Sum–product growth

Talk by Dima Zakharov

Notes by Sanjana Das

September 20, 2024

§1 Introduction

We'll talk about the following theorem: Given a set of integers A and an integer k, we can consider its iterated sumset and iterated product set

$$kA = \underbrace{A + A + \cdots + A}_{k \text{ copies}} \quad \text{and} \quad A^k = \underbrace{A \cdot A \cdot \cdots \cdot A}_{k \text{ copies}}.$$

The theorem states that one of these has to be very large.

Theorem 1.1 (Bourgain–Chang 2004)
For any $A \subseteq \mathbb{Z}$ and $k \in \mathbb{N}$, either $|kA| \geq |A|^b$ or $\left|A^k\right| \geq |A|^b$, for some $b = b(k) \to \infty$ as $k \to \infty$.

Why is this interesting? Let's consider the standard sum-product theorem.

Theorem 1.2 (Sum-product)
For any A, we have $|A + A| + |A \cdot A| \geq |A|^{1+c}$ (for some constant $c > 0$).

This was first proven by Erdős and Szemerédi for some c in the 1980s. Then there were several improvements; a couple of landmark results were $c = \frac{1}{4}$ (Elekes) and $c = \frac{1}{3}$ (Solymosi), and the current record is $\frac{1}{3} + \varepsilon$ for some ε which is a rational number whose numerator and denominator are at most 10000 (this comes from work of Shkredov and others).

You can imagine taking this bound and applying it a few times, and maybe you'd expect to get a result like the one in Theorem 1.2. But the problem is that when you iterate, you will get $A + A + A + A$ and $A \cdot A \cdot A \cdot A$, but you'll also get things like $(A + A) \cdot (A + A)$, and it's not clear what to do with this — it's not clear that if this set is large, then one of the first two sets is also large. (In general, it's easier for sets to expand when we take both products and sums, compared to when we just take one.) So we need to do something more interesting.

Another interesting thing is that these sum-product bounds were proved using geometry. For example, you can prove $c = \frac{1}{4}$ using Szemerédi–Trotter on the grid $(A + A) \times (A \cdot A)$ — it turns out that if $A + A$ and $A \cdot A$ are small, then you get lots of incidences between this grid and lines of a certain form. But geometric methods can't really handle the growing exponents in Theorem 1.1 — you can use the plane to prove things involving $A \cdot A$, for instance, but beyond that, geometry doesn't really help you.

Today Dima will tell us about some of the proof of Theorem 1.1; there'll be some piece that we can't prove, but we'll talk about it. And there have also been some interesting developments around it in the last year (related to polynomial Freiman–Ruzsa, which was proven over finite fields last year).

§2 Proof of Theorem 1.1

Now we'll start talking about how we prove Theorem 1.1.

§2.1 Energies

The first thing we need is some results about *energies* — often, the way we prove that a sumset is large is by upper-bounding its energy. The standard energy for sumsets is defined as

$$\mathsf{E}(A) = \#\{a + b = a' + b' \mid a, b, a', b' \in A\}.$$

This thing is useful because by Cauchy–Schwarz, we have

$$|A + A| \geq \frac{|A|^4}{\mathsf{E}(A)}.$$

And this generalizes to longer sumsets as well — we can define the generalized energy

$$\mathsf{E}_k(A) = \#\{a_1 + \cdots + a_k = b_1 + \cdots + b_k \mid a_i, b_i \in A\}.$$

Then by Cauchy–Schwarz, we have

$$|kA| \geq \frac{|A|^{2k}}{\mathsf{E}_k(A)}.$$

As a sanity check, what range does $\mathsf{E}_k(A)$ fall into? We can imagine taking $a_i = b_i$ for all i, and then this equality will always hold; this gives the lower bound $\mathsf{E}_k(A) \geq |A|^k$. Meanwhile, if you fix any $2k - 1$ of the variables, then the last is determined, so $\mathsf{E}_k(A) \leq |A|^{2k-1}$. (And equality holds for an arithmetic progression, since then the last determined variable will typically live in the progression as well.) So we get

$$|A|^k \leq \mathsf{E}_k(A) \leq |A|^{2k-1}.$$

§2.2 Energies with multiple sets

We'll obtain the first case of Theorem 1.2 by proving upper bounds on this energy. But this will involve some sort of induction, and for that we'll need a slight generalization of this energy — where we have more than one set, and we require that each element belongs to its corresponding set.

