BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 56, Number 1, January 2019, Pages 29–64
https://doi.org/10.1090/bull/1648
Article electronically published on October 1, 2018

NEW APPLICATIONS OF THE POLYNOMIAL METHOD: THE CAP SET CONJECTURE AND BEYOND

JOSHUA A. GROCHOW

ABSTRACT. The cap set problem asks how large can a subset of $(\mathbb{Z}/3\mathbb{Z})^n$ be and contain no lines or, more generally, how can large a subset of $(\mathbb{Z}/p\mathbb{Z})^n$ be and contain no arithmetic progressions. This problem was motivated by deep questions about structure in the prime numbers, the geometry of lattice points, and the design of statistical experiments. In 2016, Croot, Lev, and Pach solved the analogous problem in $(\mathbb{Z}/4\mathbb{Z})^n$, showing that the largest set without arithmetic progressions had size at most c^n for some $c < 4$. Their proof was as elegant as it was unexpected, being a departure from the tried and true methods of Fourier analysis that had dominated the field for half a century. Shortly thereafter, Ellenberg and Gijswijt leveraged their method to resolve the original cap set problem. This expository article covers the history and motivation for the cap set problem and some of the many applications of the technique: from removing triangles from graphs, to rigidity of matrices, and to algorithms for matrix multiplication. The latter application turns out to give back to the original problem, sharpening our understanding of the techniques involved and of what is needed for showing that the current bounds are tight. Most of our exposition assumes only familiarity with basic linear algebra, polynomials, and the integers modulo N.

1. Introduction

The published proof [39] of the Cap Set Conjecture is so elegant, elementary, and short—and others have already provided expositions of it [63, 130, 138]—that we can hardly do better here. Although we'll include a quick proof in Section 2, our main purpose here is to provide motivation for the conjecture, put it in its proper historical context, and discuss some of the consequences of the new technique.

The Cap Set Conjecture says that the largest subset of $(\mathbb{Z}/3\mathbb{Z})^n$ which contains no lines—that is, no three points x, y, z such that $x + y = 2z$, or equivalently (mod 3), $x + y + z = 0$—has size at most c^n for some c strictly less than 3.

If you are an (additive) combinatorialist, you may find this problem intrinsically interesting and immediately fall in love with it. For the rest of us, however, it is natural to wonder how one arrives at this conjecture. Why were people studying such questions in the first place? What connections does it have to other areas of mathematics? As is often the case in mathematics, if we take the time to get to know the problem a bit better—take it to dinner, ask about its history, its family, what kind of recreational activities it enjoys—we find that these questions have good answers, and we come to appreciate a problem whose upbringing is perhaps not so much like our own. My main goal in this exposition is to share some of the

Received by the editors April 22, 2018.
2010 *Mathematics Subject Classification.* Primary 11B25, 51E22.

answers to these questions. And, okay, sure, I'll show you the proof, too—how can I resist?

One final note before diving in: I have attempted to make this exposition accessible to as wide an audience as possible, starting from the early undergraduate level. As this may well include students who haven't seen group theory before, or have little exposure to combinatorics, I will spell out (nearly) all the details for expository purposes, including some combinatorial arguments that are easy exercises once you've seen a bit of combinatorics. We ask for a little patience from our readers for whom such arguments are too standard to be worth writing down. In the few cases where I wanted to make some remark that was difficult to make under these conditions, footnotes are there to help the reader along; if such remarks confuse you, you ought to be able to skip over them and still understand the rest of the article. Conversely, if such footnotes annoy you, you ought to be able to skip over them without losing much. But my hope is that the footnotes will help entice readers with less background to learn new and exciting things!

1.1. **A frivolous and fun motivation.** The popular imagination is perhaps drawn to this problem because of its connection with the card game SET®. (There are other expositions which use the game of SET to introduce readers to some of the mathematics we consider here, such as [29, 96].) The SET deck consists of 81 cards, which are illustrated in Figure 1.

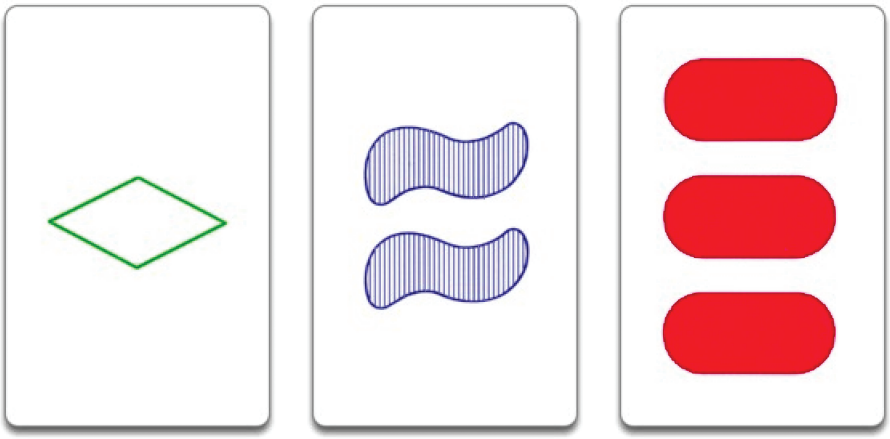

FIGURE 1. A Set in the popular card game.

Each card has four attributes, each of which can take three values: color (values: red, green, purple), shape (diamond, oval, squiggle), number (1,2,3), and fill (solid, shaded, open). A "Set" is a collection of three cards such that, in each attribute, either all cards have the same value, or all cards have distinct values. Figure 1 shows a Set in which all the cards are distinct in each attribute (they have three different colors—or, if you're reading this in black and white, are all the same color—three different shapes, three different numbers, and three different fills). As another example, the cards (red,diamond,1,solid), (red,diamond,2,shaded), (red,diamond,3,open) also form a Set. Twelve cards are laid face up, and players compete to find Sets as fast as possible. When the players agree there is no Set on the table, three more

cards are laid face up until someone finds a Set. This raises the natural question: How many cards can be on the table with no Set?

If we identify the values of each attribute with the elements of the integers mod 3, then each card corresponds to a point in $(\mathbb{Z}/3\mathbb{Z})^4$, and a Set is precisely a collection of three points $\vec{x}, \vec{y}, \vec{z} \in (\mathbb{Z}/3\mathbb{Z})^4$ such that $\vec{x} + \vec{y} + \vec{z} \equiv \vec{0} \pmod 3$ (we leave this as an exercise for the reader). Since $2 \equiv -1 \pmod 3$, this is the same as saying $\vec{y} - \vec{x} \equiv \vec{z} - \vec{y} \pmod 3$; in other words, the points $\vec{x}, \vec{y}, \vec{z}$ lie on a line. Our question thus becomes: How large can a subset of $(\mathbb{Z}/3\mathbb{Z})^4$ be and still contain no line? If we generalize this from dimension 4 to arbitrary dimension n, we get the cap set problem.

In dimension 4, the answer turns out to be 20 cards [103]. Note that this answer was found three years before Marsha Jean Falco invented the game (which she did to help her visualize the combinatorics of genes related to epilepsy in German Shepherds [119]) and almost 20 years before the game was made public. Why were people looking at such questions? Our next motivation is older and quite a bit deeper.

1.2. **Motivating additive combinatorics with the primes.** Additive combinatorics is the study of the interrelation between the additive and combinatorial structure of sets. Of course, this just begs the question of what we mean by *additive structure* in the first place. To give some meaning to the notion of additive structure and why it might be interesting, let's start with the example of the natural numbers $\mathbb{N} = \{1, 2, 3, \dots\}$. In terms of just addition, it seems pretty simple: Start from 1, and just keep adding 1. It looks like a (discrete) ray, heading off in one direction. In terms of just multiplication, it looks a bit more complicated, but not so much more: Every number can be written uniquely as a product of primes. This implies that the multiplicative structure of $\mathbb{N}$ looks like the additive structure of the set of sequences of elements of $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, only finitely many of which aren't zero (which we denote $\mathbb{N}_0^{<\infty}$): To any such sequence of natural numbers $(a_1, a_2, \dots, a_d, 0, 0, 0, \dots)$, we may associate the number $2^{a_1}3^{a_2}5^{a_3} \cdot p_d^{a_d}$ where p_d is the dth prime number. This identification shows that the multiplicative structure of $\mathbb{N}$ is the same as the additive structure of $\mathbb{N}_0^{<\infty}$, an infinite-dimensional grid, as regular as can be.

When we consider both the multiplicative and additive structure of $\mathbb{N}$ together though, something remarkable happens. Note that our notion of *size* in $\mathbb{N}$ is essentially additive: How many times you need to add 1 to get to a given number. Thus, the two simplest questions we can ask that mix the additive and multiplicative structures on $\mathbb{N}$ are the following.

(1) How big is the nth prime p_n? (Additive structure of a multiplicatively defined sequence)
(2) What is the factorization of n? (Multiplicative structure of an additively defined sequence)

These also turn out to be some of the deepest questions about the interaction between additive and multiplicative structure. The first will be traditionally recognized as deep, leading quickly to the Prime Number Theorem and the Riemann Hypothesis. The depth of the second question can already be glimpsed in the

techniques used in algorithms for factoring numbers [21, 88, 99, 105, 121] (see, e.g., [87, 107, 136] for surveys).

The algorithmic viewpoint adds evidence for the idea that it is the *mixture* of the additive and multiplicative structures that leads to complexity. When we represent numbers additively—that is, in a way that makes addition easy, say, in base 10—multiplying them is relatively easy but factoring them seems to be difficult, or at least a much deeper problem. But when we represent numbers multiplicatively—in a way that makes multiplication easy, say, using their prime factorizations—converting them to base 10 becomes as easy as multiplying them was before (just multiply the prime factors), but adding them becomes equivalent to factoring numbers written in base 10 [30]. The lesson is that, regardless of whether we view the multiplicative structure of $\mathbb{N}$ through an additive lens or its additive structure through a multiplicative lens, we run into the same, much deeper complexity than if we only considered one structure at a time.

We may also ask the next questions along these lines:

(1′) What is the additive gap $p_{n+1} - p_n$ between successive primes?
(2′) What is the relationship between the factorization of n and that of $n + 1$?

Only in 2013 was it shown that there is a universal constant C such that $p_{n+1} - p_n < C$ for infinitely many n [139]; C was eventually improved to 246 [95, 104], and the Twin Prime Conjecture is that C can be lowered all the way to 2. The relationship between the factorization of n and $n + 1$ leads [100] to notoriously difficult problems like the Collatz "$3n+1$" Conjecture, about which Erdős famously said, "Mathematics may not be ready for such problems" [72, p. 330] (see also [98]). Guy's chapter [72, Problem E16] and Lagarias's annotated bibliography [83] are excellent sources of references on this difficult problem.

Now, a zeroth-order heuristic for questions (1′) and (2′) is that the answer to both is essentially random: The factorizations of n and $n + 1$ are "independent" of one another, and the prime gap $p_{n+1} - p_n$ jumps around "randomly". There is some truth in this heuristic (see, e.g., [10, 45, 56]). But along with the randomness, there is also significant structure present, as evidenced already by the aforementioned results.

We may thus ask, for example, what further additive structure is there to the prime numbers? A natural generalization of (1′) is to ask for structure in the differences between several primes (not necessarily consecutive). Here, perhaps the simplest structure to ask for would be for a set of primes with a common distance between them, that is, which form an *arithmetic progression* $p, p + r, p + 2r, p + 3r, \ldots, p + dr$. A classical folklore conjecture, going back perhaps two centuries, is that the primes contain arithmetic progressions of every length—quite a lot of additive structure for a multiplicatively defined set!

Green and Tao [68] proved this conjecture in 2004, but we will see that its history provides motivation and impetus for many topics in additive combinatorics, including our main topic, the Cap Set Conjecture.

1.3. **Additive combinatorics more generally.** As a young boy, Erdős (re)proved that $\sum_{p \text{ prime}} 1/p = \infty$, and it has been postulated by several authors that this early exciting mathematical experience, in combination with the long-standing conjecture about arithmetic progressions in the primes, led to:

Conjecture 1.1 (Erdős, 1940s or 1950s, see [40–42, 122]).[1] *If $A \subseteq \mathbb{N}$ satisfies $\sum_{n\in A} \frac{1}{n} = \infty$, then A contains arbitrarily long arithmetic progressions.*

Apparently, Erdős particularly liked the idea that the reason the primes should contain arbitrarily long arithmetic progressions actually had little to do with the primes themselves—just a statement about their density. In his 1976 talk [41], Erdős offered $3000 USD for its resolution—the highest prize he had ever offered at that point—and in 1996 he upped the prize to $5000 USD [43], which I believe was his third-largest ever.[2] As pointed out by Soifer [122, p. 354], the high prize and the frequency with which he raised this conjecture in his talks and writings suggests it was one of his favorites.

