SOLYMOSI'S THEOREM

LEO GOLDMAKHER

ABSTRACT. In 2008, Solymosi proved that for any finite set $A \subset \mathbb{R}$,

$$\max\{|A \cdot A|, |A + A|\} \gg |A|^{4/3 - o(1)}.$$

This is the strongest bound currently available, even in the restricted case that $A \subset \mathbb{Z}$. The strength of this result is particularly remarkable in view of the simplicity of Solymosi's argument, which is accessible to a second-year undergraduate and can be explained using a single bar napkin. (These claims have been empirically tested.) Here I give a longer, self-contained exposition of the proof.

1. INTRODUCING THE PROBLEM

Given two finite subsets A and B of an integral domain[1], we define the *sumset* of A and B to be

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

and the *productset* of A and B to be

$$A \cdot B = \{ab : a \in A, b \in B\}.$$

What can be said about the structure of these sets? We begin our exploration with some simple examples.

Let $A_n = \{1, 2, \dots, n\}$. Then we have $A_n + A_n = \{2, 3, \dots, 2n\}$. In particular, we see that

$$|A_n + A_n| = 2|A_n| - 1. \tag{1}$$

Note that this is as small as possible, since for *any* set A we have $|A + A| \geq 2|A| - 1$. (Are there sets $A \neq A_n$ for which equality is achieved in the lower bound?)

Next, we consider the productset $A_n \cdot A_n$. It is a much harder problem to explicitly list the elements of this set. Nonetheless, we can get a good idea of $|A_n \cdot A_n|$ without too much work. Indeed, let $\pi(n)$ denote the number of primes in A_n. Observing that the product of any pair of primes is unique up to permutation and applying the Prime Number Theorem yields

$$|A_n \cdot A_n| \geq \frac{1}{2}\pi(n)^2 \gg |A_n|^{2 - o(1)} \tag{2}$$

where $o(1)$ is a quantity which tends to 0 as $n \to \infty$ and $f \ll g$ means the same thing as $f = O(g)$: that there exists a constant $C > 0$ such that $|f(x)/g(x)| \leq C$ for all x in the domain of f/g. The bound (2) shows that the productset $A_n \cdot A_n$ is essentially as large as possible, since $|A \cdot A| \ll |A|^2$ for any set A.

[1] An *integral domain* is any commutative ring $\mathcal{R}$ satisfying the cancellation property: if $ab = ac$, then $b = c$. For the purposes of this essay, you can think of $\mathcal{R}$ as being $\mathbb{R}$ or $\mathbb{Z}$.

Thus, the sumset of A_n with itself is very small, and the productset of A_n is very large. Are there sets which have small productset instead? A little thought leads to the set

$$B_n = \{2^0, 2^1, 2^2, \dots, 2^{n-1}\}.$$

Then we have $B_n \cdot B_n = \{2^0, 2^1, \dots, 2^{2n-2}\}$, whence

$$|B_n \cdot B_n| = 2|B_n| - 1.$$

How large is the sumset? When written in binary notation, the sum of any two distinct elements of B_n has precisely two 1's and at most n digits. Conversely, any such number is the sum of two distinct elements of B_n. It follows that there are precisely $\binom{n}{2}$ distinct elements of $B_n + B_n$, which come from summing two distinct element of B_n. The only elements of $B_n + B_n$ we are missing are the sums of an element of B_n with itself; it is easily seen that we have not yet counted these (they all have exactly one digit 1), and that there are precisely n such sums. Putting this together, we find that

$$|B_n + B_n| = \binom{n}{2} + n \geq \frac{1}{2}|B_n|^2.$$

Thus, similarly to the case of A_n, one of $|B_n + B_n|, |B_n \cdot B_n|$ is as small as possible, and the other is essentially as large as possible.

A natural question arises – is it true that for any set A, one of $|A + A|$, $|A \cdot A|$ is small, and the other is large? Some playing around shows that it is very unusual for either the sumset or the productset to be small. By contrast, one of $|A+A|, |A \cdot A|$ seems to always be exceptionally large. This guess was formalized by Erdős and Szemerédi [ES]:

Conjecture 1.1 (Erdős-Szemerédi)**.** *For any $A \subseteq \mathbb{Z}$ we have*

$$\max\{|A \cdot A|, |A + A|\} \gg |A|^{2-o(1)}$$

where $o(1) \to 0$ as $|A| \to \infty$.