Definition 2.1. For sets $A_1, \ldots, A_{2k} \subseteq \mathbb{Z}$, we define

$$\mathsf{E}_{2k}(A_1, \ldots, A_{2k}) = \#\{a_1 + \cdots + a_k = a_{k+1} + \cdots + a_{2k} \mid a_i \in A_i\}.$$

(This quantity naturally pops up when you do the arguments.)

First, there's an easy bound which helps us to not deal with this complicated quantity.

Proposition 2.2

For any sets $A_1, \ldots, A_{2k} \subseteq \mathbb{Z}$, we have

$$\mathsf{E}_k(A_1, \ldots, A_{2k}) \leq \prod_i \mathsf{E}_k(A_i)^{1/2k}.$$

This says that the energy of a tuple is at most the geometric mean of its individual energies — so if we just want an upper bound, then we can split this complicated object $\mathsf{E}_k(A_1, \ldots, A_{2k})$ into terms involving individual sets.

You might guess that this proof is some sort of Hölder, and indeed it is.

Proof. First we'll write the relevant energies as integrals. For each A_i, we consider its Fourier transform

$$f_i(\theta) = \sum_{a \in A_i} e^{2\pi i \theta a}.$$

(This is a function on the torus $\mathbb{R}/\mathbb{Z}$.)

Claim 2.3 — We can write $\mathsf{E}_k(A_i) = \int |f_i(\theta)|^{2k} \, d\theta$.

Proof. Imagine we take $|f_i(\theta)|^{2k}$ and expand it out. We have

$$|f_i(\theta)|^{2k} = \underbrace{f_i(\theta) \cdots f_i(\theta)}_{k} \cdot \underbrace{\overline{f_i(\theta)} \cdots \overline{f_i(\theta)}}_{k},$$

where each of these terms is a big sum; and if we pick a term from each, we'll get something of the form

$$e^{2\pi i \theta (a_1 + \cdots + a_k - a_{k+1} - \cdots - a_{2k})}.$$

Then when we integrate over θ, we'll get 0 unless $a_1 + \cdots + a_k - a_{k+1} - \cdots - a_{2k} = 0$, in which case we'll get 1. So we collect 1 exactly for each $2k$-tuple contributing to the energy $\mathsf{E}_k(A_i)$. □

Similarly, we can write

$$\mathsf{E}_k(A_1, \ldots, A_{2k}) = \int f_1 \cdots f_k \overline{f_{k+1} \cdots f_{2k}} \, d\theta.$$

And then we just use Hölder to finish. □

§2.3 An upper bound on energies

The next observation is what'll actually help us upper-bound $\mathsf{E}_k(A)$. We imagine starting with a set $A \subseteq \mathbb{Z}$, and then picking a prime p and decomposing A based on the power of p in each element. Then it turns out that we can bound $\mathsf{E}_k(A)$ in terms of the energies of these smaller sets.

Notation 2.4. We use $\nu_p(a)$ to denote the maximum j such that $p^j \mid a$.

Proposition 2.5

Let $A \subseteq \mathbb{Z}$, and let p be prime. Let $A_j = \{a \in A \mid \nu_p(a) = j\}$ for each j. Then we have

$$\mathsf{E}_k(A)^{1/k} \leq \binom{2k}{2} \sum \mathsf{E}_k(A_j)^{1/k}.$$

So the $\frac{1}{k}$th powers of the kth energies are kind of additive when we do such a decomposition.