Gowers [62] points out that Erdős's conjecture is "morally" about sets A such that the density of A in $\{1, \dots, N\}$ is around $1/\log N$. For if the density is $1/\log N$, then the sum diverges, while if the density is, say, $1/(\log N(\log\log N)^2)$, then it converges. Now, of course, arbitrary subsets need not have a density function that varies so smoothly as a function of N. But Erdős's conjecture is sandwiched between two statements about sets of a certain density. For if $\sum_{n\in A} 1/n$ diverges, then there are infinitely many N such that $|A\cap\{1, \dots, N\}| \geq 1/(\log N(\log\log N)^2)$, so to prove the conjecture, it suffices to show that subsets of density $1/(\log N(\log\log N)^2)$ contain arbitrarily long arithmetic progressions; to disprove the conjecture, it suffices to find a set of density $1/\log N$ that does not contain arithmetic progressions. Nonetheless, the conjecture is what it is. It remains open even to prove that a set A satisfying the hypothesis contains 3-term arithmetic progressions.

However, before Erdős put forth this conjecture, he was indeed thinking about the density of sets. In his 1936 paper with Turán [48], they conjectured:

[1] Regarding the date of this conjecture, the earliest written reference I could find for the case of 3-term arithmetic progressions was a 1973 seminar report [40, Conjecture 1.2], and for general arithmetic progressions was a talk from 1976 [41]. This conjecture is often attributed to Erdős and Turán's 1936 paper [48], but the conjecture does not appear there in print—even as a question—and in his later writings, including a touching tribute to Turán [41, p. 40], although Erdős raises the problem in connection with his work on the primes with Turán, he refers to it as an "old conjecture *of mine*" (emphasis added). Soifer [122, p. 355] found references in a 1982 talk in which Erdős says it was more than 40 years old, and a 1986 talk in which Erdős said it was "about 30 years ago," so we conclude with Soifer that the conjecture was made sometime between the early 1940s and mid-1950s.

Regarding its motivation from arithmetic progressions in the primes, we have at least the following evidence. In [40], Erdős relates it with Goldbach's Conjecture—which, in particular, implies that for any prime p, $2p = p_1 + p_2$ for some other primes p_1, p_2, and hence the primes have infinitely many 3-term arithmetic progressions—and Chowla's unconditional result [23] that there are infinitely many 3-term arithmetic progressions in the primes. (The earlier paper of van der Corput proving the same [134] was apparently forgotten until later.) In his 1977 paper [41] he writes down the conjecture as stated here, that is, for arbitrarily long arithmetic progressions, and points out that, in particular, it would imply arbitrarily long arithmetic progressions in the primes, and therefore resolving the conjecture should be quite hard. In 1981, he restates the conjecture yet again [42, p. 28], this time explicitly "in connection" with the problem of showing arbitrarily long arithmetic progressions in the primes.

[2] The only larger Erdős prizes I'm aware of are $10,000 USD to show that $p_{n+1} - p_n$ is "large" infinitely often, and $25,000 USD to show there are only finitely many consecutive pairs of primes p_n such that $p_n \leq \frac{p_{n+1}+p_{n-1}}{2}$, though he offered only $100 USD for a disproof; see [82] for details on these large prizes and a list of other Erdős prizes.

Conjecture 1.2 (Erdős and Turán [48])**.** *If $A \subseteq \mathbb{N}$ has positive upper density—that is, $\limsup_{N\to\infty} \frac{|A\cap\{1,\ldots,N\}|}{N} > 0$—then A contains infinitely many k-term arithmetic progressions, for every k.*

While the primes do not have positive upper density—indeed the Prime Number Theorem states that their density up to N is $\sim 1/\log N$, so their upper density is zero—this conjecture turned out to be a crucial step toward proving that there are arbitrarily long arithmetic progressions in the primes. Szemerédi [126] proved the Erdős–Turán Conjecture, and Szemerédi's proof was a crucial ingredient in Green and Tao's proof [68] that the primes contain arbitrarily long arithmetic progressions.

When encountering a difficult conjecture, two natural tactics are to consider special subcases or to consider analogous conjectures in slightly different settings. First, instead of arbitrarily long arithmetic progressions, let's ask for any nontrivial arithmetic progressions at all. By "nontrivial" we mean consisting of at least three distinct points. Okay, great: We'll focus on 3-term arithmetic progressions for a bit. Historically, this has indeed been a good place to start; for example, two decades before Szemerédi's Theorem was proved, in 1953 Roth [108] proved the analogous result for 3-term arithmetic progressions, by a significantly easier argument. (Szemerédi himself proved the $k = 4$ case in 1969 [125], before proving the general case.) In terms of arithmetic progressions in the primes, it's trivial to show existence of a 3-term arithmetic progression (3,5,7), and almost 70 years before Green and Tao, van der Corput [134] showed there were infinitely many 3-term arithmetic progressions in the primes. But even the existence of infinitely many 4-term arithmetic progressions in the primes—or even a *single* 24-term arithmetic progression—remained open until Green and Tao's result. Moreover, while 3-term arithmetic progressions have a nice formulation in terms of convolutions of Fourier transforms, 4-term arithmetic progressions do not; a conundrum which eventually led to higher-order Fourier analysis (see, e.g., [137] for a nice discussion of this difficulty and how it was overcome). Okay, fine, 4 is a lot harder than 3, so let's stick with 3-term arithmetic progressions.

And now, instead of only considering subsets of $\mathbb{N}$ (or $\mathbb{Z}$), let's consider subsets of arbitrary abelian groups.[3] These are, arguably, the most natural settings in which the notion of "additive structure" makes sense, since these are precisely the sets which have a notion of addition.[4] As a start, it should be clear that if we have $A \subseteq \mathbb{N}$ and we consider $A \cap \{1,\ldots,N\}$, then from the point of view of arithmetic progressions, this is essentially equivalent to considering A as a subset of $\mathbb{Z}/N\mathbb{Z}$, the integers mod N. If we're considering 3-term arithmetic progressions, maybe we should instead consider A as a subset of $\mathbb{Z}/2N\mathbb{Z}$, just to make sure there's no accidental wrapping around, but philosophically, and even mathematically, this turns out to make little difference. And note that we can indeed rephrase, for example, Szemerédi's Theorem in terms of $A \cap \{1,\ldots,N\}$: For all $k \in \mathbb{N}, \varepsilon > 0$, there exists N_0 such that if $N > N_0$ and $|A \cap \{1,\ldots,N\}| > \varepsilon N$, then A contains a

[3]If you don't know what an abelian group is, don't be scared! The integers $\mathbb{Z}$, the integers mod a number $\mathbb{Z}/N\mathbb{Z}$, and vectors of such $(\mathbb{Z}/m\mathbb{Z})^n$ are all abelian groups. In fact, these are almost the general case, so just keep these in mind as your examples, and you should have smooth sailing.

[4]Okay, technically maybe we should consider abelian *semi*groups. But many semigroups can naturally be embedded into groups, and those that can't have an addition operation which differs quite substantially from our intuition for addition: for example, if $x + x + x = x + x$ but $x \neq 0$, then our addition operation seems to be somewhat far from our main interest, namely $\mathbb{N}$.

k-term arithmetic progression. Okay, great, so other abelian groups might be good models for the phenomena we're interested in, but considering only cyclic groups (the integer mod N) seems to really be considering the *same* phenomena, rather than phenomena in analogous settings.

At the "opposite end" of some sort of natural spectrum, we might consider n-dimensional vectors over the integers modulo m, for small m and $n \to \infty$, such as $(\mathbb{Z}/2\mathbb{Z})^n, (\mathbb{Z}/3\mathbb{Z})^n, (\mathbb{Z}/4\mathbb{Z})^n$. To what extent can we use results about arithmetic progressions in these sets, which look more like vector spaces (and indeed, when m is prime, *are* vector spaces), to understand arithmetic progressions in $\mathbb{N}$?

There are two answers to this question, one historical and one formal. Historically, it's been useful to first consider the vector spaces $(\mathbb{Z}/p\mathbb{Z})^n$ (p prime), where we have lots of substructures to play with and induct on. To transfer results from this setting to $\mathbb{Z}/N\mathbb{Z}$, one fruitful approach is to use so-called Bohr sets, which are kind of an "approximate subspace" (in some Fourier-analytic sense). For any given result, however, this transference remains something of an art, but has been very effective in the past. The papers [64, 137] are entirely devoted to the analogy between vector spaces over finite fields and $\mathbb{Z}$.

Formally, there is also a tool for comparing the additive structure of subsets in one abelian group with subsets of another abelian group which I can't resist mentioning: Freiman homomorphisms [53] (for a textbook treatment, see, e.g., [132, Section 5.3]). For me, when I first learned of this notion, it helped clarify what is meant by additive structure in general; I hope it has the same effect for you. Let's define an *additive set* to be a subset A of an abelian group Z, and let's say that its "k-additive structure" is completely determined by all equalities of the form $a_1 + a_2 + \cdot + a_k = a_1' + a_2' + \cdot + a_k'$ where $a_i, a_i' \in A$. For example, the 2-additive structure captures any k-term arithmetic progressions: $a_1, a_2, a_3, \ldots, a_k$ form an arithmetic progression if and only if $a_{i+1} - a_i = a_{i+2} - a_{i+1}$ for all $i = 1, \ldots, k-2$, which we can rewrite entirely additively as $a_i + a_{i+2} = a_{i+1} + a_{i+1}$.

Definition 1.3 (Freiman homomorphism)**.** A *Freiman k-homomorphism* between additive sets $A \subseteq Z$ and $B \subseteq W$ is a function $f\colon A \to B$ such that, for all $a_i, a_i' \in A$

$$a_1 + \cdots + a_k = a_1' + \cdots + a_k \implies f(a_1) + \cdots + f(a_k) = f(a_1') + \cdots + f(a_k').$$

A *Freiman k-isomorphism* is a bijective Freiman k-homomorphism whose inverse is also a k-homomorphism; equivalently, we require f to be a bijection such that the one-way implication above becomes two-way: $a_1 + \cdots + a_k = a_1' + \cdots + a_k'$ if and only if $f(a_1) + \cdots + f(a_k) = f(a_1') + \cdots + f(a_k')$.

Although this is not how many of the results were transferred from finite fields to the integers, in principle this notion could let us transfer (some) results about arithmetic progressions in one abelian group, such as $(\mathbb{Z}/m\mathbb{Z})^n$, to another abelian group, such as $\mathbb{Z}$. At any rate, as I said, I think this notion helps clarify what we mean by additive structure.

1.4. **The Cap Set Conjecture.** Now, for arithmetic progressions, $(\mathbb{Z}/2\mathbb{Z})^n$ isn't interesting, as it only has trivial arithmetic progressions (of length 2): If x, y, z form an arithmetic progression, then $z = y + (y - x) = x$, because we're working modulo 2. So the simplest interesting "toy model" to consider in our family $(\mathbb{Z}/m\mathbb{Z})^n$ is $(\mathbb{Z}/3\mathbb{Z})^n$. By a similar argument, because $3 = 0$ in $(\mathbb{Z}/3\mathbb{Z})^n$, it has no arithmetic progressions of length larger than 3 (which are proper, that is, consisting of all

distinct elements). But that's alright, because 3-term arithmetic progressions were where we wanted to start anyway. So $(\mathbb{Z}/3\mathbb{Z})^n$ has the simultaneous virtues of (1) being analogous to, but not the same as, our original question(s) in $\mathbb{N}$, (2) among such analogous structures, being the simplest one which is still interesting; and (3) forcing us to focus our attention on the smallest case, namely that of 3-term arithmetic progressions, without having to worry about any "higher" additive structure.

Before getting into the Cap Set Conjecture itself, let's return to history to motivate it from two other angles. The first angle is geometric. Harborth [73] introduced the function $s(m, n)$ to be the smallest number s such that any s points in $\mathbb{Z}^n$ contains a subset of size m whose centroid also has integer coordinates. It is not hard to see that this is equivalent to the smallest s such that any sequence of s elements in $(\mathbb{Z}/m\mathbb{Z})^n$ contains a subsequence of length m whose sum is zero. For $m = 3$, note that this is nearly the same as the largest subset of $(\mathbb{Z}/3\mathbb{Z})^n$ which contains no 3-term arithmetic progression, since modulo 3 we have that x, y, z form an arithmetic progression if and only if $x + y = 2z$ if and only if $x + y + z = 0$ (the only difference is that $s(m, n)$ allows sequences with repeated elements). Alon and Dubiner [5, 6] were perhaps the first to raise this question for $m = 3$ explicitly, but their primary interest in those papers was the case of small n and large m, which is closer to our original motivation of $\mathbb{Z}/N\mathbb{Z}$ than to our new toy model of $(\mathbb{Z}/3\mathbb{Z})^n$. For general m, n, the current best bounds are still those due to Harborth [73]:

$$(m-1)2^n + 1 \leq s(m, n) \leq (m-1)m^n + 1.$$

(Note that the upper bound is larger than the size of $(\mathbb{Z}/m\mathbb{Z})^n$: the question is about sequences of elements which need not be distinct.) Alon and Dubiner [5, 6] asked whether there was some $c < 3$ such that $s(3, n) \leq c^n$.

