0$). The result follows. □

Lemma 5.9. *Let $C \subset G$, a subgroup of $\mathbb{F}_p^*$, with $1 < |C| < \sqrt{p}$. Then we have $|G(C-C)| \gg |C|^{3/2}$ and $|G| \cdot |C-C| \gg |C|^{5/2}$.*

Proof. First note that $|G(C-C)| \geq |G|$ and $|C-C| \geq |C|$ so the results follow unless $|C| \geq 2$ and $|G| \leq |C|^{3/2}$, which we now assume.

Now, GS is a union of cosets of G for any set $S \in \mathbb{F}_p^*$, so we can write

$$G(C-C) = \{0\} \cup \cup_{i=1}^{m} t_i G,$$

where we order these cosets of G so that $r_{G-G}(t_1) \geq r_{G-G}(t_2) \geq \ldots \geq r_{G-G}(t_m)$ (note that $r_{G-G}(t_i)$ takes the same value for any choice of t_i inside a fixed coset of G). Since $|G(C-C)| = |G|m+1$, the first result follows unless $m \leq M := [|C|^{3/2}/|G|]$.

If $J \leq M$ then $J|G|^4 \leq M|G|^4 \leq |G|^3|C|^{3/2} \leq |C|^{9/2}|C|^{3/2} = |C|^6 \leq p^3$. Therefore

$$J r_{G-G}(t_J) \leq \sum_{i=1}^{J} r_{G-G}(t_i) \leq 4(J|G|)^{2/3} \tag{21}$$

by Lemma 5 of [15],[4] and so

$$|G(C-C)| > m|G| \geq \frac{1}{8}\left(\sum_{i=1}^{m} r_{G-G}(t_i)\right)^{3/2}. \tag{22}$$

For any fixed $c_0 \in C$ the solutions to $h_1 - h_2 = t_i$ with $h_1, h_2 \in G$ are in 1-1 correspondence with the solutions $h_3 - c_0 = t_i h_4$ with $h_3, h_4 \in G$, as may be seen by taking $h_3 = h_1 c_0/h_2$ and $h_4 = c_0/h_2$. Hence

$$r_{G-G}(t_i) = \sum_{t \in t_i G} r_{G-c_0}(t) \geq \sum_{t \in t_i G} r_{C-c_0}(t). \tag{23}$$

We then deduce

$$\sum_{i=1}^{m} r_{G-G}(t_i) \geq \sum_{t \in (C-C)G \setminus \{0\}} r_{C-c_0}(t) = |C| - 1,$$

and the first result follows from (22).

We now prove the second result, no longer assuming that $m \leq M$: Since $r_{C-C}(t) \leq r_{G-G}(t) = r_{G-G}(t_i)$ for all $t \in t_i G$, we have

$$\sum_{t \in t_i G} r_{C-C}(t)^2 \leq r_{G-G}(t_i) \sum_{t \in t_i G} r_{C-C}(t) \leq r_{G-G}(t_i) \sum_{c_0 \in C} \sum_{t \in t_i G} r_{C-c_0}(t) \leq |C| r_{G-G}(t_i)^2$$

by (23). Therefore, using (21) for the bound $r_{G-G}(t_J) \leq 4|G|^{2/3}/\min\{J,M\}^{1/3}$, we obtain

$$\begin{aligned}
\sum_{t \neq 0} r_{C-C}(t)^2 &= \sum_{i=1}^{M} \sum_{t \in t_i G} r_{C-C}(t)^2 + \sum_{i=M+1}^{m} \sum_{t \in t_i G} r_{C-C}(t)^2 \\
&\leq \sum_{i=1}^{M} |C| r_{G-G}(t_i)^2 + \sum_{i=M+1}^{m} r_{G-G}(t_i) \sum_{t \in t_i G} r_{C-C}(t) \\
&\leq |C| \sum_{i=1}^{M} 16|G|^{4/3} i^{-2/3} + 4|G|^{2/3} M^{-1/3} \sum_{i=M+1}^{m} \sum_{t \in t_i G} r_{C-C}(t) \\
&\leq 48|C||G|^{4/3} M^{1/3} + 4|C|^2 |G|^{2/3} M^{-1/3} \asymp 52|C|^{3/2}|G|.
\end{aligned}$$