Proof. As before, we define

$$f_j(\theta) = \sum_{a \in A_j} e^{2\pi i \theta a}.$$

We also set $f = \sum f_j$, so that

$$\mathsf{E}_k(A) = \int |f|^{2k} \, d\theta.$$

So we want to bound this integral.

The idea is that you can again expand out the brackets in $|f|^{2k}$ (each f is a sum of the f_j), and you'll get a sum over all possible ways to choose a term f_j from each summand — so we get

$$\mathsf{E}_k(A) = \sum_{i_1, \ldots, i_{2k}} \int f_{i_1} \cdots f_{i_k} \overline{f_{i_{k+1}} \cdots f_{i_{2k}}} \, d\theta. \tag{2.1}$$

The key observation is that a lot of these terms vanish.

Claim 2.6 — If all the indices $i_1, \ldots, i_{2k}$ are distinct, then $\int f_{i_1} \cdots f_{i_k} \overline{f_{i_{k+1}} \cdots f_{i_{2k}}} \, d\theta = 0$.

Proof. Imagine we take this product and expand out each f_{i_j}, and look at one of the resulting terms — this means we pick some $a_j \in A_{i_j}$ for each j, and we get

$$e^{2\pi i \theta(a_1 + \cdots + a_k - a_{k+1} - \cdots - a_{2k})}.$$

And each of these a_j contains a different power of p (i.e., the values of $\nu_p(a_j)$ are all distinct). In particular, there's no way for the sum to be 0 — if we consider the a_j divisible by the smallest power of p, then since all the remaining terms are divisible by a larger power of p, there's no way to cancel it out. (For example, if a_1 is coprime to p and $p \mid a_j$ for all other j, then we get something of the form $a_1 + p \cdot \bullet \neq 0$.)

And if this sum isn't 0, then the integral cancels out, as desired. □

So it turns out that in the long sum (2.1), most things don't actually contribute; and we can reduce it to a sum over all tuples where two indices coincide. The $\binom{2k}{2}$ factor in the bound corresponds to choosing *which* two terms coincide; so we'll write in this factor and assume that the first two sets coincide (this is cheating slightly); then we get

$$\mathsf{E}_k(A) \leq \binom{2k}{2} \sum \int f_i^2 f_{i_3} \cdots f_{i_k} \overline{f_{i_{k+1}} \cdots f_{i_{2k}}} \, d\theta.$$

We're going to eventually apply Hölder, but before that, we can bring part of the sum inside — if we leave the sum over i (the common value of $i_1 = i_2$) on the outside, but pull everything else back in, then we get

$$\mathsf{E}_k(A) \leq \binom{2k}{2} \sum_i \int f_i^2 f^{2k-2} \, d\theta. \tag{2.2}$$

(The point is that the other indices aren't constrained by anything, so when we sum over then, we just get f back. Some terms might need conjugates, but this doesn't affect the proof.)

And now Hölder looks promising. We know that

$$\mathsf{E}_k(A_i) = \int |f_i|^{2k} \, d\theta,$$

so we want to turn the square in (2.2) into this exponent of $2k$. So we use Hölder to say

$$\mathsf{E}_k(A) \leq \binom{2k}{2} \sum_i \left(\int |f_i|^{2k} \, d\theta \right)^{1/k} \left(\int |f|^{2k} \, d\theta \right)^{(k-1)/k}.$$

And this is great, because the final term is just $\int |f|^{2k} \, d\theta = \mathsf{E}_k(A)$, so we get

$$\mathsf{E}_k(A) \leq \binom{2k}{2} \sum_i \mathsf{E}_k(A_i)^{1/k} \mathsf{E}_k(A)^{(k-1)/k}.$$

And cancelling out $\mathsf{E}_k(A)$ from both sides gets the desired bound. □

§2.4 Skew dimension

Let's try to understand what Proposition 2.5 means for us. We have a lower bound on $|kA|$ in terms of the energy of A, and we now have an iterated upper bound on this energy which depends on the prime factors of elements of A. Prime factors have something to do with multiplication, so this is kind of going in the right direction.