Finally, a third motivation. When studying the design of statistical experiments, Bose in 1947 [17] was led to study so-called *caps*: subsets of the projective geometry $\mathbb{P}(\mathbb{F}_q^n)$ over a finite field $\mathbb{F}_q$ that contain no three collinear points. In particular, he was interested in the size of the largest caps. This question was taken up by Segre, who provided upper and lower bounds on the maximum size of a cap [117, 118], and has been well-studied since (e.g., a quick search on MathSciNet reveals at least a dozen papers in the last five years alone). Over $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$, caps have become known as "cap sets",[5] and a cap set is essentially the same as a set of points in $(\mathbb{Z}/3\mathbb{Z})^n$ containing no 3-term arithmetic progression; indeed, the extremal sizes of these two objects have identical asymptotic behavior as $n \to \infty$.[6]

With these motivations in mind, let us recall what's known about 3-term arithmetic progressions in $\mathbb{Z}/N\mathbb{Z}$ and in $(\mathbb{Z}/3\mathbb{Z})^n$. For any abelian group Z, let $r_3(Z)$ be the size of the largest subset of Z without 3-term arithmetic progressions. If

[5] Apparently, Tao [128] mistakenly introduced the term "cap set". He writes [130]: "... it seems I may have inadvertently propagated a slightly incorrect terminology in [128]: sets in $\mathbb{F}_3^n$ free of collinear triples are known as affine caps or simply caps in the design theory literature, rather than cap sets. The latter terminology seems to have become rather entrenched, though, at least in additive combinatorics circles...".

[6] If you aren't familiar with how to compare the asymptotic behavior of two functions $f, g \colon \mathbb{N} \to \mathbb{N}$, it's not hard, but now would be a good time to consult Appendix A. The only difference here is the difference between affine space $\mathbb{F}_3^n$ and projective space $\mathbb{P}(\mathbb{F}_3^n)$.

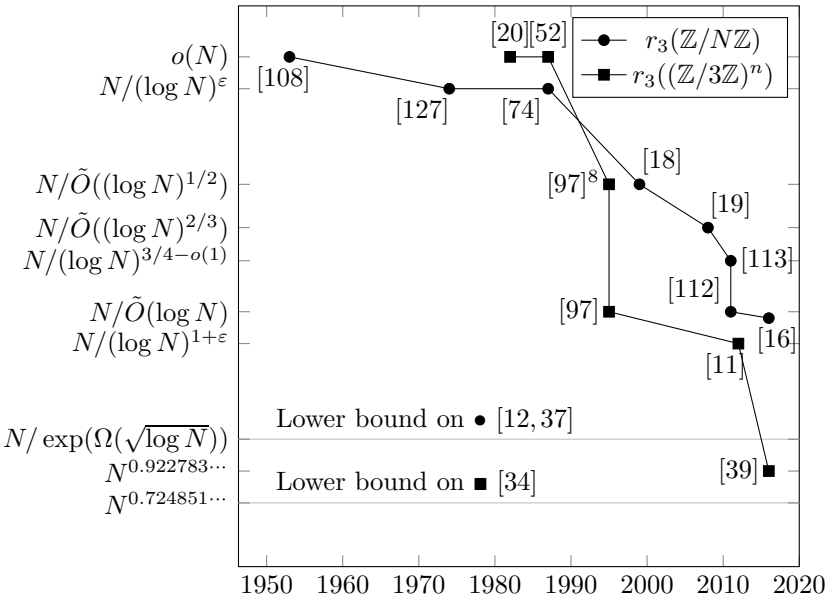

FIGURE 2. History of upper bounds on $r_3(\mathbb{Z}/N\mathbb{Z})$ and $r_3((\mathbb{Z}/3\mathbb{Z})^n)$. In all cases, ε denotes some constant strictly between 0 and 1/2, but each use of ε denotes a different constant. The $\tilde{O}$ hides terms of the form $(\log\log N)^c$ or smaller. Note that Behrend's [12] and Elkin's [37] lower bounds on $r_3(\mathbb{Z}/N\mathbb{Z})$ grow faster than $N^{1-\varepsilon}$ for any $\varepsilon > 0$.

we think of 3-term arithmetic progressions as some of the simplest nontrivial additive structures a set could have, $r_3(Z)$ is an upper bound on the size of (mostly) "unstructured" subsets of Z.

Figure 2 shows the history. Since we want to compare $\mathbb{Z}/N\mathbb{Z}$ against $(\mathbb{Z}/3\mathbb{Z})^n$, we will always use N to denote $|Z|$. Lower bounds on $r_3(Z)$ come from constructions of sets without 3-term arithmetic progressions. In $\mathbb{Z}/N\mathbb{Z}$, the current best bound is due to Elkin [37], who constructed such a set of size $\geq N(\log N)^{1/4}/e^{c\sqrt{\log N}}$ for some $c > 0$ (improving on Behrend's classic bound [12] by $\sqrt{\log N}$; see [70] for a shorter, albeit less constructive, proof). In $(\mathbb{Z}/3\mathbb{Z})^n$, the current best bound is due to Edel [34], who constructed a cap set of size $\geq N^{0.724851\cdots}$. In case you're a little rusty on your asymptotics,[7] $N/e^{O(\sqrt{\log N})}$ grows faster than $N^{1-\varepsilon}$ for arbitrarily small $\varepsilon > 0$, but slower than $N/(\log N)^c$ for arbitrarily large c. In contrast, Edel's lower bound is of the form $N^{0.72\cdots}$.

It is thus natural to ask where the truth lies for $(\mathbb{Z}/3\mathbb{Z})^n$: Does $r_3((\mathbb{Z}/3\mathbb{Z})^n)$ grow more quickly than $(3^n)^{1-\varepsilon}$ for all $\varepsilon > 0$, or is there some $c < 3$ such that $r_3((\mathbb{Z}/3\mathbb{Z})^n) \leq c^n$? While this was raised as a question by several authors as early as 1993 [5, 6, 22, 34], and it seems plausible that other experts may have believed the upper bound to be c^n for some $c < 3$, the earliest reference we can find in which

[7] What's the sound of an analytic number theorist drowning? "Log log log log$\cdots$" The same is true for algorithms researchers.

[8] Due to I. Rusza, see [97].

TABLE 1. Lower bounds on $r_3((\mathbb{Z}/3\mathbb{Z})^n)$. The bounds for $d \leq 6$ are tight [106]. The current best bound, due to Edel [34], uses a recursive construction to build a cap set of size $\approx 7.21 \times 10^{65}$ in $(\mathbb{Z}/3\mathbb{Z})^{480}$ ($N \approx 10^{229}$).

d	$r_3((\mathbb{Z}/3\mathbb{Z})^d) \geq$	$r_3((\mathbb{Z}/3\mathbb{Z})^n) \geq$
1	2	2^n
3	9 [17]	$9^{n/3} \approx 2.08^n$
4	20 [103]	$20^{n/4} \approx 2.11^n$
5	45 [36]	$45^{n/5} \approx 2.14^n$
6	112 [22, 35]	$112^{n/6} \approx 2.19^n$
62	$2(2 \cdot 112^{10} + 2 \cdot 10 \cdot 112^9 \cdot 12)$ [34]	$\approx 2.20^n$
480	$32^{80} + 8^5 \binom{10}{5} 112^{75} \times 12^5$ [34]	$\approx 2.217^n$

someone explicitly expressed the belief that $r_3((\mathbb{Z}/3\mathbb{Z})^n) \leq c^n$ for some $c < 3$ is from 2004 [64]; other authors expressed the opposite belief [128].

Conjecture 1.4 (Cap Set Conjecture [64] (cf. [5, 6, 22, 34]))**.** *$r_3((\mathbb{Z}/3\mathbb{Z})^n) \leq c^n$ for some $c < 3$.*

Before we come to its resolution, let's pause to discuss the lower bounds known on the cap set problem. Within $\mathbb{Z}/3\mathbb{Z}$, we have that the set $\{0, 1\}$ is a cap set; it follows that $\{0, 1\}^n$ is a cap set of size 2^n in $(\mathbb{Z}/3\mathbb{Z})^n$, for $x + y + z = 0$ in $(\mathbb{Z}/3\mathbb{Z})^n$ if and only if $x_i + y_i + z_i = 0$ for each coordinate $1 \leq i \leq n$. Similarly, if we have a cap set C of size s in $(\mathbb{Z}/3\mathbb{Z})^d$, then we get a cap set of size $s^{n/d}$ in $(\mathbb{Z}/3\mathbb{Z})^n$ (when d divides n) by partitioning the n coordinates into n/d groups of d, and considering the cap set $C^{n/d}$. Table 1 shows the bounds achieved using this idea.

Note that this technique always produces lower bounds of the form c^n for some $c < 3$. The only hope to disprove the conjecture this way would be to find an infinite family of better and better such constructions, and given the level of complexity of Edel's construction, finding such an infinite family (of course, before we knew it was impossible) seemed like a tall order.

Given the lack of consensus on which way this conjecture should be resolved, it was then quite a surprise to see it resolved in 2016. All the previous upper bounds on $r_3((\mathbb{Z}/3\mathbb{Z})^n)$ had used Fourier analytic techniques (with one exception due to Lev [90], though the techniques there were still philosophically close to Fourier analysis), and several people had speculated on ways to extend these techniques to get better bounds (e.g., [61, 128]). Then in 2016, Croot, Lev, and Pach [28] left Fourier analysis behind and introduced a beautiful new use of the polynomial method to show that $r_3((\mathbb{Z}/4\mathbb{Z})^n) \leq (4^n)^{0.926\cdots}$. In retrospect, the case of $\mathbb{Z}/4\mathbb{Z}$ in some ways seems harder than the $\mathbb{Z}/p\mathbb{Z}$ case with p prime. Why were they looking at the $\mathbb{Z}/4\mathbb{Z}$ question? Lev was kind enough to provide an answer (personal communication, 2018): their initial motivation was to improve Sanders's result "Roth's Theorem in $(\mathbb{Z}/4\mathbb{Z})^n$" [111]. Very shortly thereafter, and nearly simultaneously with one another, Ellenberg and Gijswijt leveraged the Croot–Lev–Pach technique to give a positive resolution to the Cap Set Conjecture [39]. This exponentially small upper bound on $r_3((\mathbb{Z}/3\mathbb{Z})^n)$ was all the more surprising because it shows a striking asymptotic difference between arithmetic progressions in $\mathbb{Z}/N\mathbb{Z}$ and those

in $(\mathbb{Z}/3\mathbb{Z})^n$. Moreover, while Fourier methods work in both settings, the Croot–Lev–Pach use of the polynomial method yields only trivial bounds for $\mathbb{Z}/N\mathbb{Z}$.

Such a strong upper bound is also tantalizing in its connection with the integers. Gowers observed [62, p. 273] that if the upper bound on $r_3(\mathbb{Z}/N\mathbb{Z})$ could be improved from its current record of $N(\log\log N)^4/\log N$ [16] to $N\log\log N/\log N$, it would give an alternative proof of *Roth's theorem in the primes* (a theorem due to Green [65] with improved bounds by Helfgott and de Roton [75]). The new greatly improved bound on cap sets perhaps provides renewed hope that the upper bound on $r_3(\mathbb{Z}/N\mathbb{Z})$ could at least be improved by this seemingly tiny amount to give a purely combinatorial proof of Roth's theorem in the primes.

Even for cap sets, there is still an exponential gap between the upper bound of 2.756^n [39] and the lower bound of 2.217^n [34], and closing this gap is an interesting problem. Improving the lower bound "just" requires a new construction; the smallest dimension in which a finite construction could get within .01 of 2.756 is at least $d = 597$.[9] We'll see in Section 3.2, through its connection with algorithms for matrix multiplication (of all things!), that the Croot–Lev–Pach–Ellenberg–Gijswijt technique extends from sets without arithmetic progressions to so-called *tricolored sum-free sets*, and that in the tricolored setting the upper bound of $\approx 2.756^n$ is indeed tight [79]. Thus any attempt to improve the upper bound must differ substantially enough to *not* extend to the tricolored setting. Alternatively, to improve the lower bound, one might try to turn the Kleinberg–Sawin–Speyer construction of a tricolored sum-free set [79] into an ordinary cap set of the same size, but this too seems difficult. See the end of Section 3 for a more detailed discussion of the difficulties in improving the upper bound.

2. Proof of the Cap Set Conjecture

The method introduced by Croot, Lev, and Pach in [28], which ultimately led to the resolution of the Cap Set Conjecture in [39], is an application of the polynomial method. In general, the polynomial method is the application of algebraic geometry to combinatorics (and sometimes other fields that, at first blush, seem unrelated to algebraic geometry). In particular, natural combinatorial structures can often be defined in terms of polynomial equations, and then by reasoning about these systems of polynomial equations (the domain of algebraic geometry) we can often learn about the original combinatorial structures. In addition to the Cap Set Conjecture, the polynomial method was also instrumental in the recent solution of another long-standing combinatorial problem, the finite field Kakeya problem [31] (see also [33] for nearly tight bounds, again using the polynomial method). For general introductions to the polynomial method see [76, Chapter 16] (in the context of extremal combinatorics), [132, Chapter 9] (in the context of additive combinatorics), and [129] for a recent tutorial and survey.

The version of the Ellenberg–Gijswijt proof we will follow here is the *symmetric* version due to Tao [130]. The idea is essentially the same as Ellenberg–Gijswijt,

[9] We calculated this using the exact upper bound from [28], reproduced below as Lemma 2.5. That is, for each $d = 2, 3, \ldots, 597$, we raised the formula from Lemma 2.5 to the $1/d$ power and reported the first value of d where this value was at least $2.756 - .01 = 2.746$. While we could ask the same for even more generous bounds, such as getting within .1 of 2.756, the result there is $d = 8$ which is so small that we believe that the upper bound given by Lemma 2.5 is far from being tight. For example, at $d = 6$ this formula gives an upper bound on $r_3((\mathbb{Z}/3\mathbb{Z})^6)$ of 153, but the true value is 112 [106].