[4] What is this reference? HBK

Hence

$$(|C|^2-|C|)^2=\left(\sum_{t\neq 0}r_{C-C}(t)\right)^2\leq |C-C|\sum_{t\neq 0}r_{C-C}(t)^2\ll |C-C||C|^{3/2}|G|,$$

and the result follows. □

Theorem 6. *If $|A|<\sqrt{p}$ then $|I(A,A)|\gg |A|^{3/2}$.*

Proof. We have $|I(A,A)|\geq 2|AA|-1$ by Lemma 5.4(ii), and the result follows unless $|AA|\ll |A|^{3/2}$, which we now assume.

Let $k=|A|^2/100|AA|$ ($\gg |A|^{1/2}$), and define $R_k(A),G_k(A),H_k(A)$ as above. By Lemma 5.8 there exists $h\neq 0$ such that if $C=\{g\in G_k(A):\ gh\in A\}=G_k\cap h^{-1}A$, then

$$|C|=|A\cap hG_k(A)|\geq \frac{49}{50}|A|.$$

Therefore, using the fact that $H=G(A-A)/(A-A)$ we have

$$|H|+1\geq |G(A-A)|\geq |G(hC-hC)|=|G(C-C)|\gg |C|^{3/2}\gg |A|^{3/2}$$

by Lemma 5.9. The result follows by Lemma 5.4 and Proposition 5.6 since now $|I(A,A)|\gg \min\{k|A|,|H_k|\}\gg |A|^{3/2}$. □

Theorem 7. *We have*

$$E_\times(A,A)\leq 4|A+A|^2\max\left\{\frac{|A|^2}{p},\frac{p}{|A|}\right\}\log|A|,$$

and

$$E_\times(A,A)^4<32|A+A|^8|A|^2\max\left\{\frac{|A|^3}{p},2|A+A|\right\}(\log|A|)^4.$$

If $|A|\leq\sqrt{p}$ then $|A|^3/p\leq |A|\leq |A+A|$, so the above becomes $E_\times(A,A)^4<64|A+A|^9|A|^2(\log|A|)^4$. This yields a sum-product bound which is non-trivial for all $|A|\leq\sqrt{p}$:

Corollary 5.10. *We have*

$$E_\times(A,B)\leq 4|A+A||B+B|\log(|A||B|)\left(\max\left\{\frac{|A|^2}{p},\frac{p}{|A|}\right\}\max\left\{\frac{|B|^2}{p},\frac{p}{|B|}\right\}\right)^{1/2},$$

and

$$E_\times(A,B)^8<2^{10}(|A+A||B+B|\log|A||B|)^8(|A||B|)^2\max\left\{\frac{|A|^3}{p},2|A+A|\right\}\times\max\left\{\frac{|B|^3}{p},2|B+B|\right\},$$

which implies that if $|A|,|B|\leq p^{1/2}$ then

$$|AB|^8(|A+A||B+B|)^9>\frac{(|A||B|)^{14}}{2^{12}(\log|A||B|)^8},$$

Proof. By the Cauchy–Schwarz inequality we have

$$E_\times(A,B)^2 = \left(\sum_n r_{A/A}(n) r_{B/B}(n)\right)^2 \le \sum_m r_{A/A}(m)^2 \sum_n r_{B/B}(n)^2 = E_\times(A,A)E_\times(B,B)$$

and then the first two results follow from Theorem 7. Now, if $|A|, |B| \le p^{1/2}$ then $|A|^3/p \le |A| < |A+A|$. Therefore, by the second inequality and (7), we obtain our third and final inequality □

If $|A|, |B| \ge p^{2/3}$ then by the first inequality in Corollary 5.10, and (7), we obtain

$$|A+A||B+B||AB| \ge \frac{p|A||B|}{4\log(|A||B|}.$$

which is weaker than Theorem 3.