Question 2.7. What's the best bound we can get by iterating Proposition 2.5?

For example, if we could find a prime p such that all powers of p in A were distinct, then Proposition 2.5 would give a great upper bound.

The correct notion here is *skew-dimension*. First, suppose we have some $a \in A$ with prime factorization $a = p_1^{t_1} \cdots p_D^{t_D}$. Then we can associate to it the vector $(t_1, \ldots, t_D) \in \mathbb{Z}^D$. This lets us turn A into a subset of a large-dimensional integer grid.

Definition 2.8. We define $\Pi(A) \subseteq \mathbb{Z}^D$ as the set of vectors $(t_1, \ldots, t_D)$ over all $a = p_1^{t_1} \cdots p_D^{t_D} \in A$.

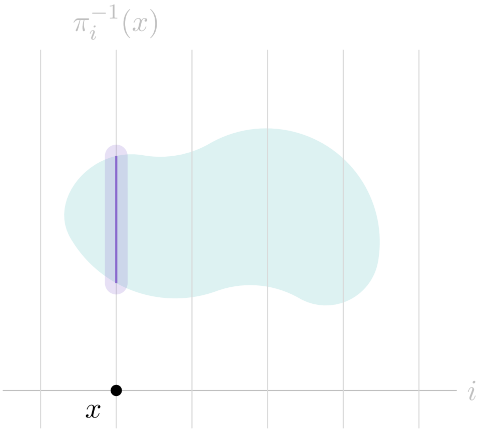
This is convenient because we're interested in the functions ν_p, which correspond to projecting onto certain coordinates — ν_{p_i} corresponds to projecting onto the ith coordinate.

Definition 2.9. Given a set $B \subseteq \mathbb{Z}^D$, we define its skew-dimension, denoted $\dim_*(B)$, as follows:

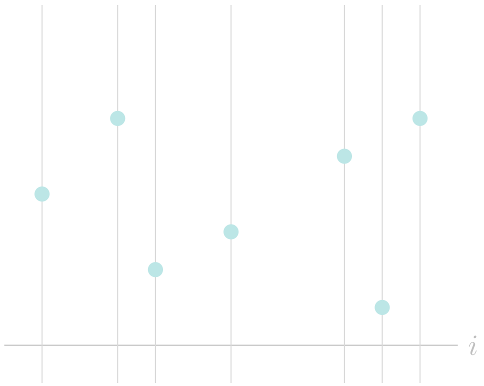
(i) $\dim_*(B) = 0$ if $|B| = 1$.

(ii) $\dim_*(B) \leq r$ if there exists a coordinate i such that $\dim_*(B \cap \pi_i^{-1}(x)) \leq r - 1$ for all x.

So we have a set B, and to say that $\dim_*(B) \leq r$, we're trying to find a coordinate i such that all the slices in the direction corresponding to i have dimension at most $r - 1$.



For example, a set has skew dimension 1 if there exists a projection where all its coordinates are distinct.



And it has dimension 2 if we can find some projection i such that each fiber has dimension 1, meaning that for each fiber, there is some other projection under which all its coordinates are distinct. Note that each of these fibers could have its own direction. For example, suppose that $\dim_*(A) = 2$, and the first projection is π_1. So when we project in this direction, we get a collection of coordinates x. And for each, if we define $A_x = \pi^{-1}(x) \cap A$, then this set has dimension 1 — i.e., there is some coordinate i_x such that $\pi_{i_x} \colon A_x \to \mathbb{Z}$ is injective. (The point is that i_x depends on x.) The word *skew* refers to this — if the coordinates were required to be the same for different values of x then we'd get a genuine dimension with respect to coordinate subspaces, but the skew version is a bit more flexible.