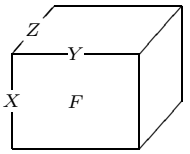

FIGURE 3. Visualizing a three-variable function $F\colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$ as a three-dimensional array or "3-tensor."

only the syntax is different. In the original proof [39], when studying solutions to the equation $x + y + z = 0$—which is clearly symmetric in all three of x, y, and z—they single out two of the variables and consider the rank of a matrix of the form $M_{x,y} = f(x + y)$ for some polynomial(s) f. What makes Tao's version more symmetric is that each of the three variables is put on an equal footing.

2.1. **Tensors.** Since we want to consider solutions to the equation $x+y+z = 0$ with $x, y, z \in (\mathbb{Z}/3\mathbb{Z})^n$, we will consider three-variable functions $F\colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$, where $X, Y, Z \subseteq (\mathbb{Z}/3\mathbb{Z})^n$. (Almost everything we say will apply to $\mathbb{Z}/p\mathbb{Z}$ with p a prime, but we'll stick with $p = 3$ for consistency.) Just as we could visualize a two-variable function on a finite domain $F\colon X \times Y \to \mathbb{Z}/3\mathbb{Z}$ as a $|X| \times |Y|$ matrix with entries from $\mathbb{Z}/3\mathbb{Z}$, when X, Y, Z are finite sets we may visualize a three-variable function as a three-dimensional array of numbers (see Figure 3)—sometimes called a 3-tensor—where the rows are indexed by the elements of X, the columns by the elements of Y, and the "depths" (the row-like thing, but in the third dimension) by the elements of Z. We thus refer to $|X|$, $|Y|$, and $|Z|$ as the *side lengths* of the 3-tensor F.

The first observation follows.

Observation 2.1. If $A \subseteq (\mathbb{Z}/3\mathbb{Z})^n$ is a cap set, then the function

$$F(x, y, z) = \delta_0(x + y + z) = \begin{cases} 1 & x + y + z = 0, \\ 0 & \text{otherwise,} \end{cases}$$

when restricted to $A \times A \times A$, satisfies

$$F|_{A\times A\times A}(x, y, z) = \begin{cases} 1 & \text{if } x = y = z, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $F|_{A\times A\times A}$ looks like a three-dimensional version of the identity matrix (see Figure 4). The proof of this observation follows directly from the definition of cap set.

If a tensor F has this property—that is, if $F(x, y, z) \neq 0$ *if and only if* $x = y = z$—we call F a *diagonal* tensor. (For the purposes of this article, unlike a "diagonal matrix," which can have zero entries on its diagonal, when we speak of a "diagonal tensor", we mean that all of its diagonal entries are nonzero.)

Continuing the analogy with matrices, what we would like is some notion of rank for 3-tensors such that

(1) diagonal tensors have rank equal to their side length (just as diagonal matrices do); but

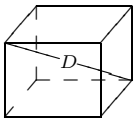

FIGURE 4. Cap sets correspond to diagonal tensors; in this picture, the only nonzero entries are along the body diagonal D.

(2) the function $\delta_0(x+y+z)$ (on all of $(\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n$) has rank exponentially smaller than its side length 3^n; and
(3) the rank of a tensor is always at least the rank of any of its sub-tensors, gotten by restricting $F\colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$ to $X' \times Y' \times Z'$ for $X' \subseteq X, Y' \subseteq Y, Z' \subseteq Z$.

Given such a notion of rank, we quickly prove the Cap Set Conjecture:

Proof of the Cap Set Conjecture, assuming a notion of rank satisfying (1)–(3).
Let $F(x,y,z) = \delta_0(x+y+z)$ be the tensor above. Then for any cap set A,

$$\begin{array}{rcll} |A| & = & \operatorname{rank}(F|_{A\times A\times A}) & \text{by property (1), since } F|_{A\times A\times A} \text{ is diagonal} \\ & \leq & \operatorname{rank}(F) & \text{by property (3)} \\ & \leq & c^n & \text{for some } c<3, \text{ by property (2).} \end{array}$$

And that's it! □

The following notion is a direct generalization of the rank of matrices, which we'll see has the desired properties. Showing that it satisfies properties (1)–(3) above will then turn the above proof of the Cap Set Conjecture into a Proof.

Definition 2.2 (Tao [130])**.** Given a 3-tensor $F\colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$, its *slice-rank* is the least r such that F can be written as the following sum of r terms:

$$F(x,y,z) = \sum_{i=1}^{a} f_i(x)g_i(y,z) + \sum_{i=a+1}^{b} f_i(y)g_i(x,z) + \sum_{i=b+1}^{r} f_i(z)g_i(x,y).$$

Any such expression for F, even with r not minimal, is called a *slice decomposition*.

This notion was introduced by Tao [130] in developing this symmetric version of the Ellenberg–Gijswijt proof, and following [14] we call it *slice-rank*. Further properties of slice-rank were elaborated in [14, 15, 131].

A 3-tensor of slice-rank 1 thus has the form $f(x)g(y,z)$ (or any of its symmetric versions, gotten by permuting the variables). In terms of our three-dimensional array visualization, we may think of $g(y,z)$ as a matrix placed on a horizontal slab coming out of the page, and then the function $f(x)g(y,z)$ consists of stacking up a bunch of scalar multiples of this slab on top of one another. In other words, a 3-tensor has slice-rank 1 iff its two-dimensional layers (in at least one of the three directions) are all scalar multiples of one another.

Property (3), that slice-rank cannot increase when passing to subtensors, is the easiest to see:

Observation 2.3. For any tensor $F\colon X \times Y \times Z \to \mathbb{Z}/3\mathbb{Z}$, and any $X' \subseteq X, Y' \subseteq Y, Z' \subseteq Z$:

$$\text{slice-rank}(F|_{X'\times Y'\times Z'}) \leq \text{slice-rank}(F).$$

Proof. Given a slice decomposition for F with functions $f_i(\bullet), g_i(\bullet,\bullet)$, we get a slice decomposition for $F|_{X'\times Y'\times Z'}$ by restricting each f_i and g_i to the primed subsets. That is, for $f_i\colon X\to\mathbb{Z}/3\mathbb{Z}$, we restrict it to $f_i' = f_i|_{X'}\colon X'\to\mathbb{Z}/3\mathbb{Z}$, and restrict $g_i\colon Y\times Z\to\mathbb{Z}/3\mathbb{Z}$ to $g_i|_{Y'\times Z'}\colon Y'\times Z'\to\mathbb{Z}/3\mathbb{Z}$, and so on. □

Property (1), that the slice-rank of diagonal tensors is maximal, can be shown by induction from the 2-variable (i.e., 2-tensor, i.e., matrix) case, which we leave as an exercise.

Lemma 2.4 (Tao [130])**.** *The slice-rank of a diagonal tensor is equal to its number of nonzero entries.*

Finally, Property (2), the exponential upper bound, brings us to the key idea of the proof, which is an application of the polynomial method.

2.2. **Key idea of the proof.** The following lemma is the **key idea** from Croot, Lev, and Pach in [28] and Ellenberg and Gijswijt in [39], which unlocks the whole proof. Since a polynomial $F(\vec{x},\vec{y},\vec{z})$ on $((\mathbb{Z}/3\mathbb{Z})^n)^3$, is, in particular, a function $F\colon(\mathbb{Z}/3\mathbb{Z})^n\times(\mathbb{Z}/3\mathbb{Z})^n\times(\mathbb{Z}/3\mathbb{Z})^n\to\mathbb{Z}/3\mathbb{Z}$, we may view it as a 3-tensor of side length 3^n. In the language of slice-rank we have the following.

Lemma 2.5 (Croot, Lev, and Pach in [28], slightly generalized by Ellenberg and Gijswijt in [39])**.** *If $F(x_1,\dots,x_n,y_1,\dots,y_n,z_1,\dots,z_n)$ is a polynomial over $\mathbb{Z}/3\mathbb{Z}$ of degree $\le d$, then, when viewed as a 3-tensor of side length 3^n as above, we have*

$$\text{slice-rank}(F)\le 3\sum_{\substack{a,b,c\ge 0\\ a+b+c=n\\ b+2c\le d/3}}\frac{n!}{a!b!c!}.$$

Proof. Each monomial m in F has the form $m = x_1^{e_1}x_2^{e_2}\cdots x_n^{e_n}y_1^{e_1'}\cdots y_n^{e_n'}z_1^{e_1''}\cdots z_n^{e_n''}$, where the exponents satisfy $\sum_i(e_i+e_i'+e_i'')\le d$. If we consider the degrees of this monomial in the x's, the y's, and the z's separately, namely $d_x(m) := \sum_i e_i$, $d_y(m)=\sum_i e_i'$, $d_z(m)=\sum_i e_i''$, then we have $d_x(m)+d_y(m)+d_z(m)=\deg(m)\le d$. Therefore, for each m at least one of $d_x(m)$, $d_y(m)$, and $d_z(m)$ must be $\le d/3$.

Now, let M_x be the set of terms (=monomials together with their coefficients) for which $d_x(m)\le d/3$, let M_y be the set of terms for which $d_y(m)\le d/3$, and define M_z similarly. Although not strictly necessary, it will make things simpler if M_x, M_y, M_z are disjoint, so let's remove from M_y anything in M_x, and then remove from M_z anything in M_x or in M_y. Then we can write F as

$$F(x,y,z)=\sum_{m\in M_x}m+\sum_{m\in M_y}m+\sum_{m\in M_z}m.$$

The key trick here is to rewrite each of these three sums by factoring out the relevant variables. For example, factor out the x variables as much as possible from $\sum_{m\in M_x}m$,

$$\sum_{m\in M_x}m=\sum_{\substack{e_1,\dots,e_n\ge 0\\ \sum_i e_i\le d/3}}x_1^{e_1}x_2^{e_2}\cdots x_n^{e_n}\overline{g}_{e_1,\dots,e_n}(\vec{y},\vec{z}),$$

where the $\overline{g}_{e_1,\dots,e_n}$ are precisely what they need to be to make this equality hold; but the only fact we need about the $\overline{g}$'s is that they only depend on $\vec{y},\vec{z}$, and not on $\vec{x}$, as then the right-hand side here is a slice decomposition of the left-hand side. (The $\overline{g}$'s here are overlined because we're about to replace them.)

The next thing to note is that we can also restrict the e_i so that they are all at most 2. For any $\alpha \in \mathbb{Z}/3\mathbb{Z}$, note that $\alpha^3 = \alpha$, and thus the polynomial x^3, as a function on $\mathbb{Z}/3\mathbb{Z}$, computes the same function as the polynomial x. This lets us reduce the degree of each variable in each monomial until it is strictly less than 3. We are then left with

$$\sum_{m \in M_x} m = \sum_{\substack{e_1,\dots,e_n \in \{0,1,2\} \\ \sum_i e_i \leq d/3}} x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n} g_{e_1,\dots,e_n}(\vec{y}, \vec{z}).$$

The g's here may be combinations of some of the $\overline{g}$'s from before, but again, all we care about is that they do not depend on $\vec{x}$. Thus each term in this sum has slice-rank 1, and we have

$$\text{slice-rank}\left(\sum_{m \in M_x} m\right) \leq \left|\{(e_1,\dots,e_n) \in \{0,1,2\}^n : \sum_i e_i \leq d/3\}\right|.$$

By swapping the role of $\vec{x}$, $\vec{y}$, and $\vec{z}$, we get the same bound on $\sum_{m \in M_y} m$ and $\sum_{m \in M_z} m$, and thus have

$$\text{slice-rank}(F) \leq 3\left|\{(e_1,\dots,e_n) \in \{0,1,2\}^n : \sum_i e_i \leq d/3\}\right|. \tag{2.1}$$

All that remains is to show that the set on the right-hand side has the size claimed in the statement of the lemma. Given $(e_1,\dots,e_n) \in \{0,1,2\}^n$, let a be the number of e_i that are 0, let b be the number of e_i that are 1, and let c be the number of e_i that are 2. Then

$$a + b + c = n, \tag{2.2}$$

for each e_i takes exactly one of these three values. Also, we have that $\sum_i e_i = a \cdot 0 + b \cdot 1 + c \cdot 2$, so $\sum_i e_i \leq d/3$ if and only if

$$b + 2c \leq d/3. \tag{2.3}$$

Thus we can rewrite (2.1) as

$$\text{slice-rank}(F) \leq 3 \sum_{\substack{a,b,c \geq 0 \\ a+b+c=n \\ b+2c \leq d/3}} |\{(e_1,\dots,e_n) \text{ with } a \text{ 0's, } b \text{ 1's, and } c \text{ 2's}\}|. \tag{2.4}$$

Finally, given values of a, b, c satisfying the constraints (2.2) and (2.3), we need to know how many vectors $(e_1,\dots,e_n)$ have a 0's, b 1's, and c 2's. Every such vector comes from permuting the coordinates of the vector $(0,0,\dots,0,1,1,\dots,1,2,2,\dots,2)$ (with a 0's, b 1's, and c 2's). There are $n!$ such permutations. However, this is significantly overcounting, since if we permute only those coordinates with the same value, we get back the same vector. Thus we have overcounted by a factor of $a!b!c!$, so our final count is

$$|\{(e_1,\dots,e_n) \in \{0,1,2\}^n : \text{ there are } a \text{ 0's, } b \text{ 1's, and } c \text{ 2's}\}| = \frac{n!}{a!b!c!}. \tag{2.5}$$

Combining (2.4) with (2.5) yields the lemma. □

2.3. **Finishing it off.** We now use the key Lemma 2.5 to prove property (2) for slice-rank, which will thus complete the proof of the Cap Set Conjecture. The techniques used here are completely standard, going back probably at least two centuries; we include them so all our readers can see how everything fits together.