Corollary 5.11. *If $4p^4/|A|^6 \ge |A+A| > |A|^3/2p$ and $|A| \le 2p^{5/9}$ then*

$$|AA|^4|A+A|^9 \ge \frac{|A|^{14}}{2^6(\log|A|)^4}$$

If $|A+A| \le |A|^3/2p$ and $p^{1/2} \le |A| \le p^{5/9}$ then

$$|AA||A+A|^2 \ge \frac{|A|^{11/4}p^{1/4}}{2^{5/4}\log|A|}$$

Proof. This follows by Theorem 7 and (7) with $B = A$, which gives

$$|A|^4 \le |AA|E_\times(A,A).$$

□

Proof of Theorem 7. We begin by noting that

$$E_\times(A,A) = \sum_{b\in A} \sum_{k=0}^{[\log|A|]} \sum_{\substack{a\in A \\ 2^k \le |aA\cap bA| < 2^{k+1}}} |aA \cap bA|$$

where the logarithm here is in base 2. Hence there exists $2^k \le |A|$ for which there is $A_1 \subset A$ and $b_0 \in A$ such that

$$2^k \le |aA \cap b_0A| < 2^{k+1}$$

for every $a \in A_1$, where $|A_1|2^{k+1} \ge E_\times(A,A)/(|A|\log|A|)$.

By Proposition 5.2 with $S = A_1$ there exists $a \in A_1$ such that

$$|aA - b_0A| \ge \frac{1}{2}\min\{p, |A|^2|A_1|/p\};$$

and $|aA - b_0A| \le |A+A|^2/|aA \cap b_0A|$ by Corollary 4.4. Hence

$$|A+A|^2 \ge \frac{E_\times(A,A)}{4\log|A|} \min\left\{\frac{p}{|A_1||A|}, \frac{|A|}{p}\right\},$$

and the first result follows.

If $\frac{A_1-A_1}{A_1-A_1} = \mathbb{F}_p$ then by Corollary 5.3 there exists $a_1,a_2,a_3,a_4 \in A_1$ such that

$$|(a_1-a_2)A_1+(a_3-a_4)A_1| > \frac{1}{2}\ \min\left\{p,|A_1|^2\right\};$$

By Proposition 4.1 we have $Y \subseteq b_0A$ such that

$$|(a_1-a_2)A+(a_3-a_4)A| \le |Y+a_1A-a_2A+a_3A-a_4A| \le |Y|\prod_{i=1}^{4}\frac{|a_iA \pm b_0A|}{|b_0A|}$$
$$\le |b_0A|\prod_{i=1}^{4}\frac{|A+A|^2}{|b_0A|\ |a_iA\cap b_0A|} \le \frac{|A+A|^8}{|A|^3\ 2^{4k}}$$

using Corollary 4.4. Hence

$$|A+A|^8 > \frac{(|A_1|2^{k+1})^4}{32}\ \min\{p/|A|,|A|\},$$

and so $E_\times(A,A)^4 \le 32|A+A|^8|A|^3(\log|A|)^4\max\left\{1,|A|^2/p\right\}$.

If $\frac{A_1-A_1}{A_1-A_1} \neq \mathbb{F}_p$ then by Corollary 5.3 there exists $a_1,a_2,a_3,a_4 \in A_1$ such that if $A_2 \subset A_1$ then

$$|(a_1-a_2)A_2+(a_1-a_2+a_3-a_4)A_1| = |A_1|\ |A_2|.$$

By Corollary 4.3 there exists $(a_1-a_2)A_2 \subset (a_1-a_2)A_1$ with $|A_2| \ge \frac{1}{2}|A_1|$ such that

$$|(a_1-a_2)A_2+(a_1-a_2)A_1+(a_3-a_4)A_1|$$
$$\le 4|A_2|\frac{|A_1+A_1|}{|(a_1-a_2)A_1|}\cdot\frac{|(a_1-a_2)A_1+(a_3-a_4)A_1|}{|(a_1-a_2)A_1|}.$$

Bounding the last term as above we obtain $|A+A|^9 > (|A_1|2^{k+1})^4|A|^2/64$, so that $E_\times(A,A)^4 \le 64|A+A|^9|A|^2(\log|A|)^4$. □