Another way to think about this is query complexity — how many coordinates you need to look at in order to reconstruct an element of your set A (where you're allowed to pick one direction and ask for the coordinate of a in that direction; you can look at the answer you get and then pick the next direction to ask about; and so on).

Then by iterating Proposition 2.5 appropriately, we get the following bound.

Proposition 2.10

Let $A \subseteq \mathbb{Z}$ be such that $\dim_* \Pi(A) = d$. Then

$$\mathsf{E}_k(A)^{1/k} \le \binom{2k}{2}^d |A|\,.$$

So we get a bound which is something exponential in d, times the size of A (k is a constant).

Proof. The idea is to look at A and choose a prime p corresponding to (ii) in the definition of skew dimension — i.e., we pick p such that

$$\dim_* \Pi(A_j) < \dim_* A$$

for all j (where the sets A_j are defined as in Proposition 2.5). Then we can plug this into Proposition 2.5 and use induction to bound each individual term $\mathsf{E}_k(A_j)$; this gives

$$\mathsf{E}_k(A)^{1/k} \le \binom{2k}{2} \sum_j \mathsf{E}_k(A_j)^{1/k} \le \binom{2k}{2} \sum_j \binom{2k}{2}^{d-1} |A_j|\,.$$

And $|A| = \sum_j |A_j|$ (because A is a disjoint union of the A_j), so the right-hand side is precisely $\binom{2k}{2}^d |A|$. □

§2.5 Skew-dimension vs. additive structure

Now the moral is that in order to prove something like Theorem 1.1, we're interested in understanding the skew-dimension of $\Pi(A)$; specifically, we want to relate this skew-dimension to the multiplicative expansion of A (we want to say that if the skew-dimension is large, then A has lots of multiplicative expansion).

Letting $B = \Pi(A) \subseteq \mathbb{Z}^D$, we're in a setting where we know $\dim_* B$ is large, and we'd like to show that B doesn't have too much *additive* structure — i.e., that $|B + B| \gg |B|$ (where how much larger depends on how large the skew-dimension is).

> **Remark 2.11.** Note that we have addition rather than multiplication here because taking a product of numbers corresponds to adding the exponents in their prime factorizations.

We'll state the precise result of this form, and then see how this result and what we've discussed so far piece together to prove the theorem.

> **Theorem 2.12** (Pálvölgyi–Zhelezov 2017)
> Let $A \subseteq \mathbb{Z}^D$ be such that $|A + A| \leq K\,|A|$. Then there exists a dense subset of A with small skew-dimension — more precisely, there exists $A' \subseteq A$ with $|A'| \geq K^{-C}\,|A|$ and $\dim_*(A') \leq C \log K$.

(This is a precise version of the goal we just stated, with reversed quantifiers; the set A here corresponds to B in our discussion.)

If you've seen the weak Freiman–Ruzsa conjecture, this statement is very similar, but with $\dim_*$ in place of dim; Dima will try to explain the connections between them later. But first, Theorem 2.12 is precisely what we need to finish the proof of Theorem 1.1; we'll now talk about this argument.

§2.6 Proof of Theorem 1.1

We have a set $A \subseteq \mathbb{Z}$; suppose we know that $|A^k| \leq |A|^b$ (where b is some small but growing power that we'll specify later).

The first step is a pigeonhole argument — let $B_j = A^{2^j}$ be the product of 2^j copies of A. Then $B_{j+1} = B_j \cdot B_j$, and if we only consider $j \leq \log k$, then the last one has to be smaller than A^k. So we can pigeonhole a step where the set doesn't grow too much — we can find some ℓ for which

$$|B_\ell \cdot B_\ell| \leq |A|^\varepsilon\,|B_\ell|\,.$$

For what the parameters are, we have $\log k$ steps and a total increase of $|A|^b$, so to make the pigeonhole work out we want $b < \varepsilon \log k$; we'll specify ε later.