The conjecture remains wide open, and the best exponent currently known is $4/3 - o(1)$; much smaller than the conjectured exponent of $2 - o(1)$.

The examples above give some empirical evidence for the conjecture. In the next section, I'll outline an argument which lends Conjecture 1.1 further credibility.

2. Motivation: Freiman-Ruzsa theorem

Most sets have exceptionally large sumset and productset.[2] But what sort of set has exceptionally small sumset or productset? Building on our first example A_n, we see that *any* arithmetic progression has small sumset. Remarkably, a converse to this statement also holds:

Theorem 2.1 (Freiman-Ruzsa Theorem)**.** *Suppose $A \subseteq \mathbb{Z}$ satisfies $|A + A| \leq k|A|$. Then A is contained in a generalized arithmetic progression[3] of dimension $\ll_k 1$ and size $\ll_k |A|$.*

[2] I'm cheating when I say *most* – there are multiple ways to make this statement precise. I don't know of a natural way to do this over $\mathbb{Z}$.

[3] A d-dimensional *generalized arithmetic progression* (often abbreviated gAP) is a natural extension of an arithmetic progression: rather than having a single constant difference, one has d constant differences.

The k in the subscripts indicates that the implicit constant is allowed to depend on k, but on nothing else. It might seem odd to leave these dependencies implicit, and actually some explicit bounds are known. I have purposely suppressed the explicit statement in order to highlight the heart of the theorem: that the dimension does *not* depend on the size of A, and that the gAP containing A isn't much larger than A itself.

This deep theorem was originally discovered and proved by Freiman in the 1960s [Fr]. However, it was an elegant proof due to Ruzsa [Ru] which ignited wide interest in the subject. The two underlying ideas of the proof are:

> (i) for any A, sets of the form $mA - \ell A$ have a lot of additive structure once m and ℓ are large enough[4]; and
>
> (ii) if $mA - \ell A$ contains a large finite-dimensional arithmetic progression, then so does A.

The latter idea is not difficult to make precise, but proving a precise form of (i) seems to be quite hard. The starting point is a result of Bogolyubov, who proved that for any $A \subseteq \mathbb{F}_p$ of positive density, the set $2A - 2A$ contains a large subset with a lot of diophantine structure (called a *Bohr set*). A deep result from the geometry of numbers (Minkowski's 2^{nd} theorem) implies that any finite-dimensional Bohr set contains a long finite-dimensional arithmetic progression, thus giving a version of (i) over finite fields (with $m = \ell = 2$). Ruzsa realized that by allowing m and ℓ to be larger, one can bootstrap from the finite field case to the integers. To accomplish this, he first employed a bound of Plünnecke on graph connectivity to show that if $A \subset \mathbb{Z}$ satisfies $|2A| \ll |A|$, then $|mA - \ell A| \ll |A|$ as well. Next, Ruzsa built on Freiman's original arguments to prove that given any $A \subset \mathbb{Z}$ satisfying $|mA - mA| \ll |A|$, it is possible to embed almost all of A into some finite field $\mathbb{F}_p$ (with $p \asymp |A|$) in such a way that finite-dimensional arithmetic progressions are preserved under the embedding. Thus, one can translate the problem back and forth between $\mathbb{Z}$ and $\mathbb{F}_p$ without much loss, which is how the argument succeeds. (Obviously I'm being vague here; the actual arguments are quite involved.[5] The effort necessary to understand them is amply rewarded, however.)

It is natural to look for a version of the Freiman-Ruzsa theorem for productsets, and a simple computation shows that any geometric progression has exceptionally small productset. However, no converse to this statement has been proved, and a multiplicative analogue of Freiman-Ruzsa remains a tantalizing open problem. Assuming the existence of a multiplicative Freiman-Ruzsa theorem, we see that it should be impossible for both $A + A$ and $A \cdot A$ to be simultaneously small, since this should only happen in the event that A simultaneously looks like an arithmetic progression and a geometric progression. Thus, we expect at least one of $A + A$ or $A \cdot A$ to be large. This gives theoretical evidence that some statement along the lines of the Erdős-Szemerédi conjecture should hold.

---

[4] The notation mA means the sum of m copies of A.

[5] A readable and thorough exposition of the Ruzsa's proof can be found in Soundararajan's lecture notes on additive combinatorics, available on his homepage.