Observation 2.6. Let $F_0 \colon (\mathbb{Z}/3\mathbb{Z})^n \to \mathbb{Z}/3\mathbb{Z}$ be any function. Then the 3-tensor $F \colon (\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n \times (\mathbb{Z}/3\mathbb{Z})^n \to \mathbb{Z}/3\mathbb{Z}$ defined by $F(x,y,z) = F_0(x+y+z)$ can be written as a polynomial of degree at most $2n$.

Proof. The idea is to use interpolation to write the function F_0 as a polynomial. More formally, for $\alpha \in \mathbb{Z}/3\mathbb{Z}$, we can write the indicator function $\delta_\alpha(x)$, which is 1 if and only if $x = \alpha$, and it is 0 otherwise, as

$$\delta_\alpha(x) = 1 - (x-\alpha)^2.$$

Then for any $\vec{\alpha} \in (\mathbb{Z}/3\mathbb{Z})^n$, we can write the indicator function $\delta_{\vec{\alpha}}(\vec{x})$ as

$$\delta_{\vec{\alpha}}(\vec{x}) = \prod_{i=1}^n \delta_{\alpha_i}(x_i) = \prod_{i=1}^n \left(1 - (x_i - \alpha_i)^2\right).$$

Note that $\deg \delta_{\vec{\alpha}} = 2n$. Thus, we can write any function F_0 as a polynomial of degree at most $2n$,

$$F_0(\vec{x}) = \sum_{\vec{\alpha} \in (\mathbb{Z}/3\mathbb{Z})^n} \delta_{\vec{\alpha}}(\vec{x}) F_0(\vec{\alpha}). \qquad \square$$

In particular, since the 3-tensor we care about, $F(x,y,z) = \delta_0(x+y+z)$, has the form in the preceding observation, we may apply the key Lemma 2.5 to a function of degree $\leq 2n$.

And now we come to the crucial (if standard) estimate. This estimate is best phrased in terms of the Shannon entropy of a probability distribution with three outcomes. Recall that for $\alpha_1, \ldots, \alpha_k \geq 0$ summing to 1, the entropy of the corresponding k-outcome distribution is $h(\alpha_1, \ldots, \alpha_k) = \sum_{i=1}^k \alpha_i \log(1/\alpha_i)$. To see how entropy might arise in estimating the sum from Lemma 2.5, recall Stirling's Formula: $a! \sim (a/e)^a \sqrt{2\pi a}$. In particular, $\log(a!) \sim a \log a$, and this is indeed how entropy arises here. In the proof we'll need two standard properties of the entropy: (1) it is strictly convex; and (2) it attains its maximum value uniquely when $\alpha_1 = \cdots = \alpha_k = 1/k$, namely, $h(1/k, \ldots, 1/k) = \log k$.

Lemma 2.7. *Let*

$$\theta = \max_{\substack{\alpha,\beta,\gamma \geq 0 \\ \alpha+\beta+\gamma=1 \\ \beta+2\gamma \leq 2/3}} e^{h(\alpha,\beta,\gamma)}.$$

Then θ is strictly less than 3*, and the sum from Lemma* 2.5 *is asymptotically equal to*

$$\sum_{\substack{a,b,c \geq 0 \\ a+b+c=n \\ b+2c \leq 2n/3}} \frac{n!}{a!b!c!} \sim \theta^{n(1+o(1))}.$$

Proof. Let θ be as in the statement of the lemma. By the standard properties of entropy mentioned above, the maximum value of $h(\alpha,\beta,\gamma)$ is $\log 3$, uniquely attained at the uniform distribution $\alpha = \beta = \gamma = 1/3$. However, the constraint $\beta + 2\gamma \leq 2/3$ is violated by the uniform distribution, so no distribution satisfying

this constraint can have entropy as large as $\log 3$. Thus every term in the expression for θ is strictly less than $e^{\log 3} = 3$, hence θ is also.

Now, to see the second part, if $a+b+c=n$, Stirling's Formula gives us

$$\begin{aligned}\frac{n!}{a!b!c!} &\sim \left(\frac{n}{e}\right)^n \left(\frac{e}{a}\right)^a \left(\frac{e}{b}\right)^b \left(\frac{e}{c}\right)^c \sqrt{\frac{2\pi n}{8\pi^3 abc}} \\ &= \frac{n^n}{a^a b^b c^c}\sqrt{\frac{n}{4\pi^2 abc}} \\ &= \frac{1}{2\pi}\left(\frac{n^n}{a^a b^b c^c}\right)^{1+o(1)}.\end{aligned}$$

(If you're not familiar with the $o(1)$ notation in the exponent, now might be a good time to quickly consult Appendix A.)

Now, since $a+b+c=n$, and we want to consider the behavior as $n\to\infty$, it's useful to rescale these three to get a probability distribution that essentially doesn't depend on n: For $\alpha = a/n, \beta = b/n, \gamma = c/n$, we have $\alpha,\beta,\gamma \geq 0$ and $\alpha+\beta+\gamma = 1$. This allows us to rephrase the preceding estimate in terms of entropy:

$$\begin{aligned}\frac{n!}{a!b!c!} &\sim \frac{1}{2\pi}\left(\frac{1}{\alpha^a\beta^b\gamma^c}\right)^{1+o(1)} \\ &= \frac{1}{2\pi}\left(\exp\left(-(a\log\alpha + b\log\beta + c\log\gamma)\right)^{1+o(1)}\right. \\ &= \frac{1}{2\pi}\left(\exp\left(-n(\alpha\log\alpha + \beta\log\beta + \gamma\log\gamma)\,(1+o(1))\right)\right. \\ &= \frac{1}{2\pi}\exp\left(nh(\alpha,\beta,\gamma)(1+o(1))\right).\end{aligned}$$

Our sum can thus be rewritten asymptotically as

$$\sim \frac{1}{2\pi}\sum_{\substack{a,b,c\geq 0\\ a+b+c=n\\ b+2c\leq 2n/3}} \exp\left(nh(a/n,b/n,c/n)(1+o(1))\right).$$

Since this is a sum of exponentials, and the sum only has $O(n^2)$ terms, as n gets large this will be dominated by the single largest exponential. For if the largest is e^{Cn} and the next largest is $e^{C'n}$, with $C' < C$, then even if all the remaining terms had magnitude $e^{C'n}$, they would still only add up to $O(n^2 e^{C'n}) \leq e^{C'n(1+o(1))}$, which is still $o(e^{Cn})$.

To find the single largest term, we need to maximize the entropy $h(\alpha,\beta,\gamma)$ subject to the constraints that α,β,γ form a probability distribution satisfying $\beta + 2\gamma \leq 2/3$. This is precisely the maximization in the definition of θ. □

While the preceding lemma is already enough to prove the Cap Set Conjecture, a routine Lagrange multiplier calculation will find the exact values of α,β,γ that give the value of θ (see [130]). Since entropy is convex, simple numerical hill-climbing will also yield the correct maximum value to any desired accuracy. The largest value of $h(\alpha,\beta,\gamma)$ subject to the constraint $\beta+2\gamma\leq 2/3$ is ≈ 1.013455, resulting in an upper bound of $\sim \exp(1.013455\cdots n(1+o(1))) = 2.756\cdots^{n(1+o(1))}$.

2.4. **Bread crumbs of the proof.** This completes the proof of the Cap Set Conjecture. Here are the bread crumbs of the proof (this is how I remember it).

(1) Observe that a cap set corresponds to a diagonal 3-tensor, which is a subtensor of $F(x, y, z)$, the indicator function of $x + y + z = 0$. Since diagonal 3-tensors are like diagonal matrices, the size of the cap set is the slice-rank of the diagonal 3-subtensor, which is therefore upper bounded by the slice-rank of F itself.
(2) Write the indicator function F as a polynomial. Note that it has degree $\leq 2n$ (it's a product of n indicator functions, each of which has degree 2).
(3) Pigeonhole the monomials by degree, $d = 2n$ pigeons into three holes (one for each of x, y, z).
(4) Group together the monomials which have x-degree $\leq d/3$, (and similarly for y-degree and z-degree)
(5) Count monomials and estimate the growth rate using entropy maximization.

3. Tricolored sum-free sets and the question of tight bounds

For both $r_3(\mathbb{Z}/N\mathbb{Z})$ and $r_3((\mathbb{Z}/3\mathbb{Z})^n)$ there is still a gap between the best upper and lower bounds known: in the former case, between $N/(\log N)^{1+\varepsilon}$ [16] and $N/e^{c\sqrt{\log N}-(1/4)\log\log N}$ [12, 37], and in the latter between 2.756^n [39] and 2.217^n [34]. In both cases, however, if we broaden our scope to slightly more general objects, we find that the upper bounds are essentially tight.

3.1. **Tight bounds in the integers?** An arithmetic progression is a sequence of integers satisfying $x+z = 2y$. In the case of the integers (or $\mathbb{Z}/N\mathbb{Z}$), if we generalize to other translation-invariant linear equations—that is, of the form $\sum_{i=1}^k a_i x_i = 0$ where $\sum_i a_i = 0$—we find essentially tight bounds. Translation invariance is a natural condition here, as we want to consider subsets of $\mathbb{Z}$ satisfying some condition on their relative differences $x_i - x_j$, which are unaffected by translating the entire subset by an additive constant.

Recently, Schoen and Sisask [116] (following [115], who showed the $k = 6$ case) showed that any subset $A \subseteq \{1, \ldots, N\}$ of size $\geq N/\exp(c(\log N)^{1/7})$ contains distinct elements $x_1, \ldots, x_4$ such that

$$x_1 + x_2 + x_3 = 3x_4.$$

Behrend's construction for arithmetic progressions adapts easily to this setting, resulting in essentially tight bounds for the preceding linear equation: The difference is only between 1/2 and 1/7 in the exponent of the exponent (and that's not a typo!).

In fact, their argument works for any translation-invariant equation with at least four terms. As $x + y = 2z$ is of this form with only three terms, improving their result from four terms to three terms would show that Behrend's construction is essentially tight for $r_3(\mathbb{Z}/N\mathbb{Z})$. It's worth noting, given our discussion above, that their techniques also give similar bounds for such equations over finite fields (but over finite fields no construction as large as Behrend's is known for four-term equations, even though such a construction is not ruled out by the Cap Set Conjecture). In fact, they first present the argument over finite fields, as it is simpler, and then use Bohr sets to extend the argument to the integers, as discussed above.

This result has the following interesting implication. Either:

- their result extends to the three-term case, in which case Behrend's bound is essentially tight for $r_3(\mathbb{Z}/N\mathbb{Z})$; or
- showing a significantly better lower bound must use techniques that are sensitive to the difference between 3-term linear equations and 4-term linear equations.

(A lower bound of the form $N/(\log N)^c$ for some $c > 1$ would show that Bloom's upper bound is nearly tight.) We note that, although the difference here is between 3-term and 4-term linear equations, this situation actually seems quite different than the syntactically similar difference between 3-term and 4-term arithmetic progressions. In particular, since 3-term arithmetic progressions can be captured by a single equation, they are relatively easy to analyze using Fourier analysis. Since 4-term arithmetic progressions require two equations ($x_1 + x_3 = 2x_2$ and $x_2 + x_4 = 2x_3$), there is no single Fourier expression that captures them, thus necessitating the "higher-order Fourier analysis" suggested by Gowers [59, 60] and developed by Green, Tao, and Ziegler [69]. In contrast, both the $k = 3$ and $k = 4$ cases considered in this section are still just a single equation and thus are—at least in principle—amenable to standard Fourier-analytic techniques. The difference, if any, between 3-term and 4-term linear equations is apparently more subtle.

3.2. **Tight bounds in vector spaces over finite fields?** In the case of $(\mathbb{Z}/3\mathbb{Z})^n$, a slightly different generalization yields 2.756^n as a tight bound, matching the Ellenberg–Gijswijt upper bound. This generalization was motivated by algorithms for matrix multiplication (see Section 4.2), but our starting point here will be a simple observation about slice rank, which generalizes the fact that the rank of a matrix is invariant under change of basis. Given a tensor $F\colon X \times Y \times Z \to \mathbb{F}$ and an invertible $|X| \times |X|$ matrix S, $|Y| \times |Y|$ matrix T, and $|Z| \times |Z|$ matrix U, we may use these as change of basis matrices on $\mathbb{F}^X$, $\mathbb{F}^Y$, and $\mathbb{F}^Z$ to get a new tensor $F'(x', y', z') = \sum_{x,y,z} S(x, x')T(y, y')U(z, z')F(x, y, z)$.