Now by Theorem 2.12, taking $K = |A|^\varepsilon$, we can find $S \subseteq B_\ell$ with $|S| \geq K^{-C}\,|B_\ell|$ and $\dim_* \Pi(S) \leq C \log K$.

The idea is that if we knew the dimension of A itself was at most $C \log K$, then we'd be happy — we could just apply Proposition 2.10 and we'd be done. And the point is that this is more or less true — S is obtained from A by taking some iterated product set and then passing to some dense subset, so intuitively it should remember a lot of structure from A. And you can prove this using a bit of additive combinatorics — using the fact that $|A \cdot S| \leq K^C\,|S|$ and some additive combinatorics, you can show there exists x such that

$$|xS \cap A| \geq K^{-C}\,|A|\,.$$

Then we can look at this large piece of A, which we denote by $A' = xS \cap A$. We know

$$\dim_* A' \leq \dim_* xS = \dim_* S \leq C \log K \leq C\varepsilon \log |A|\,.$$

And now we can apply Proposition 2.10; this tells us

$$\mathsf{E}_{k'}(A')^{1/k'} \leq \binom{2k'}{2}^{C\varepsilon \log|A|} |A'| \tag{2.3}$$

(for any k' — we're going to prove an upper bound on $\mathsf{E}_{k'}(A')$ for some $k' \ll k$, which will give a lower bound on $|k'A'|$ and therefore $|kA|$).

To get a lower bound on $|k'A|$ with growing exponent, all we want is for $\mathsf{E}_{k'}(A)$ to be quite a bit smaller than the maximum possible — for example, we'll be happy if

$$\mathsf{E}_{k'}(A') < |A'|^{3k'/2}.$$

(The same would work if we replaced $\frac{3}{2}$ with any number between 1 and 2.) To get this from (2.3), we want

$$|A'|^{C\varepsilon \log k'} < |A'|^{1/2}.$$

And we can achieve this by taking $\varepsilon \ll \frac{1}{\log k'}$. (We want k' to grow slowly compared to k so that $b \approx \varepsilon \log k$ grows with k.)

Once we do this, we get $|kA| \geq |k'A'| > |A'|^{k'/2}$, which concludes the proof.

So basically, there's a neat pigeonholing argument which sets you in a position to apply Theorem 2.12; then you get a dense set A' with very little additive structure, and this implies the bound we want.

§3 More about Theorem 2.12

Now the only mysterious bit of the proof is Theorem 2.12. There are two (maybe three) proofs of this theorem.

The proof in the Pálvölgyi–Zhelezov paper is fairly combinatorial. It uses some nontrivial additive inequality, whose proof is some induction where you look at numbers carefully and they work out (this requires being careful somehow).

Then the result was reproven in a paper about entropy methods in combinatorics by Green, Manners, and Tao — they gave a proof of Theorem 2.12 using entropy. On a basic level, you have a set A, and you're looking at some projection, so you have a bunch of fibers A_x. And you want to do induction here — applying the inductive hypothesis to each fiber. So you look at how additively structured the 'base' is (the projection onto the ith coordinate, i.e., the set of all values of x). And you use the fact that if we take two fibers A_x and A_y, then $A_x + A_y$ lives above $x + y$. And you can combine these bounds together in some way; but the proof seems a bit delicate, and you have to do some computation.

But there's actually a nice theorem which says you can replace the notion of skew-dimension with an *actual* dimension.

Definition 3.1. For a set $A \subseteq \mathbb{Z}^D$, we define its dimension, denoted $\dim A$, as the dimension of its affine span (as a vector space).

Theorem 3.2 (Weak PFR, Green–Gowers–Manners–Tao)

Let $A \subseteq \mathbb{Z}^D$ be such that $|A + A| \leq K\,|A|$. Then there exists $A' \subseteq A$ of size $|A'| \geq K^{-C}\,|A|$ and such that $\dim A' \leq C \log K$.