There has been a lot of work towards Conjecture 1.1, involving increasingly ingenious and complicated arguments. In 2008, Solymosi [So] discovered a short, beautiful, and completely elementary argument which gives the strongest result known towards Erdős-Szemerédi. Moreover, his result holds for all subsets of $\mathbb{C}$, not just of $\mathbb{Z}$. We state and analyze his theorem in the next section.

3. Solymosi's theorem: statement and discussion

Without further ado, we state Solymosi's theorem. I should point out that although the theorem is stated only for subsets of the positive real numbers, it can be extended to arbitrary subsets of $\mathbb{C}$ (see for example [KR]).

Theorem 3.1 (Solymosi, 2009)**.** *Given two finite sets of positive real numbers A and B, we have*

$$|AB||A+A||B+B| \gg \frac{|A|^2|B|^2}{\log|A|}$$

The implicit constant is absolute and effective.

This immediately gives

$$\max\{|A\cdot A|, |A+A|\} \gg |A|^{4/3-o(1)},$$

which is the largest exponent known toward the Erdős-Szemerédi conjecture. Moreover, in the extreme case $|A+A| \ll |A|^{1+\epsilon}$, Solymosi's theorem yields the lower bound $|AA| \gg |A|^{2-\delta}$ (where $\delta \to 0$ as $\epsilon \to 0$); this essentially confirms Erdős-Szemerédi. On the other hand, if $|AA| \ll |A|^{1+\epsilon}$, Solymosi's theorem only yields the bound $|A+A| \gg |A|^{3/2-\delta}$. Even so, this recovers the strongest result previously known, due to Elekes and Ruzsa [ER].[6]

We make several observations about Solymosi's lower bound. First, the $\log|A|$ is standing in for $\max\{\log|A|, \log|B|\}$, since without loss of generality A is the larger set of the two. More importantly, it is well-known[7] that

$$|A_n \cdot A_n| \ll \frac{|A_n|^2}{(\log|A_n|)^\delta} \tag{3}$$

for some constant $\delta > 0$, which shows that the logarithm cannot be entirely removed from Solymosi's lower bound.

The rest of this section is devoted to proving the upper bound (3). Since this plays no role in the proof of Solymosi's theorem, the reader may freely choose to skip ahead to Section 4, where the proof of Theorem 3.1 is presented.

The bound (3) is well-known to the experts; indeed, much more precise statements are known. However, (3) suffices for many applications, and can be proved with relative ease. My main motivation in writing down a proof here is that I couldn't find it in the literature; authors tend to either prove a strong version of the bound (in which case the proof is quite technical), or else a qualitative version of the bound in the form $|A_n\cdot A_n| = o(|A_n|)$, which has an easy proof but is hard to employ in applications. It's worth pointing out that even the Erdős-Kac theorem is not strong enough to

[6] In the special case that $A \subset \mathbb{N}$, Chang [Ch] has shown that if $|AA| < \alpha|A|$ then $|A+A| > 36^{-\alpha}|A|^2$. This once again approaches Erdős-Szemerédi, but only when α is quite small.

[7] This is the famous *Multiplication Table* problem posed by Erdős: how many distinct integers appear in the $N \times N$ multiplication table? The definitive results on this are due to Ford [Fo]; see also the Ph.D. thesis of D. Koukoulopoulos.

yield the bound (3); it gives only a $\log\log|A_n|$ savings.

We begin by reformulating the upper bound (3):

Theorem 3.2. *Let $\mathcal{A}(N) := \{ab : a, b \le N\}$. There exists an absolute constant $\delta > 0$ such that*

$$|\mathcal{A}(N)| \ll \frac{N^2}{(\log N)^\delta}. \tag{4}$$

Proof. [8] Rather than dealing directly with $\mathcal{A}(N)$, it will be easier to work with the quantity

$$\mathcal{A}^*(N) := \{ab : a, b \le N \text{ and } (a,b) = 1\}.$$

Note that for each $n \in \mathcal{A}(N)$, there exists some $d \le N$ such that $\frac{n}{d^2} \in \mathcal{A}^*\left(\frac{N}{d}\right)$; it follows that

$$|\mathcal{A}(N)| \le \sum_{d \le N} \left|\mathcal{A}^*\left(\frac{N}{d}\right)\right|.$$