Remark (A few ideas). In the case that $\frac{A-A}{A-A} = \mathbb{F}_p$, we get in the above proof that there exists $a_1,a_2,a_3,a_4 \in A$ such that

$$\frac{1}{2}\ \min\left\{p,|A|^2\right\} \le |(a_1-a_2)A+(a_3-a_4)A| \le |bA|\prod_{i=1}^{4}\frac{|A+A|^2}{|bA|\ |a_iA\cap bA|}$$

Now note that $\sum_{a,b\in A}|aA\cap bA| = E_\times(A,A) \ge |A|^4/|AA|$ so $|aA\cap bA|$ is $|A|^2/|AA|$ on average. If we somehow get that, even with the loss of a constant (or even $|A|^\varepsilon$) for our $|a_iA\cap bA|$ then our bound here would be $|A+A|^8|AA|^4 \gg |A|^{11}\min\left\{p,|A|^2\right\}$ which is what we get in Corollary 5.11, but in a less complicated way. If we could take $b=a_1$ so we can replace one term in our product by $|A+A|/|A|$. Then we would get the bound $|A+A|^7|AA|^3 \gg |A|^9\min\left\{p,|A|^2\right\}$; this improves the exponent from $\frac{13}{12}$ when $|A| \le \sqrt{p}$ to $\frac{11}{10}$.

If $\frac{A-A}{A-A} \neq \mathbb{F}_p$ then we can change $\min\left\{p,|A|^2\right\}$ to $\min\left\{\left|\frac{A-A}{A-A}\right|,|A|^2\right\}$ using the same argument.

Combining all the results to this point, here are the results we obtained on sum-product in finite fields:

Corollary 5.12. *If $A \subseteq \mathbb{F}_p$ then*

$$\max\{|AA|,|A+A|\} \gg (\log|A|)^{O(1)} \cdot \begin{cases} \sqrt{p|A|} & \text{if } |A| \ge p^{2/3} \\ |A|^2/\sqrt{p} & \text{if } p^{2/3} > |A| \ge p^{7/13} \\ |A|^{11/12}p^{1/12} & \text{if } p^{7/13} > |A| \ge p^{13/25} \\ |A|^{14/13} & \text{if } p^{13/25} > |A| \end{cases}$$

As one can conjecture there is room for improvements. Indeed Rudnev recently published [25] a new bound

$$\max\{|AA|,|A+A|\} \gg (\log|A|)^{O(1)} \cdot |A|^{12/11} \text{ if } p^{1/2} > |A|.$$

More strikingly after completing this survey we have learned that Roche-Newton, Rudnev, and Shkredov announced [23] a fantastic bound

$$\max\{|AA|,|A+A|\} \gg |A|^{6/5} \text{ if } p^{5/8} > |A|.$$

In their proof they use incidence bounds in "Elekes style". Misha Rudnev [24] used a result of Guth and Katz on point-line incidences in space (see in [14] and in [20]) to obtain an unexpectedly strong point-plain incidence bound in K^3 for arbitrary field K. This beautiful result led to new sum-product bounds in $\mathbb{F}_p$ even for composite p.

5.1 Getting the Full Field

We now give a result of Glibichuk discussed at the start of this section:

Theorem 8. *If $|A||B| \ge 3p/2+\sqrt{p}$ then $8AB = \mathbb{F}_p$*

Proof. Suppose $|B| \ge |A|$, let $B_+ = \{b \in B : -b \in B\}$ and $B_- = \{b \in B : -b \notin B\} \cup \{b \in B_+ : 1 \le b \le \frac{p-1}{2}\}$. By definition B_+ is symmetric ($b \in B_+ \Leftrightarrow -b \in B_+$) and B_- is anti-symmetric ($b \in B_- \implies -b \notin B_-$). Let B_* be the larger of the two. By a simple counting argument we know that $|B_*| = \max\{|B_+|,|B_-|\} \ge \max\{\frac{2|B|-1}{3}, 2|B|-p\}$. Note that $|A||B_*| \ge |A| \cdot \frac{2|B|-1}{3} \ge p+2$.