The moral of this is that all small-doubling sets come from low-dimensional subspaces. The $\log K$ is natural because we can imagine taking A to be a box $[n]^d$; then its doubling is roughly 2^d. So in this case, we get a log dependency between the dimension and doubling.

Remark 3.3. PFR stands for the polynomial Freiman–Ruzsa conjecture. Strong PFR would state that if you have small doubling, then you actually look like such a box (more precisely, you contain a dense subset of the box).

This result is tricky and uses entropy methods, but it's a good thing to know.

§4 Theorem 1.1 over $\mathbb{R}$

Finally, Theorem 1.1 was about subsets of the integers; what if we instead consider subsets of the reals?

Our proof quite strongly used the fact that our numbers were integers, so it's unclear if you can modify the argument to make it work over reals. But you can do something — Bourgain and Chang showed that if we work over a finite-degree field extension of $\mathbb{Q}$, then the argument still works.

Theorem 4.1 (Bourgain–Chang)

If $A \subseteq \mathbb{F}$ for a field $\mathbb{F}$ with $[\mathbb{F}/\mathbb{Q}] = d$, then Theorem 1.1 still holds, where b also depends on d.

In this setting, instead of primes, you can use prime ideals. This *almost* works — in our argument we didn't care how many primes there were, and most of what we did with primes can also be done with prime ideals. But you get into trouble with units — what if A is a subset of the set of units of your field? Prime ideals can't distinguish units between each other, so for this the mod-p trick is useless. So in this case, you need some actual number theory to deal with units — a statement that if A is a set of units, then it has to have little additive structure.

For a while it wasn't known if you can extend Theorem 1.1 to all subsets of $\mathbb{R}$ (without a dependence on degree). But now we know you can.

Theorem 4.2 (Mudgal, GGMT, PZh)

Theorem 1.1 holds for any $A \subseteq \mathbb{R}$.

The proof of this heavily uses weak PFR (Theorem 3.2) together with a very deep result from algebraic number theory. We will just state this result, because it's an amazing theorem. It has several versions; here's one.

Theorem 4.3 (Subspace theorem, Evertse–Schlickewei–Schmidt)

Let $\Gamma \subseteq \mathbb{C}^\times$ be a rank-d (multiplicative) subgroup. Let $\ell \geq 1$ and $c_1, \dots, c_\ell \in \mathbb{C}$. Then

$$\#\{c_1\gamma_1 + \cdots + c_\ell\gamma_\ell = 1 \mid \gamma_i \in \Gamma, \text{ no subsums are } 0\} \leq e^{C_\ell d}.$$

A rank-d multiplicative subgroup means that we take d complex numbers and generate all possible products and ratios using them. (This is what you get when you apply weak PFR to a small-doubling subspace.)

Here we're counting solutions to the linear equation $c_1\gamma_1 + \cdots + c_\ell\gamma_\ell = 1$ (such that no smaller pieces sum to 0). It's not even clear that this count is finite, since Γ is infinite. But the subspace theorem says it's bounded, and there is a really good upper bound on it — one that's exponential in the rank of Γ. (And the dependence on ℓ is not that bad either — you can take $C_\ell \leq \ell^6$.)

If you combine the subspace theorem and weak PFR, you more or less immediately get Theorem 1.1. We start with a set $A \subseteq \mathbb{C}$. If A has small multiplicative doubling (i.e., $|A \cdot A| \leq K\,|A|$ — we can get to such a situation in the same way as in our proof of Theorem 1.1), then we can basically assume that A lives in a subgroup with rank $\log K$, using weak PFR. Then we can apply the subspace theorem to upper-bound the energy — this tells you that the number of energy tuples is at most exponential in the rank, and the rank is $\log K$, so we get the bound we need.

But the subspace theorem is really hard.