We deduce that (4) is implied by the weaker bound

$$|\mathcal{A}^*(N)| \ll \frac{N^2}{(\log N)^\delta}. \tag{5}$$

Observe that

$$|\mathcal{A}^*(N)| \ll \left|\left\{n \in \mathcal{A}^*(N) : \omega(n) \le \frac{3}{2}\log\log N\right\}\right| + \left|\left\{n \le N^2 : \omega(n) > \frac{3}{2}\log\log N\right\}\right| \tag{6}$$

where $\omega(n)$ denotes the number of distinct prime factors of n (counted without multiplicity). We can make the two terms on the right hand side look more similar by noting that if $n \in \mathcal{A}^*(N)$, then $\omega(n) = \omega(a) + \omega(b)$ for some $a, b \le N$. It follows that

$$\begin{aligned}\left|\left\{n \in \mathcal{A}^*(N) : \omega(n) \le \frac{3}{2}\log\log N\right\}\right| &\le \left|\left\{(a,b) : a, b \le N \text{ and } \omega(a) + \omega(b) \le \frac{3}{2}\log\log N\right\}\right| \\ &\le 2N \times \left|\left\{m \le N : |\omega(m) - \log\log N| \ge \frac{1}{4}\log\log N\right\}\right|.\end{aligned}$$

Plugging this into (6) yields

$$\begin{aligned}|\mathcal{A}^*(N)| \ll N \times &\left|\left\{m \le N : |\omega(m) - \log\log N| \ge \frac{1}{4}\log\log N\right\}\right| + \\ &+ \left|\left\{n \le N^2 : |\omega(n) - \log\log N| \ge \frac{1}{2}\log\log N\right\}\right|. \end{aligned} \tag{7}$$

Recall Hardy and Ramanujan's celebrated result that the normal order of $\omega(n)$ is $\log\log n$. We therefore expect both terms on the right hand side of (7) to be small. More precisely, we claim that for all $\alpha \in [0, 1/2]$,

$$\frac{1}{N}\left|\{n \le N : |\omega(n) - \log\log N| \ge \alpha\log\log N\}\right| \ll (\log N)^{-\alpha^2/4}. \tag{8}$$

Applying this bound in (7) yields (5), and hence the theorem.

[8] I'm grateful to M. Radziwill for introducing me to this approach. I've also borrowed heavily from the thesis of D. Koukoulopoulos and from Montgomery and Vaughan's book [MV].

It therefore suffices to prove (8). To do so, we first transform the problem into one of multiplicative number theory. Let

$$S = S_\alpha(N) = \{n \le N : |\omega(n) - \log\log N| \ge \alpha \log\log N\}$$

and let χ denote the characteristic function of S (i.e. $\chi(n) = 1$ if $n \in S$ and $\chi(n) = 0$ otherwise). Observe that for all $n \le N$,

$$\chi(n) \le \kappa_1^{\omega(n) - (1+\alpha)\log\log N} + \kappa_2^{(1-\alpha)\log\log N - \omega(n)}$$

for any constants $\kappa_i \ge 1$. It follows that

$$\begin{aligned} \frac{1}{N}|S_\alpha(N)| &= \frac{1}{N}\sum_{n\le N}\chi(n) \\ &\le (\log N)^{-(1+\alpha)\log(1+\alpha)}\frac{1}{N}\sum_{n\le N}(1+\alpha)^{\omega(n)} \\ &\qquad + (\log N)^{-(1-\alpha)\log(1-\alpha)}\frac{1}{N}\sum_{n\le N}(1-\alpha)^{\omega(n)} \end{aligned} \tag{9}$$

where we have made the choices $\kappa_1 = 1+\alpha$ and $\kappa_2 = (1-\alpha)^{-1}$. Thus, we have reduced the problem to estimating the mean value of multiplicative functions of the form $\lambda^{\omega(n)}$. From our work below (namely, from Corollary 3.4 applied to the function $f(n) = \lambda^{\omega(n)}$) we shall deduce:

$$\frac{1}{N}\sum_{n\le N}\lambda^{\omega(n)} \ll (\log N)^{\lambda-1} \tag{10}$$

for all $\lambda \in [1/2, 3/2]$. Taking this on faith for the moment, we can use it to bound the right hand side of (9):

$$\frac{1}{N}|S_\alpha(N)| \ll (\log N)^{-\lambda_1\log\lambda_1+\lambda_1-1} + (\log N)^{-\lambda_2\log\lambda_2+\lambda_2-1}. \tag{11}$$

where $\lambda_1 = 1+\alpha$ and $\lambda_2 = 1-\alpha$. It is a straightforward calculus exercise to show that

$$-\lambda\log\lambda + \lambda - 1 \le -\frac{(\lambda-1)^2}{4}$$

whenever $\frac{1}{2} \le \lambda \le \frac{3}{2}$. Employing this in (11) yields

$$\frac{1}{N}|S_\alpha(N)| \ll (\log N)^{-\alpha^2/4},$$

thus concluding the proof of (8) and, hence, of the theorem. □

To complete the above proof, it remains only to prove the bound (10). We will deduce this from a more general result:

Lemma 3.3. *Suppose f is a real-valued, non-negative multiplicative function satisfying*

$$\sum_{p\le x} f(p)\log p \ll x \qquad \text{and} \qquad \sum_{\substack{p^k \\ k\ge 2}} \frac{f(p^k)}{p^k}\log p \ll 1.$$

Then

$$\frac{1}{x}\sum_{n\le x} f(n) \ll \frac{1}{\log x}\sum_{n\le x}\frac{f(n)}{n}$$

Before proving Lemma 3.3, we observe that the hypotheses on f are fairly restrictive. Indeed, we have

$$\sum_{p\le x}\log p \gg x \qquad \text{and} \qquad \sum_{\substack{p^k\\ k\ge 2}}\frac{\log p}{p^k}\gg 1$$

which indicates that any f satisfying the hypotheses must be small most of the time. For example, $f(t)=\log t$ does not satisfy the hypotheses.

Proof of Lemma 3.3 (see Theorem 2.14 in [MV]*).* First, I claim that for *any* multiplicative function,

$$\sum_{n\le x} f(n)\log n \ll \sum_{p^k\le x}(\log p) f(p^k)\sum_{m\le x/p^k} f(m). \tag{12}$$

This comes from using the identity $\log n = \sum_{d|n}\Lambda(n)$ (the von Mangoldt function) and applying standard algebraic manipulations. Separating the RHS of (12) into two pieces (one with $k=1$, the other with $k\ge 2$) and applying the hypotheses on the size of f, we deduce that

$$\sum_{n\le x} f(n)\log n \ll x\sum_{n\le x}\frac{f(n)}{n}.$$

Next, observe that

$$\sum_{n\le x} f(n)\log\frac{x}{n} \le \sum_{n\le x} f(n)\frac{x}{n}.$$

Summing these two inequalities yields the lemma. □

Corollary 3.4. *If f satisfies the hypotheses of Lemma 3.3, then*

$$\frac{1}{x}\sum_{n\le x} f(n) \ll \frac{1}{\log x}\exp\left(\sum_{p\le x}\frac{f(p)}{p}\right).$$

Proof. By Lemma 3.3, it suffices to prove

$$\sum_{n\le x}\frac{f(n)}{n} \ll \exp\left(\sum_{p\le x}\frac{f(p)}{p}\right).$$

First, observe that

$$\begin{aligned}\exp\left(\sum_{p^k\le x}\frac{f(p^k)}{p^k}\right) &= 1+\sum_{p^k\le x}\frac{f(p^k)}{p^k}+\frac{1}{2}\left(\sum_{p^k\le x}\frac{f(p^k)}{p^k}\right)^2+\cdots\\ &\ge 1+\sum_{p^k\le x}\frac{f(p^k)}{p^k}+\sum_{p^k,q^\ell\le x}\frac{f(p^kq^\ell)}{p^kq^\ell}+\cdots\\ &\ge \sum_{n\le x}\frac{f(n)}{n}.\end{aligned}$$

On the other hand, we have

$$\sum_{p^k\le x}\frac{f(p^k)}{p^k} = \sum_{p\le x}\frac{f(p)}{p}+O(1)$$

by one of the hypotheses on f. Substituting this into the bound above yields the claim. □

4. SOLYMOSI'S THEOREM: PROOF

It will be convenient to measure redundancies in the sets we're dealing with. Given two subsets A and B and a binary operation $\otimes$ on $\mathbb{R}$, let

$$\rho_{A\otimes B}(x) := |\{(a,b) \in A \times B : a \otimes b = x\}|.$$

This notation will allow us to outline Solymosi's proof with relative ease (a detailed proof follows our outline). At the heart of the proof is the identity

$$\sum_{x \in AB} \rho_{AB}(x)^2 = \sum_{m \in B/A} \rho_{B/A}(m)^2, \tag{13}$$

which is unexpected but trivial to verify. The strategy is to bound this quantity from above and below in terms of sumsets and productsets of A and B; the resulting inequality is Solymosi's theorem.