Noting that $a+tb = c+td$ iff $a-td = c-tb$ we have $R(t) = R(-t)$ in Proposition 5.2 so there exists $t \ne 0$ such that $R(t) = R(-t) \le |A||B_*| + |A|^2|B_*|^2/(p-1)$ so that

$$|A+tB_*|, |A-tB_*| \ge \frac{|A|^2|B_*|^2}{R(t)} \ge \frac{|A|^2|B_*|^2}{|A||B_*| + |A|^2|B_*|^2/(p-1)} > p/2$$

as $|A||B_*| \ge p+2$. If $B_* = B_+$ then $I(A,B) = A(B+B) + B(A+A) \subset 4AB$ and so, by Lemma 5.4(i), we have $|4AB_+| \ge |I(A,B_+)| \ge |A+tB_+| > p/2$, and thus $8AB_+ = \mathbb{F}_p$ by the pigeonhole principle. If $B_* = B_-$ then, by the pigeonhole principle, there exists $a_1,a_2 \in A, b_1,b_2 \in B_-$ such that $a_1 - tb_1 = -(a_2 - tb_2)$ so that $t = (a_1+a_2)/(b_1+b_2)$ (and $b_1+b_2 \ne 0$ as B_- is antisymmetric). But then $|4AB_-| \ge |(b_1+b_2)A + (a_1+a_2)B_-| = |A+tB| > p/2$, and thus $8AB_- = \mathbb{F}_p$ by the pigeonhole principle.

The result follows. □

Corollary 5.13. *If $|A||B| \ge 3p/4+\sqrt{p}$ then $16AB = \mathbb{F}_p$.*

Proof. Suppose $|B| \ge |A|$ so that $|A||B+B| \ge |A|(2|B|-1) \ge 2|A||B| - \sqrt{|A||B|} \ge 3p/2 + \sqrt{p}$, and the result follows by applying Theorem 8 with B replaced by $B+B$, so that $16AB \subset 8A(B+B) = \mathbb{F}_p$. □

We now give a result of Hart and Iosevitch [16]:

Theorem 9. *If $\prod_{j=1}^m |A_j||B_j|/p > (p-1)$ then $\sum_{i=1}^m A_iB_i \supseteq \mathbb{F}_p^*$. In particular, if $|A||B| > p(p-1)^{1/m}$ then $mAB \supseteq \mathbb{F}_p^*$. If $\prod_{j=1}^m |A_j||B_j|/p > (p-1)^2$ then $\sum_{i=1}^m A_iB_i \supseteq \mathbb{F}_p$.*

Proof. Let $r(t)$ be the number of representations of t as $\sum_{i=1}^m a_ib_i$, so that

$$r(t) = \sum_{\substack{a_i\in A_i\\ b_j\in B_j}} \frac{1}{p}\sum_k e\left(\frac{k(\sum_i a_ib_i - t)}{p}\right) = \frac{\prod_i |A_i||B_i|}{p} + \frac{\text{Error}}{p},$$

where, by the Cauchy–Schwarz inequality, and writing $u \equiv k/l \pmod p$

$$\begin{aligned}
|\text{Error}|^2 &= \left|\sum_{a_i\in A_i}\sum_{k\ne 0} e\left(\frac{-kt}{p}\right)\prod_j\sum_{b_j\in B_j} e\left(\frac{ka_jb_j}{p}\right)\right|^2\\
&\le \sum_{a_i\in A_i}\left|\sum_{k\ne 0} e\left(\frac{-kt}{p}\right)\prod_j\sum_{b_j\in B_j} e\left(\frac{ka_jb_j}{p}\right)\right|^2\\
&\le \prod_i |A_i|\cdot\sum_{a_i}\sum_{k,l\ne 0} e\left(\frac{(l-k)t}{p}\right)\prod_j\sum_{b_j,b_j'\in B_j} e\left(\frac{a_j(kb_j - lb_j')}{p}\right)\\
&\le \prod_i |A_i|\cdot\sum_{u,l\ne 0} e\left(\frac{(1-u)lt}{p}\right)\prod_j p\sum_{b_j,ub_j\in B_j} 1.
\end{aligned}$$

Assume, for now, that $t \ne 0$. If $u \ne 1$ then the sum over l equals -1, otherwise it equals $p-1$. Hence the above is $\le (p-1)\prod_i |A_i|\cdot\prod_j p\sum_{b_j\in B_j} 1 = (p-1)p^m\prod_i |A_i||B_i|$. We deduce that $pr(t) \ge \prod_i |A_i||B_i| - ((p-1)p^m\prod_i |A_i||B_i|)^{1/2}$ and the result follows.