Observation 3.1. If $F, F'\colon X \times Y \times Z \to \mathbb{F}$ are two 3-tensors that differ by a change of basis (as above), then slice-rank(F) = slice-rank(F').

Proof. Suppose slice-rank$(F) = r$. Then there is a slice decomposition

$$F(x, y, z) = \sum_{i=1}^{a} f_i(x)g_i(y, z) + \sum_{i=a+1}^{b} f_i(y)g_i(x, z) + \sum_{i=b+1}^{r} f_i(z)g_i(x, y).$$

Given a change of basis (S, T, U), let us apply it to the preceding decomposition. Now, for simplicity, let's just consider the first summation $\sum_i f_i(x)g_i(y, z)$ of the slice decomposition of F, and how it appears in the expression for F'. The other two summations will be handled similarly (one advantage of the symmetry of the notion of slice-rank). We have

$$\begin{aligned}\sum_{x,y,z} S(x, x')T(y, y')U(z, z') \sum_{i=1}^{a} f_i(x)g_i(y, z)\\ = \sum_{i=1}^{a} \sum_{x} S(x, x')f_i(x) \sum_{y,z} T(y, y')U(z, z')g_i(y, z)\end{aligned}$$

Note that, by defining $f_i'(x') = \sum_x S(x,x')f_i(x)$, f_i' only depends on x', since we sum over all values of x. Similarly, the function

$$g_i'(y',z') = \sum_{y,z} T(y,y')U(z,z')g_i(y,z)$$

depends only on y', z'. Using these new functions of x', y', z', the preceding expression becomes

$$\sum_{i=1}^{a} f_i'(x')g_i'(y',z'),$$

thus our first sum can be written after the change of basis using exactly as many slice-rank 1 terms as before, and similarly for the second and third sums. Thus slice-rank(F) = slice-rank(F'). □

Although this may seem a rather trivial consequence of the definition, note that when we considered a cap set $A \subseteq (\mathbb{Z}/3\mathbb{Z})^n$, it at least *felt* important that we were using values from the *same* set A for all three variables x, y, z. But the above change-of-basis observation says that we can change basis in X independent from Y independent from Z. In particular, if we change bases using permutation matrices, this corresponds to simply reordering the elements of $(\mathbb{Z}/3\mathbb{Z})^n$ in each of X, Y, Z. What does a diagonal subtensor of our favorite tensor, $F(x,y,z) = \delta_0(x+y+z)$, look like after permuting basis elements? It's a restriction of F to $A \times B \times C$ with $A, B, C \subseteq (\mathbb{Z}/3\mathbb{Z})^n$ such that $F(a_i, b_j, c_k) = 1$ if and only if $i = j = k$, where $A = \{a_1, \ldots, a_{|A|}\}$, and similarly for B, C. This leads to the following notion.

Definition 3.2 (Tricolored sum-free set [8, 14])**.** A *tricolored sum-free set* of cardinality ℓ in an abelian group Z consists of three subsets $(a_1, \ldots, a_\ell)$, $(b_1, \ldots, b_\ell)$, $(c_1, \ldots, c_\ell)$ with all $a_i, b_j, c_k \in Z$, such that

$$(\forall i,j,k)[a_i + b_j + c_k = 0 \iff i = j = k].$$

(The indexing is only relevant for identifying the matching between A, B, and C.)

Cap sets are examples of tricolored sum-free sets, but they are far from the only ones.

From our observations above, we thus have the following.

Lemma 3.3. *For any finite abelian group Z, let $r_3^{\mathrm{col}}(Z)$ denote the size of the largest tricolored sum-free set in Z, and let F_Z be the $|Z| \times |Z| \times |Z|$ tensor defined by $F_Z(x,y,z) = \delta_0(x+y+z)$. Then we have*

$$r_3^{\mathrm{col}}(Z) \leq \text{slice-rank}(F_Z).$$

Proof. From the discussion above, tricolored sum-free sets yield diagonal subtensors in some basis. Apply Observation 3.1 and Definition 2.4. □

Finally, for tricolored sum-free sets, Kleinberg, Sawin, and Speyer proved that the Ellenberg–Gijswijt bound is essentially exactly tight.

Theorem 3.4 (Kleinberg, Sawin, and Speyer [79], with a lemma from Norin [101] and Pebody [102])**.** *Let $\theta \approx 2.756$ be the base of the exponent in the Ellenberg–Gijswijt bound (as in Lemma 2.7). There is a tricolored sum-free set in $(\mathbb{Z}/3\mathbb{Z})^n$ of size $\geq \theta^{n(1-o(1))}$.*

To me, one of the really cool things here is not just that they achieved a tight bound, but also their method of proof. They use a pullback of Behrend's construction in the integers! Namely, in outline, they essentially

(1) choose three random mappings $h_1, h_2, h_3 \colon \mathbb{Z}^n \to \mathbb{Z}/p\mathbb{Z}$ for a large prime $p \sim \exp(cn)$ for some c;
(2) use Behrend's construction to get a large set $S \subseteq \mathbb{Z}/p\mathbb{Z}$ without arithmetic progressions; and
(3) build their tricolored sum-free set as a large subset of $\{(a, b, c) \in (\mathbb{Z}/3\mathbb{Z})^n : h_1(a) = h_2(b) = h_3(c) \in S\}$.

The $o(1)$ in the exponent of Theorem 3.4 hides a factor which is nearly exactly the density of the Behrend/Elkin construction relative to the prime p. One striking aspect of this construction is that it uses a construction which is *not* known to be tight in the integers, to prove a *tight* lower bound in $(\mathbb{Z}/3\mathbb{Z})^n$.

Finally, let us return to the question of tight bounds on $r_3((\mathbb{Z}/3\mathbb{Z})^n)$. Of course, we may take Theorem 3.4 as some indication that the Ellenberg–Gijswijt bound is already tight for cap sets. At the end of Section 1.4 we began discussing what is needed for improving the lower bound. But already this little bit about the proof of Theorem 3.4 may give us some inspiration: Perhaps by taking pullbacks of Behrend's construction, we can indeed get an infinite family of better and better cap sets in $(\mathbb{Z}/3\mathbb{Z})^d$ for $d \to \infty$ that would meet the Ellenberg–Gijswijt bound.

To improve the upper bound, the key barrier to be avoided (at the moment) is the use of slice-rank itself. This is not particular to Tao's symmetric formulation; rather, any technique (such as the Croot–Lev–Pach or Ellenberg–Gijswijt techniques), which yields a slice-rank upper bound falls prey to this limitation. For slice-rank is an upper bound on the size of tricolored sum-free sets, and Theorem 3.4 says that these bounds cannot be further improved. Thus, to improve the upper bound what is needed is a method that is somehow sensitive to the difference between a tricolored sum-free set and a cap set, or equivalently, between a diagonal tensor in arbitrary bases versus a diagonal tensor in three identical bases. Said another way, one needs a property of 3-tensors that is invariant under change of bases of the form (S, S, S), but *not* invariant under change of bases of the form (S, T, U).

4. Applications and extensions

Next we come to the question of the relationship between the Cap Set Conjecture and other problems or even other areas of mathematics. In this section we'll cover several applications of the Croot–Lev–Pach polynomial method, as well as extensions of the Cap Set Conjecture motivated by other questions. Since Croot, Lev, and Pach first posted their preprint, progress on these applications has happened very rapidly, and there is more than we can possibly cover in this short space. I will cover those with which I am most familiar. Here are some that are left out (I cannot hope to be exhaustive): relations between polynomials, namely, a polynomial Sárközy's Theorem [67]; sum-sets as unions of sum-sets of subsets [38]; k-colored sum-free sets [94]; Harborth's original question about zero-sum-avoiding sequences and Erdős–Ginzburg–Ziv constants [51]; subsets containing no right angles [57]; and ordered tricolored sum-free sets [78]. As our purpose here is just to highlight a few of the many connections the Cap Set Conjecture has with other areas of mathematics, we won't be quite as expository in this section as we've been

so far, but will point the reader to the relevant literature for further details. We will, however, give at least some motivation for each of the problems considered.

4.1. **Sunflowers.** If you thought addition and lines were pretty basic mathematical objects, let's leave them behind for a moment to get even more basic: we'll just consider sets and their intersections. A *sunflower* is a collection of sets $A_1, \dots, A_k$ such that their pairwise intersections are the same as their k-wise intersection: $A_i \cap A_j = A_1 \cap A_2 \cap \cdots \cap A_k$ for all $i \neq j$. (If you draw the Venn diagram of such a collection of sets, you'll see where the name comes from.) This notion was introduced by Erdős and Rado [46] in 1960 as a generalization of Dirichlet's box argument, and it has since found many uses in combinatorics, number theory, and computer science (see the introduction to [8] for many excellent references).

Dirichlet's box argument says that for any finite list $x_0, \dots, x_{a^2}$ of a^2+1 elements, there is a sublist of size $(a+1)$ such that either all the elements of the sublist are equal, or all are distinct from one another. This should sound a little familiar, as $x + y + z = 0$ in $\mathbb{Z}/3\mathbb{Z}$ if and only if $x = y = z$ or x, y, z are all distinct from one another. We'll see that the connection between sunflowers and cap sets is very tight indeed.

Theorem 4.1 (Erdős and Rado [46])**.** *Let $\mathcal{F}$ be a family of sets each of size s. If $|\mathcal{F}| \geq (k-1)^s s!$, then $\mathcal{F}$ contains a k-sunflower.*

Conjecture 4.2 (The Sunflower Conjecture [46])**.** *For every $k > 0$, there is a constant c_k such that "$(k-1)^s s!$" in the above theorem can be replaced by "$(c_k)^s$."*

This conjecture itself has also had many applications in extremal graph theory, the construction of Ramsey graphs, and circuit complexity (again, see [8] for references).

A slight variant of the Sunflower Conjecture, which will bring us even closer to cap sets, follows.

Conjecture 4.3 (The Erdős–Szemerédi Sunflower Conjecture [47])**.** *There is a constant $c < 2$ such that any family $\mathcal{F}$ of subsets of $[n] = \{1, \dots, n\}$ of size $|\mathcal{F}| \geq c^n$ contains a 3-sunflower.*

The difference between this conjecture and the preceding one is that this one doesn't require every set in $\mathcal{F}$ to have the same size, it *does* depend on the size of the ambient set from which $\mathcal{F}$ is built, and it only posits the existence of a 3-sunflower (instead of k-sunflowers for arbitrary k).

In connection with the complexity of matrix multiplication (see Section 4.2), Alon, Shpilka, and Umans [8] studied the Sunflower Conjecture and several of its variants, showing implications and equivalences between them. They introduced the following notion, which draws out the connection with cap sets.

Definition 4.4 (Sunflowers in $(\mathbb{Z}/m\mathbb{Z})^n$ [8, Definition 2.7])**.** A *k-sunflower in $(\mathbb{Z}/m\mathbb{Z})^n$* is a collection of k vectors $v_1, \dots, v_k \in (\mathbb{Z}/m\mathbb{Z})^n$ such that for every coordinate $i \in [n]$, either all the v_j have the same value in their ith coordinate or these values are all distinct. Equivalently, for each i, $|\{(v_1)_i, (v_2)_i, \dots, (v_k)_i\}|$ must be either 1 or k.

This is equivalent to a k-sunflower of sets (the usual notion) if the ambient set is partitioned into n pairwise disjoint blocks of size m, and every set in $\mathcal{F}$ contains exactly one element from each block.

Observation 4.5. A 3-sunflower in $(\mathbb{Z}/3\mathbb{Z})^n$ is the same as a cap set.

Conjecture 4.6 (Sunflower conjecture in $(\mathbb{Z}/m\mathbb{Z})^n$ [8, Conjecture 2.8])**.** *For every k, there is a constant b_k such that for all m, n, any set of $\geq (b_k)^n$ vectors in $(\mathbb{Z}/m\mathbb{Z})^n$ contains k vectors forming a k-sunflower.*

While this sounds different from the original Sunflower Conjecture, Alon, Shpilka, and Umans showed that the two are actually equivalent [8, Theorem 2.9].

By Observation 4.5, the Cap Set Conjecture thus resolves the $k = 3$ case of a weak form of Conjecture 4.6, in which we also restrict m to be 3. We'll see in the next section that the same method used to resolve the Cap Set Conjecture was then applied to resolve the full Sunflower Conjecture.

4.2. **Algorithms for matrix multiplication, and tricolored sum-free sets in other abelian groups.** Multiplying matrices—and its computationally equivalent sibling, solving linear systems of equations—is a fundamental linear algebra primitive used throughout the algorithmic world. Understanding its complexity is a central question in algebraic complexity theory that has led to new insights and conjectures in the representation theory of finite groups and algebraic geometry (see, e.g., Landsberg [84] and references therein).