The lower bound comes from an application of Cauchy-Schwarz to the identity

$$|A \times B| = \sum_{x \in AB} \rho_{AB}(x). \tag{14}$$

(The same holds with multiplication replaced by any binary operation, but we won't require this level of generality.) This bounds the left hand side of (13) from below by $|A|^2|B|^2/|AB|$.

The upper bound comes from the observation that $\rho_{B/A}(m)$ admits a natural geometric interpretation: it counts the number of lattice points (i.e. points of $A \times B$) lying on the line $\mathcal{L}_m$ through the origin with slope m. Using some elementary geometric arguments, for example the relation

$$\rho_{\mathcal{L}_m + \mathcal{L}_n}(x,y) = 1, \tag{15}$$

it will be seen that the right hand side of (13) counts lattice points lying in sets of the form $\mathcal{L}_m + \mathcal{L}_n$. All such lattice points are trivially contained in $(A \times B) + (A \times B)$, which has size $|A + A| \cdot |B + B|$. Combining this upper bound with the lower bound from above, we find that $|AB| \cdot |A + A| \cdot |B + B| \gtrapprox |A|^2|B|^2$, which is the conclusion of Solymosi's theorem.

Proof of Solymosi's theorem. Applying the Cauchy-Schwarz inequality to (14), we find

$$|A \times B|^2 = \left(\sum_{x \in AB} \rho_{AB}(x) \right)^2 \leq |AB| \sum_{x \in AB} \rho_{AB}(x)^2.$$

The identity (13) therefore gives the lower bound

$$\sum_{m \in B/A} \rho_{B/A}(m)^2 \geq \frac{|A|^2|B|^2}{|AB|}$$

(the proof of (13) is left as an exercise to the reader).

We now translate the problem into a geometric setting. To simplify the exposition, we introduce a couple of pieces of notation. First, given two sets S and I, define $S^I := S \cap I$. Next, let $\mathcal{L}_m$ denote the line of slope m which passes through the origin. Observe that

$$\rho_{B/A}(m) = |\mathcal{L}_m^{A \times B}|,$$

the number of lattice points (i.e. points in $A \times B$) on the line $\mathcal{L}_m$. In this new language, our lower bound from above reads

$$\sum_{m \in B/A} \left|\mathcal{L}_m^{A\times B}\right|^2 \geq \frac{|A|^2|B|^2}{|AB|}. \tag{16}$$

The strategy is to find an upper bound on this sum in terms of sumsets, by coming up with an appropriate geometric interpretation of the sum. Unfortunately, I have no idea what a natural interpretation of this sum is. So, instead, we make a slight detour: we estimate the sum by another sum which *does* admit a clear geometric interpretation.

First, we restrict to a special set of lines which give the bulk of the contribution. Pick any $M \in \mathbb{N}$ which maximizes the quantity

$$\sum_{\substack{m \in B/A \\ 2^M \leq \left|\mathcal{L}_m^{A\times B}\right| < 2^{M+1}}} \left|\mathcal{L}_m^{A\times B}\right|^2,$$

and set

$$\mathcal{M} = \{m \in B/A : 2^M \leq \left|\mathcal{L}_m^{A\times B}\right| < 2^{M+1}\}.$$

It seems likely that the set of lines $\mathcal{L}_m$ parametrized by $\mathcal{M}$ has some natural geometric interpretation, but again it's not obvious to me what this should be. In any event, we can approximate the entire sum in (16) by just this one maximal piece:

$$\begin{aligned}\sum_{m \in B/A} \left|\mathcal{L}_m^{A\times B}\right|^2 &= \sum_{j \in \mathbb{N}} \sum_{\substack{m \in B/A \\ 2^j \leq \left|\mathcal{L}_m^{A\times B}\right| < 2^{j+1}}} \left|\mathcal{L}_m^{A\times B}\right|^2 \\ &\ll (\log |A|) \sum_{m \in \mathcal{M}} \left|\mathcal{L}_m^{A\times B}\right|^2. \end{aligned} \tag{17}$$

Recall that our immediate goal is to rewrite the above sum in a way which admits a geometric interpretation. Our key tool will be the following result.