If $t = 0$ the sum over l is $p-1$ and it is feasible that $ub_j \in B_j$ for all u, j, so the above is $\le (p-1)^2p^m\prod_i |A_i||B_i|$ and the result follows (one can also prove this bound more directly using (16)). □

6 Helfgott's More General Bounds

Theorem 10 (Helfgott's Theorem). *Let G be a group and Γ be an abelian group of automorphisms. Let $S \subset \Gamma$ with the property*

$$\text{If } g^\sigma = g \text{ for some } g \in G, \sigma \in S^{-1}S \text{ then } g = 1 \text{ or } \sigma = 1. \tag{24}$$

Then for any $A \subset G$ we have one of the following:

(i) *There exists $g \in A$ such that $|Ag^S| = |A|\,|S|$*

Or there exists $c \in A^{-1}A$ and $\lambda \ne \tau \in S$ such that

(ii) *There exists $b \in A \cup A^{-1}$ such that $\{(ab^\sigma)^\tau c^\sigma (ab^\sigma)^{-\lambda} : a \in A,\ \sigma \in S\}$ contains $|A|\,|S|$ distinct elements.*

Or (iii) *There exists* $\eta \in S \cup S^{-1}$ *such that* $\{a^\tau (c^\eta)^\sigma a^{-\lambda} : \ a \in A, \ \sigma \in S\}$ *contains* $|A| \ |S|$ *distinct elements;*
or (iv) $\{a^\tau c^\sigma a^{-\lambda} : \ a \in A, \ \sigma \in S\}$ *contains* $\ge \frac{1}{2}\min\{|A||S|, |\mathscr{O}|\}$ *distinct elements, where* $\mathscr{O}$ *is the union of the orbits of elements of* A *under the two maps* $a \to ba$ *for any* $b \in A \cup A^{-1}$, *and* $a \to a^\eta$ *for any* $\eta \in S \cup S^{-1}$.

Proof. Define $U := \{g \in G : \text{ There exist } \lambda \ne \tau \in S \text{ such that } g^{\tau-\lambda} \in A^{-1}A\}$. For any $g \in G$ define $\phi_g : A \times S \to G$ by $\phi_g(a, \sigma) = ag^\sigma$. Note that ϕ_g is injective if and only if $g \notin U$, and in this case $|Ag^S| = |A| \ |S|$.

Next define $\delta_{\lambda,\tau}(g) = g^{\tau-\lambda}$ for $g \in G$, for each $\lambda \ne \tau \in S$. This is always injective, for if $\delta_{\lambda,\tau}(g_1) = \delta_{\lambda,\tau}(g_2)$ then $(g_1^{-1}g_2)^{\lambda\tau^{-1}} = g_1^{-1}g_2$, and so $g_1 = g_2$ by (24). Hence if $g \notin U$ then $|\delta_{\lambda,\tau}(Ag^S)| = |A| \ |S|$. We observe that

$$\delta_{\lambda,\tau}(ag^\sigma) = (ag^\sigma)^{\tau-\lambda} = a^\tau g^{\sigma(\tau-\lambda)} a^{-\lambda}, \tag{25}$$

using the fact that Γ is abelian.

Suppose that $\mathscr{O} \not\subset U$. By following how the orbits are created from A by applying the two maps, we consider the first element of $\mathscr{O}$ that is not in U. Then one of the following must be true:

(i) There exists $g \in A$ such that $g \notin U$;
(ii) There exists $u \in \mathscr{O} \cap U$ such that $g = bu \notin U$ with $b \in A \cup A^{-1}$;
(iii) There exists $u \in \mathscr{O} \cap U$ such that $g = u^\eta \notin U$ with $\eta \in S \cup S^{-1}$, where

In case (i) we have that ϕ_g is injective so that $|Ag^S| = |A| \ |S|$.