The naive method of multiplying two $n \times n$ matrices takes $O(n^3)$ steps, which was thought to be optimal until Strassen showed [124] that this could be done in only $O(n^{2.81\cdots})$ steps. This led to the introduction of the *exponent ω of matrix multiplication*, namely

$$\omega = \inf\{w : n \times n \text{ matrices can be multiplied in } O(n^w) \text{ steps }\}.$$

The best lower bound known [85], although significantly nontrivial to prove, is still only a constant multiple of the obvious $\Omega(n^2)$: Any algorithm must at least read all $2n^2$ entries of the input matrices. Currently, the best algorithm known takes $O(n^{2.3729\cdots})$ steps [86], but it is a folklore conjecture that $\omega = 2$. Closing this gap is a major open problem in algebraic complexity theory.

Starting in 1969 with Strassen's result, there was a relatively steady stream of improvements to the best upper bound for ω. This culminated in 1990 when Coppersmith and Winograd [26] used the Salem–Spencer construction [110] of arithmetic-progression-free sets to develop an infinite family of matrix multiplication algorithms, whose exponent was limited to $2.375477\cdots$. This was the first hint of a relationship between matrix multiplication and arithmetic progressions. Then progress on ω hit a standstill for 20 years.

Although improvements in the upper bound on ω would wait until 2010, in 2003 Cohn and Umans [25] introduced a new approach to algorithms for matrix multiplication, which will draw out just how deep the connection is between such algorithms and arithmetic progressions. Briefly, their approach requires finding finite groups with only low-dimensional irreducible representations and containing three subsets satisfying a certain condition (see Definition 4.7 below). In 2005, with Kleinberg and Szegedy [24] they showed how to use this approach to develop new algorithms, and to capture the Coppersmith–Winograd algorithm as a Cohn–Umans-style construction in abelian groups of bounded exponent (that is, an infinite family of finite abelian groups such that every element of every group in the family had order $\leq b$). Starting in 2010, by analyzing higher tensor powers of the basic object used by Coppersmith and Winograd, Stothers [123], then Vassilevska Williams [135],

and finally Le Gall [86] made improvements, resulting in the current world record $\omega < 2.3728639\cdots$. However, it was then shown [9] that this particular technique—analyzing higher tensor powers of Coppersmith–Winograd—could get an exponent no better than $2.3078\cdots$.

Shortly after the resolution of the Cap Set Conjecture [39], Blasiak, Church, Cohn, Grochow, Naslund, Sawin, and Umans in [14], and independently N. Alon, showed that Cohn–Umans-style constructions in an abelian group yielded not cap sets, but tricolored sum-free sets (first introduced in connection with matrix multiplication in [8], and elaborated in [14]). To get a sense for where these come from, let's see how Cohn and Umans proposed using a finite group to multiply matrices. The idea is that the group algebra $\mathbb{C}[G]$ is a direct sum of matrix algebras $\mathbb{C}[G] \cong M_{d_1}(\mathbb{C}) \oplus \cdots \oplus M_{d_c}(\mathbb{C})$, where the d_i are the dimensions of the irreducible representations of G. If we could somehow embed $n \times n$ matrix multiplication into $\mathbb{C}[G]$ with $n > \max\{d_i\}$, then we could recursively multiply the smaller $d_i \times d_i$ matrices in order to multiply $n \times n$ matrices, thereby getting a nontrivial algorithm. One then gets a nontrivial upper bound on ω as the infimum of w falsifying the inequality $n^w \leq \sum_i d_i^w$ [25].

To embed a matrix product larger than any of the d_i into a group algebra, Cohn and Umans proposed the following construction. If we want to multiply $A \cdot B = C$, we'll use three subsets $S, T, U \subseteq G$, such that A is an $|S| \times |T|$ matrix and B is a $|T| \times |U|$ matrix. We embed A into the group algebra as $\iota_1(A) = \sum_{i,j} a_{ij} s_i t_j^{-1}$ and B into the group algebra as $\iota_2(B) = \sum_{j,k} b_{j,k} t_j u_k^{-1}$. We would like to be able to read off the entries of C as the coefficients of the group elements $s_i u_k^{-1}$ in the product $\iota_1(A)\iota_2(B)$. When we perform this multiplication, however, we end up with $\sum_{i,j,j',k} a_{i,j} b_{j',k} s_i t_j^{-1} t_{j'} u_k^{-1}$. If the only way that a group element $s_i t_j^{-1} t_{j'} u_k^{-1}$ can be of the form $s_{i'} u_{k'}^{-1}$ is with $i = i'$, $j = j'$, and $k = k'$, then indeed we get that the coefficient of $s_i u_k^{-1}$ in the product is precisely $\sum_j a_{i,j} b_{j,k} = c_{i,k}$, as desired. Rewriting this condition we have the following.

Definition 4.7 (Triple product property (TPP) [25])**.** Given a group G, three subsets $S, T, U \subseteq G$ satisfy the *triple product property* if

$$s_{i'}^{-1} s_i t_j^{-1} t_{j'} u_k^{-1} u_{k'} = 1 \iff i = i' \text{ and } j = j' \text{ and } k = k'.$$

In fact, the constructions of [24, 26, 86, 123, 135] are all instances of a generalization of this called the *simultaneous* triple product property—in which one embeds several independent copies of matrix multiplication simultaneously—but the preceding definition is already enough to give us the flavor of the connection with tricolored sum-free sets. For if we write $Q(S) = S^{-1}S = \{s^{-1}s' : s, s' \in S\}$, then the TPP can be rewritten as $\forall q_1 \in Q(S)$, $q_2 \in Q(T)$, $q_3 \in Q(U)$,

$$q_1 q_2 q_3 = 1 \iff q_1 = q_2 = q_3 = 1.$$

This condition is precisely the nonabelian generalization of the defining condition of a tricolored sum-free set (the nonabelian version is sometimes called a *multiplicative matching* [1, 114]). And here, we finally see where this notion of tricolored came from: it is because we wanted three *different* sets to index the rows of A, the rows of B, and the columns of C.

Blasiak et al. [14] extended the Ellenberg–Gijswijt bound from vector spaces $(\mathbb{Z}/p\mathbb{Z})^n$ to $(\mathbb{Z}/m\mathbb{Z})^n$ for arbitrary m, and even more generally to abelian groups

of bounded exponent. This generalization implies the $(\mathbb{Z}/m\mathbb{Z})^n$ Sunflower Conjecture 4.6, which was previously shown equivalent to the original Sunflower Conjecture 4.2 [8]. Additionally, using the connection between TPP constructions and tricolored sum-free sets, they showed the following.

Theorem 4.8 (Blasiak, Church, Cohn, Grochow, Nasland, Sawin, and Umans [14])**.** *One cannot show that $\omega = 2$ using simultaneous TPP constructions in families of abelian groups of bounded exponent.*

This includes and goes significantly beyond the class of Coppersmith–Winograd-style algorithms [26, 86, 123, 135]. This result was recently extended to certain non-abelian groups as well [15] by generalizing the Croot–Lev–Pach method to the group rings $\mathbb{F}_p[G]$ when G is an arbitrary (not necessarily abelian!) finite p-group. For limitations on a different generalization of the Coppersmith–Winograd technique, see [2, 3]. Thus, if $\omega = 2$, proving so requires significantly new techniques—perhaps Cohn–Umans-style constructions in nonnilpotent groups?

4.3. **Triangle removal.** Szemerédi's Regularity Lemma is a powerful tool in graph theory, essentially giving the structure of an arbitrary graph. A well-known consequence of the regularity lemma is the following.

Theorem 4.9 (Triangle Removal Lemma, Szemerédi and Ruzsa [109])**.**[10] *If a graph G on n vertices contains only $o(n^3)$ triangles, then by removing only $o(n^2)$ edges one can make the resulting graph triangle-free. More precisely, if G has $\leq \delta n^3$ triangles, then one can remove $\varepsilon(\delta)n^2$ edges, where $\varepsilon(\delta) \to 0$ as $\delta \to 0$.*

While on the surface this seems to have little to do with arithmetic progressions and cap sets, we note that the Triangle Removal Lemma can be used to give a very simple proof [109] of Roth's result [108] that $r_3(\mathbb{Z}/N\mathbb{Z}) \leq o(N)$. To see the connection, we give the brief proof here.

Proof of Roth's Theorem from the Triangle Removal Lemma [109]. Suppose $A \subseteq [N]$ has size $|A| \geq \varepsilon N$. We build a graph G as follows. Its vertex set V will be the disjoint union of three sets V_1, V_2, V_3, each of size $3N$, which we identify with $[3N]$ (so $|V| = 9N$). The edges are as follows: $(i, j) \in V_1 \times V_2$ is an edge if and only if $j - i \in A$; $(j, k) \in V_2 \times V_3$ is an edge if and only if $k - j \in A$; and $(k, i) \in V_3 \times V_1$ is an edge if and only if $\frac{k-i}{2} \in A$. There are no other edges. Then $(i, j, k) \in V_1 \times V_2 \times V_3$ form a triangle if and only if $j - i = a_1 \in A$ and $k - j = a_3 \in A$ and $\frac{k-i}{2} = a_2 \in A$ if and only if a_1, a_2, a_3 is an arithmetic progression in A, for we have $a_2 - a_1 = \frac{k+i}{2} - j = a_3 - a_2$. Note that this also allows the trivial arithmetic progression (a, a, a), as nothing here forces the difference $\frac{k+i}{2} - j$ to be nonzero. For each $i \in [3N]$ and each $a \in A$ we get a triangle corresponding to the trivial arithmetic progression (a, a, a), namely the triangle with vertices $i \in V_1$, $i + a \in V_2$, and $i + 2a \in V_3$. Since $|A| \geq \varepsilon N$, we thus have at least $|A||V_1| \geq 3\varepsilon N^2$ triangles in G. Furthermore, these triangles are all disjoint from one another, so to make G triangle-free would require removing at least $3\varepsilon N^2$ edges (one for each such triangle). As this is *not* $o(|V|^2) = o(81N^2) = o(N^2)$, it must *not* be the case

[10] This well-known version of the lemma appears to have first been stated in [55, Theorem 7], where the result is actually proved for arbitrary graphs H in place of triangles, and in [7, Proposition 4.4] where it is shown for cliques of arbitrary size. See also [44]. Although not phrased this way in [109], in some sense this was just because the community at the time was focused on other questions and missed asking for such a clean result. We thank Rödl for this historical insight.

that G has only $o(N^3)$ triangles, by the Triangle Removal Lemma. In other words, there is some $\delta > 0$ that depends only on ε (but not on A nor N) such that G contains at least $\delta|V|^3 = 729\delta N^3$ triangles. However, the total number of triangles corresponding to the trivial arithmetic progressions is $|A||V_1| \leq 3N^2$, so at least $729\delta N^3 - 3N^2$ of the triangles correspond to proper arithmetic progressions in A. In particular, since δ is independent of N, for sufficiently large N it must be the case that A contains at least one proper arithmetic progression of length 3. □

With this connection in mind, it is natural to define a *triangle* in an abelian group Z to be three elements $x, y, z \in Z$ such that $x + y + z = 0$.

Theorem 4.10 (Green [66])**.** *Let Z be an abelian group of order N. If $A \subseteq Z$ has only $o(N^2)$ triangles, then by removing only $o(N)$ elements from A, one can make the resulting set triangle-free.*

As with the original Triangle Removal Lemma, we can rephrase this in terms of δ, ε. Unfortunately, the (previous) best quantitative upper bound was that $1/\delta$ was of the form $2^{2^{2^{\cdot^{\cdot^{\cdot}}}}}$, where the height of this tower was $\log(1/\varepsilon)$ [49]. But with the techniques used to resolve the Cap Set Conjecture (and the result of [14]), Fox and Lovász showed a tight bound, dropping this from an exponential tower all the way down to a polynomial! They also generalized it from a single set A to a tricolored version: given $A, B, C \subseteq Z$, we say $(a, b, c) \in A \times B \times C$ is a triangle if $a+b+c = 0$.

Theorem 4.11 (Fox and Lovász [50])**.** *For each prime p there is a constant C_p such that the following holds. If $A, B, C \subseteq (\mathbb{Z}/p\mathbb{Z})^n$ have only δN^2 tricolored triangles where $\delta = (\varepsilon/3)^{C_p}$ (and $N = p^n$, as usual), then by removing only εN elements from $A \cup B \cup C$, one can make the resulting set triangle-free. Furthermore, this is essentially tight, in that it only holds with $\delta \leq \varepsilon^{C_p - o(1)}$.*

Continuing the recurring theme of the relationship between $(\mathbb{Z}/p\mathbb{Z})^n$ and $\mathbb{Z}/N\mathbb{Z}$, Aaronson [1] extended this connection between tricolored sum-free sets and triangles to $\mathbb{Z}/N\mathbb{Z}$.

4.4. **Matrix rigidity.** A natural question in computational complexity is: For a fixed matrix A, how hard is it to compute the function $x \mapsto Ax$? The naive approach, for $n \times n$ matrices, takes $O(n^2)$ arithmetic operations. There are several famous matrices for which this number can be reduced, most notably Fourier matrices, which can be applied in only $O(n \log n)$ operations, nearly linear in the dimension of the vector space. Aside from a few other highly structured classes of matrices, very little is known about this question in general. (And if we can't even answer this question with modern techniques, what hope do we have of proving $\mathsf{P} \neq \mathsf{NP}$ any time soon?) Two natural properties of a matrix A that make the corresponding linear function easy to compute are (1) sparsity, that is, if A has only very few nonzero entries, or (2) low rank. And any two such "easy" cases can be added together. That is, if an $n \times n$ matrix A is the sum of a matrix A' with only s nonzero entries and a matrix A'' of rank r, then $x \mapsto Ax$ can be computed in $O(s + rn)$ arithmetic operations. For a given matrix A, this raises the question of how few entries need to be changed to make the difference have low rank.