Lemma 4.1. *Let $\mathcal{L}_m$ denote the line through the origin of slope m, and define*

$$\mathcal{P} = \left\{(x, y) \in \mathbb{R}^2 : x > 0 \text{ and } y > 0\right\}.$$

Suppose $0 < m < n$ and $\mathcal{S} \subseteq \mathcal{P}$. Then the addition map

$$\begin{array}{ccc} \mathcal{L}_m \times \mathcal{L}_n & \longrightarrow & \mathbb{R}^2 \\ (\alpha, \beta) & \longmapsto & \alpha + \beta \end{array}$$

restricts to an injection $\mathcal{L}_m^{\mathcal{S}} \times \mathcal{L}_n^{\mathcal{S}} \hookrightarrow \bigcup_{m<t<n} \mathcal{L}_t^{\mathcal{P}}$.

Proof. The lemma can be viewed as having two separate claims:

(1) the restricted addition map $\mathcal{L}_m^{\mathcal{S}} \times \mathcal{L}_n^{\mathcal{S}} \longrightarrow \mathcal{L}_m^{\mathcal{S}} + \mathcal{L}_n^{\mathcal{S}}$ is an injection; and
(2) $\mathcal{L}_m^{\mathcal{S}} + \mathcal{L}_n^{\mathcal{S}} \subseteq \bigcup_{m<t<n} \mathcal{L}_t^{\mathcal{P}}$.

It is a straightforward exercise to prove that the unrestricted addition map $\mathcal{L}_m \times \mathcal{L}_n \longrightarrow \mathcal{L}_m + \mathcal{L}_n$ is an injection. The first claim immediately follows, since the restricted map $\mathcal{L}_m^{\mathcal{S}} \times \mathcal{L}_n^{\mathcal{S}} \longrightarrow \mathcal{L}_m^{\mathcal{S}} + \mathcal{L}_n^{\mathcal{S}}$ automatically inherits injectivity from the unrestricted map.

To prove the second claim, observe that $\mathcal{L}_m^{\mathcal{P}} + \mathcal{L}_n^{\mathcal{P}} = \bigcup_{m<t<n} \mathcal{L}_t^{\mathcal{P}}$. Since $\mathcal{L}_m^{\mathcal{S}} + \mathcal{L}_n^{\mathcal{S}}$ is trivially a subset of $\mathcal{L}_m^{\mathcal{P}} + \mathcal{L}_n^{\mathcal{P}}$, we conclude. □

Recall that we had identified a special set $\mathcal{M}$ parametrizing the lines which contribute most to the sum in (16). Denote this (finite) set $\mathcal{M} = \{m_1, m_2, \ldots\}$ where the m_i are increasing. By the lemma, the set $\mathcal{L}_{m_i}^{A\times B} + \mathcal{L}_{m_{i+1}}^{A\times B}$ consists of points lying in the sector bounded by the lines $\mathcal{L}_{m_i}$ and $\mathcal{L}_{m_{i+1}}$ in the first quadrant. It follows that

$$\bigcup_i \left(\mathcal{L}_{m_i}^{A\times B} + \mathcal{L}_{m_{i+1}}^{A\times B}\right) \tag{18}$$

is a *disjoint* union. Moreover, by the injectivity part of the lemma, we can measure the size of each of these disjoint pieces:

$$\begin{aligned}\left|\mathcal{L}_{m_i}^{A\times B} + \mathcal{L}_{m_{i+1}}^{A\times B}\right| &= \left|\mathcal{L}_{m_i}^{A\times B} \times \mathcal{L}_{m_{i+1}}^{A\times B}\right| \\ &= \left|\mathcal{L}_{m_i}^{A\times B}\right| \cdot \left|\mathcal{L}_{m_{i+1}}^{A\times B}\right|.\end{aligned}$$

It follows that

$$\left|\bigcup_i \left(\mathcal{L}_{m_i}^{A\times B} + \mathcal{L}_{m_{i+1}}^{A\times B}\right)\right| = \sum_i \left|\mathcal{L}_{m_i}^{A\times B}\right| \cdot \left|\mathcal{L}_{m_{i+1}}^{A\times B}\right|. \tag{19}$$