In cases (ii) and (iii) we have $u \in \mathscr{O} \cap U$ and so there exists $\lambda \ne \tau \in S$ and $c \in A^{-1}A$ such that $u^{\tau-\lambda} = c$.

In case (ii) we then have $g^{\tau-\lambda} = (bu)^{\tau-\lambda} = b^\tau u^{\tau-\lambda} b^{-\lambda} = b^\tau c b^{-\lambda}$, and so $\delta_{\lambda,\tau}(ag^\sigma) = (ab^\sigma)^\tau c^\sigma (ab^\sigma)^{-\lambda}$ by (25) and the commutativity of Γ. Hence

$$\{(ab^\sigma)^\tau c^\sigma (ab^\sigma)^{-\lambda} : \ a \in A, \ \sigma \in S\} = \delta_{\lambda,\tau}(Ag^S)$$

which has size $|\delta_{\lambda,\tau}(Ag^S)| = |A| \ |S|$.

In case (iii) we then have $g^{\tau-\lambda} = u^{(\tau-\lambda)\eta} = c^\eta$ and so, proceeding as above,

$$\{a^\tau (c^\eta)^\sigma a^{-\lambda} : \ a \in A, \ \sigma \in S\} = \delta_{\lambda,\tau}(Ag^S)$$

which has size $|\delta_{\lambda,\tau}(Ag^S)| = |A| \ |S|$.

Now suppose that $\mathscr{O} \subset U$ and define $R_g := \{a, b \in A, \lambda \ne \tau \in S : \ ag^\lambda = bg^\tau\}$. Note these are disjoint for if $(a, b, \lambda\tau) \in R_g \cap R_h$ then $g^{\tau-\lambda} = b^{-1}a = h^{\tau-\lambda}$ which is impossible as $\delta_{\lambda,\tau}$ is injective. Therefore,

$$\min_{g \in U} |R_g| \le \frac{1}{|U|} \#\{a, b \in A, \lambda \ne \tau \in S\} < \frac{(|A||S|)^2}{|\mathscr{O}|}$$

and, since

$$\sum_n r_{Ag^S}(n)^2 = \#\{a, b \in A, \lambda, \tau \in S : \ ag^\lambda = bg^\tau\} = |R_g| + |A||S|,$$

we deduce that for some $g \in U$,

$$|A|^2|S|^2 = \left(\sum_n r_{Ag^S}(n)\right)^2 \le |Ag^S| \sum_n r_{Ag^S}(n)^2 \le |Ag^S| \left(\frac{(|A||S|)^2}{|\mathscr{O}|} + |A||S|\right)$$

by the Cauchy–Schwarz inequality. As $g \in U$, there exists $\lambda \neq \tau \in S$ and $c \in A^{-1}A$ such that $g^{\tau-\lambda} = c$, and so, by (25),

$$\{a^\tau c^\sigma a^{-\lambda} : a \in A,\ \sigma \in S\} = \delta_{\lambda,\tau}(Ag^S)$$

which has size

$$|\delta_{\lambda,\tau}(Ag^S)| = |Ag^S| \ge \frac{|A||S||\mathscr{O}|}{|A||S| + |\mathscr{O}|} \ge \frac{1}{2}\min\{|A||S|, |\mathscr{O}|\}.$$

□

References

1. F. Amoroso and E. Viada. Small points on subvarieties of a torus. *Duke Mathematical Journal*, 150(3):407–442, 2009.
2. Ayyad, Anwar, Cochrane, Todd, Zheng, Zhiyong, The congruence $x_1x_2 \equiv x_3x_4 \pmod{p}$, the equation $x_1x_2 = x_3x_4$, and mean values of character sums. *J. Number Theory* **59** (1996), 398–413.
3. Y. Bilu. The Many Faces of the Subspace Theorem (after Adamczewski, Bugeaud, Corvaja, Zannier...). Séminaire Bourbaki, Exposé 967, 59ème année (2006-2007); Astérisque 317 (2008), 1-38., May 2007.
4. Bourgain, Jean, Glibichuk, Alexey A., Konyagin, Sergei V., Estimates for the number of sums and products and for exponential sums in fields of prime order. *J. London Math. Soc* **73** (2006), 380–398.
5. Bourgain, Jean, Katz, Nets, Tao, Terry, A sum-product estimate in finite fields, and applications. *GAFA* **14** (2004), 27–57.
6. M.-C. Chang. Sum and product of different sets. *Contributions to Discrete Mathematics*, 1(1), 2006.
7. Erdős, Paul, Szemerédi, Endre, On sums and products of integers. *Studies in Pure Mathematics* (Birkhäuser, Basel), 1983, pp. 213–218.
8. J.-H. Evertse, H.P. Schlickewei, and W.M. Schmidt. Linear equations in variables which lie in a multiplicative group. *Annals of Mathematics*, 155(3):807–836, 2002.
9. Ford, K., The distribution of integers with a divisor in a given interval. *Ann. of Math.* **168** (2008), 367–433.
10. G.A. Freiman. *Foundations of a Structural Theory of Set Addition*. Translations of Mathematical Monographs. American Mathematical Society, 1973.
11. Garaev, M. Z., An explicit sum-product estimate in $\mathbb{F}_p$. *IMRN* **35** (2007), 11.
12. Garaev, M. Z., The sum-product estimate for large subsets of prime fields. *Proc. Amer. Math. Soc.* **136** (2008), 2735–2739.
13. Glibichuk, Alexey and Rudnev, Misha Additive properties of product sets in an arbitrary finite field. Journal d'Analyse Mathématique May 2009, Volume 108, Issue 1, pp 159–170.
14. Guth, L. and Katz. N. H., On the Erdős distinct distances problem in the plane Pages 155–190 from Volume 181 (2015), Issue 1
15.

16. Hart, Derrick, Iosevitch, Alex, Sums and products in finite fields: an integral geometric viewpoint. in: Radon Transforms, Geometry, and Wavelets: AMS Special Session, January 7-8, 2007, New Orleans, Louisiana, Contemporary Mathematics, Volume 464, 2008. 129–135.
17. Hart, Derrick, Iosevitch, Alex, Solymosi, József, Sums and products in finite fields via Kloosterman sums *IMRN* Art. ID rnm007 (2007), 114.
18. Helfgott, Harald, Growth and generation in $SL\_2(Z/pZ)$, Ann. of Math. 167 (2008), 601–623.
19. Hawk Katz, Nets, Shen, Chun-Yen, A slight improvement to Garaev's sum-product estimate. Proc. Amer. Math. Soc. 136 (2008), 2499–2504
20. János Kollár, Szemerédi-Trotter-type theorems in dimension 3, Advances in Mathematics, Volume 271, 5 February 2015, Pages 30–61,
21. Konyagin, S V and Rudnev, M , On new sum-product type estimates, SIAM Journal on Discrete Mathematics, vol 27, no. 2, pp. 973–990.
22. Konyagin, S V and Shkredov, I D, On sum sets, having small product set, arXiv:1503.05771 [math.CO]
23. Oliver Roche-Newton, Misha Rudnev, Ilya D. Shkredov, New sum-product type estimates over finite fields, arXiv:1408.0542 [math.CO]
24. Rudnev, M. On the number of incidences between planes and points in three dimensions, arXiv:1407.0426 [math.CO]
25. Rudnev, M, An improved sum-product inequality in fields of prime order International Mathematics Research Notices, vol 2012., pp. 3693–3705
26. Jozsef Solymosi and Gabor Tardos. 2007. On the number of k-rich transformations. In Proceedings of the twenty-third annual symposium on Computational geometry (SCG '07). ACM, New York, NY, USA, 227-231.
27. Solymosi, J., On the number of sums and products. *Bull. London Math. Soc* **37** (2005), 491–494.
28. Solymosi, J., Bounding multiplicative energy by the sumset, Advances in Mathematics, Volume 222, Issue 2, 1 October 2009, Pages 402–408,
29. Tao, Terry and Vu, Van H., Additive Combinatorics. In: *Cambridge studies in advanced math* **105**, Cambridge University Press, Cambridge, 2006.
30. Le Anh Vinh, The Szemerédi-Trotter type theorem and the sum-product estimate in finite fields, European Journal of Combinatorics, Volume 32, Issue 8, November 2011, Pages 1177–1181.