We have therefore found a sum – namely, the right hand side of (19) – which has a clear geometric interpretation, and simultaneously looks like our original sum from (17). I now claim the sums don't just resemble each other, but are actually almost equal (up to a factor of 2 and a small error). Note that from the definition of $\mathcal{M}$, all lines $\mathcal{L}_m$ with $m \in \mathcal{M}$ contain the same number of lattice points (up to a factor of 2). Thus, we have

$$\sum_{m\in\mathcal{M}} \left|\mathcal{L}_m^{A\times B}\right|^2 \ll \sum_i \left|\mathcal{L}_{m_i}^{A\times B}\right| \cdot \left|\mathcal{L}_{m_{i+1}}^{A\times B}\right| + O(|A|),$$

where the error term comes from the largest (unpaired) value of i. Actually, we can say more:

$$\sum_{m\in\mathcal{M}} \left|\mathcal{L}_m^{A\times B}\right|^2 \asymp \sum_i \left|\mathcal{L}_{m_i}^{A\times B}\right| \cdot \left|\mathcal{L}_{m_{i+1}}^{A\times B}\right|.$$

This is because we already have a lower bound on the sum which is larger than the error term of $O(|A|)$.

We have thus obtained an interpretation of the sum from (17): it counts (roughly) the number of points in the union (18). Since this union is clearly a subset of $(A\times B)+(A\times B)$, we deduce that

$$\begin{aligned}\sum_{m\in\mathcal{M}} \left|\mathcal{L}_m^{A\times B}\right|^2 &\ll \sum_i \left|\mathcal{L}_{m_i}^{A\times B}\right| \cdot \left|\mathcal{L}_{m_{i+1}}^{A\times B}\right| \\ &= \left|\bigcup_i \left(\mathcal{L}_{m_i}^{A\times B} + \mathcal{L}_{m_{i+1}}^{A\times B}\right)\right| \\ &\leq |(A\times B)+(A\times B)| \\ &= |(A+A)\times(B+B)| \\ &= |A+A|\cdot|B+B|.\end{aligned}$$

Combining this with (16) and (17) yields the bound claimed in Solymosi's theorem. □

5. Related problems

Despite a lot of progress, the Erdős-Szemerédi conjecture remains wide open. There are also many related open problems. One recent example stems from a theorem originally due to Green, which asserts that every sufficiently large subset of $\mathbb{Z}/p\mathbb{Z}$ is of the form $A+A$ for some $A \subseteq \mathbb{Z}/p\mathbb{Z}$. The precise meaning of 'sufficiently large' remains an interesting open question (see [Al] for the strongest results).

Also, two key steps in Solymosi's proof are statements about moments:

$$\sum_{x \in AB} \rho_{AB}(x) = \sum_{m \in B/A} \rho_{B/A}(m)$$

and

$$\sum_{x \in AB} \rho_{AB}(x)^2 = \sum_{m \in B/A} \rho_{B/A}(m)^2$$

Can one generalize such relations, either to other operations or to higher moments? And what would such information give?

References

[Al] N. Alon, *Large sets in finite fields are sumsets* J. Number Theory **126** (2007), no. 1, 110-118. 11

[Ch] M.-C. Chang, *The Erdős-Szemerédi problem on sum set and product set*, Ann. of Math. **157** (2003), 939-957. 4

[ER] G. Elekes and I. Z. Ruzsa, *Few sums, many products*, Studia Sci. Math. Hungar. 40 (2003), 301-308. 4

[ES] P. Erdős and E. Szemerédi, *On sums and products of integers*, Studies in Pure Mathematics, Birkhäuser, Basel, 1983, 213-218. 2

[Fo] K. Ford, *The distribution of integers with a divisor in a given interval*, Ann. of Math. **168** (2008), 367-433. 4

[Fr] G. A. Freiman, *Elements of a Structural Theory of Set Addition*, Amer. Math. Soc., Providence, RI (1973) [English translation]. 3

[KR] S. V. Konyagin and M. Rudnev, *On new sum-product type estimates*, preprint available on the arXiv. 4

[MV] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory: I. Classical Theory*, Cambridge Studies in Advanced Mathematics **97**, Cambridge University Press, Cambridge (2007). 5, 7

[Ru] I. Z. Ruzsa, *Generalized arithmetical progressions and sumsets*, Acta Math. Hungar. 65 (1994), no. 4, 379-388. 3

[So] József Solymosi, *Bounding multiplicative energy by the sumset*, Adv. in Math. 222 (2009), 402-408. 4

Department of Mathematics, University of Toronto, Toronto, ON, Canada

E-mail address: leo.goldmakher@utoronto.ca