Definition 4.12 (Matrix rigidity [133])**.** The *rank-r rigidity* of a matrix A, denoted $R_A(r)$, is the least number s such that A is the sum of a matrix A' with $\leq s$ nonzero entries and a matrix A'' of rank $\leq r$.

Perhaps the most natural way to express the computation of an n-dimensional linear function $x \mapsto Ax$ is with a *linear circuit*: a sequence of instructions $g_1, \dots, g_\ell$ either of the form $g_i = x_j$ for some coordinate x_j of the input, or a linear combination $g_j = \alpha_{j1} g_{j_1} + \alpha_{j2} g_{j_2}$ for constants α_{jk} and previous instructions $j_1, j_2 < j$. The "output" of such a sequence is its last n values. The *size* of the linear circuit is the number ℓ of instructions. The $O(n \log n)$-step algorithm for the Fourier transform, for example, translates into a linear circuit of size $O(n \log n)$. To any such linear circuit we may naturally associate a directed acyclic graph on vertex set $[\ell]$ with arrows from $g_i \to g_j$ if g_i appears as a summand in the instruction g_j. The *depth* of a linear circuit is the length of the longest directed path in this graph.

Theorem 4.13 (Valiant [133])**.** *For every n, let A_n be an $n \times n$ matrix over a field. If $R_{A_n}(n/\log\log n) \geq \Omega(n^{1+\varepsilon})$ for some $\varepsilon > 0$, then for sufficiently large n, the linear function A_n cannot be computed by linear circuits of size $O(n)$ and depth $O(\log n)$.*

It is not hard to see that $R_A(r) \leq (n-r)^2$ for all r, and Valiant proved that almost all matrices are at least this rigid. However, to date, the best lower bound on any *explicit* matrix A is $R_A(r) \geq \Omega(\frac{n^2}{r} \log \frac{n}{r})$ [54, 120]. Other techniques that have been used to study rigid matrices include elimination theory [81], degree bounds [91, 92], spectral methods [77], and algebraic geometry [58]. For a mostly up-to-date survey that also includes relations to other areas, see [93]. For a long time it was believed that the Hadamard matrices were sufficiently rigid to apply Theorem 4.13, but this was recently disproved [4].

And now, we can add to this list of techniques the Croot–Lev–Pach polynomial method.

Theorem 4.14 (Dvir and Edelman [32])**.** *Let p be a fixed prime, and let $\varepsilon > 0$. For any function $f\colon (\mathbb{Z}/p\mathbb{Z})^n \to \mathbb{Z}/p\mathbb{Z}$, let $N = p^n$, and let M_f be the $N \times N$ matrix $M_f(x, y) = f(x + y)$. Then there is a $\delta > 0$ such that for all sufficiently large n, $R_{M_f}(N^{1-\delta}) \leq N^{1+\varepsilon}$. In particular, such matrices are not rigid enough to apply Theorem* 4.13*.*

5. Conclusion and outlook

Originally motivated by trying to find structure in the prime numbers, we were led to study arithmetic progressions in vector spaces over $\mathbb{Z}/p\mathbb{Z}$ as a model for arithmetic progression in the primes or in $\mathbb{Z}$. This turns out to be quite a fruitful toy model, and the Cap Set Conjecture was developed as a keystone problem whose solution was expected to unlock the mysteries of many other problems in combinatorics and number theory. And indeed, as evidenced by the long list of applications already, the technique used to resolve the Cap Set Conjecture had precisely the desired effect! (It may be worth noting that almost none of these applications follow as corollaries of the result itself; they only followed by using the Croot–Lev–Pach *technique*.)

Of course, the Erdős Conjecture on arithmetic progression (Conjecture 1.1) stands out as one of the most significant open problems we've discussed. Closing the gap between the best known upper and lower bounds for $r_3(\mathbb{Z}/N\mathbb{Z})$ and $r_3((\mathbb{Z}/p\mathbb{Z})^n)$ is also an interesting open problem. On the one hand, closing the gap for $r_3((\mathbb{Z}/p\mathbb{Z})^n)$ may seem like something of a "clean-up operation", given how close the bounds now are. On the other hand, the discussion in Section 3 reveals

that closing these gaps seems to require really novel methods, and one might hope that such new methods would have other applications.

In addition to the many variants of the cap set question already resolved by the polynomial method, there are many other variants of of the cap set question that remain open. For example, what about $r_4((\mathbb{Z}/p\mathbb{Z})^n)$, the size of the largest subset of $(\mathbb{Z}/p\mathbb{Z})^n$ free of *four*-term arithmetic progressions? As discussed by Tao [130], the most natural extension of the slice-rank method yields only trivial bounds for this question. Another example asks about sets that are free of arithmetic progressions *in particular directions*. Namely, if Z is an abelian group and $D \subseteq Z$, we can ask for the size of the largest subset $A \subseteq Z$ which contains no arithmetic progressions of the form $a, a+d, a+2d$ *with* $d \in D$. In particular, Lev asked about the size of the largest subset of $(\mathbb{Z}/3\mathbb{Z})^n$ free of arithmetic progressions whose common difference was in $\{0,1\}^n$; is it bounded by c^n for some $c < 3$?

As another example, while in $(\mathbb{Z}/3\mathbb{Z})^n$, the concepts of an arithmetic-progression-free set and a line-free set coincide; in $(\mathbb{Z}/m\mathbb{Z})^n$ with $m > 3$ they differ. What is the size of the largest line-free set in $(\mathbb{Z}/4\mathbb{Z})^n$? In particular, is it $< c^n$ for some $c < 4$? In $\mathbb{F}_q^n$ this question was recently answered by Bennett [13]. For more open questions in along these lines, and in additive combinatorics more generally, see [27].

Along the way, we've discussed many constructions of additive sets with various properties [12, 17, 22, 34–37, 70, 79, 103]. Are these constructions native to the groups they were designed for? Or can they be used in other groups as well? It may be fruitful to study this question from the following angle. Along with the notion of Freiman isomorphism, for any additive set A there is a notion of a "universal ambient group" for that additive set:

Definition 5.1 (Universal ambient group (see, e.g., [132, Section 5.5]))**.** Let $A \subseteq Z$ be an additive set and $k \in \mathbb{N}$. An abelian group U is a *universal ambient group* (of order k) for A if there is a Freiman k-isomorphism $A \cong A' \subseteq U$, and for every additive set $B \subseteq W$, every Freiman k-homomorphism $A' \to B$ extends to a unique group homomorphism $U \to W$.

Tao and Vu [132] showed that universal ambient groups always exist, and coined the term, though the idea was embedded in [80] (see [71, Chapter 20 Notes] for details; that chapter is also a good source for further results on universal ambient groups). As a way of getting at the preceding questions, it might be interesting to develop a generalization of Freiman homomorphisms for *multi-colored additive sets* (whatever that ought to mean), and to determine the universal ambient groups for the constructions mentioned in this exposition.

Finally, given all the applications of the Croot–Lev–Pach technique in such a short period of time, what other applications of the polynomial method are waiting to be explored?

Appendix A. Quick review of asymptotic growth

In this appendix we consider many quantities as a function of some (usually integer) parameter N, as $N \to \infty$. This allows us to get at the essence of certain constructions and bounds—and to compare different constructions with one another—without getting caught up in the details of their *exact* sizes (which can often be hard to compute) or how large N must be before one sees a difference between two techniques.

The most common notations we will be using follow.

- $f(N) \sim g(N)$. $\lim_{N\to\infty} \frac{f(N)}{g(N)} = 1$, and we say that f and g are *asymptotically equal*. The advantage of this notation is that it lets us focus on the highest-order terms only.
- $f(N) \leq O(g(N))$. There is a constant $c > 0$ such that for all sufficiently large N (that is, "there is an N_0 such that for all $N > N_0$") $f(N) \leq cg(N)$. For most purposes, it is equivalent to say that $\lim_{N\to\infty} \frac{f(N)}{g(N)}$ is finite (these are not entirely equivalent as there are functions f, g such that this limit doesn't exist, yet nonetheless $f(N) \leq O(g(N))$, but I don't think we encounter any such pathologies here).

  Similarly, we may use the notation $O(g(N))$ in a formula to denote an unspecified function f such that $f \leq O(g)$, e.g., $O(n^2)e^n$. The advantage of this notation is that it lets us focus on the highest-order terms *and* not worry about multiplicative constants, that are independent of N. For example, N and $100N$ are both $O(N)$, even though $N \not\sim 100N$.
- $f(N) \leq o(g(N))$. For *all* $c > 0$, for all sufficiently large N, $f(N) \leq cg(N)$. Equivalently, $\lim_{n\to\infty} \frac{f(N)}{g(N)} = 0$. If you think of O as the asymptotic version of $\leq$, then o is the asymptotic version of $<$. Again, we may use $o(g(N))$ in a formula to denote an unspecified f such that $f \leq o(g)$.
- $o(1)$. This is a particular case of the preceding that shows up frequently, namely an unspecified function of N that goes to zero as N goes to infinity. This is especially useful when it appears in exponents, such as $c^{n(1+o(1))}$. For example, $c^n n^2 = c^{n+2\log_c n} = c^{n(1+(1/n)2\log_c n)} = c^{n(1+o(1))}$. The advantage of this notation is that it lets us focus on the exponential growth rate without worrying about lower-order *multiplicative* terms (even when they depend on N). For we have that $c < d$ if and only if $c^{n(1+o(1))} = o(d^{n(1+o(1))})$, even if the two $o(1)$ terms are different. Let $f(N), g(N) \leq o(1)$. Then

$$\begin{aligned}\lim_{n\to\infty} \frac{c^{n(1+f(n))}}{d^{n(1+g(n))}} &= \lim_{n\to\infty} \left(\frac{c^{1+f(n)}}{d^{1+g(n)}}\right)^n \\ &= \lim_{n\to\infty} \left(\frac{c}{d}\right)^n = 0.\end{aligned}$$

  The jump from the first to second line here is allowed because if $c < d$, then there is some $\varepsilon_0 > 0$ such that if $0 < \varepsilon < \varepsilon_0$, we have $c^{1+\varepsilon} < d^{1+\varepsilon}$, and there is some n_0 such that for all $n > n_0$, we have $f(n), g(n) < \varepsilon_0$.

Here is a list of the most frequent growth rates we'll be considering and the relations between them. If you haven't seen these before, working out the relations for yourself is a nice but not terribly difficult exercise that helps acquaint you with these growth rates. They are listed in strictly increasing order, so that if you see $f(N), g(N)$ in this list, it means that $f(N) \leq o(g(N))$.

$$(\log N)^{1/2}, \log N, (\log N)^2, (\log N)^3, \ldots, \sqrt{N}, N^{0.75}, N, N^2, N^3, \ldots, 1.01^N, 2^N, e^N, 3^N.$$

Indeed, $(\log N)^c \leq o((\log N)^d)$ for constants $c, d > 0$ if and only if $c < d$, and, similarly, $N^c \leq o(N^d)$ and $\exp(cn) \leq o(\exp(dn))$ if and only if $c < d$. In particular, for exponentials of the form c^N $(c > 1)$, the base of the exponent matters: $c^N \leq o(d^N)$ if and only if $1 < c < d$.

Furthermore, for growth rates that are exponentially separated, altering constant exponents never changes this, for example, $(\log\log N)^c \leq o((\log N)^\varepsilon)$ for any constants $c > 0, \varepsilon > 0$, no matter how large c is and how small ε is. Similarly, $(\log N)^c \leq o(N^\varepsilon)$, and $N^c \leq o((1+\varepsilon)^n)$ for all $c > 0, \varepsilon > 0$.

Acknowledgments

I thank Jonah Blasiak, Thomas Church, Henry Cohn, and Chris Umans for collaborating with me on [14, 15], which got me into this whole area, a collaboration partially supported by an AIM SQuaRE grant. I thank Vojtěch Rödl for the history of the Triangle Removal Lemma (see footnote 10 on Theorem 4.9). I thank Thomas Church, Jordan Ellenberg, Dion Gijswijt, Ben Green, Vsevolod Lev, Terence Tao, Avi Wigderson, and an anonymous reviewer for reading drafts and providing references and useful feedback which improved the exposition.

About the author

Joshua A. Grochow is assistant professor in the departments of computer science and mathematics at University of Colorado, Boulder. Prior to that he was an Omidyar Postdoctoral Fellow at Sante Fe Institute, and a postdoc in the computer science theory group at University of Toronto. His research has two closely related foci: interactions between computational complexity and other areas of mathematics (representation theory, group theory, and algebraic geometry), and developing rigorous mathematical theory for the study of complex systems.

References

[1] J. Aaronson, *A connection between matchings and removal in abelian groups*, arXiv:1612.05172 [math.CO] (2016).

[2] J. Alman and V. V. Williams, *Further limitations of the known approaches for matrix multiplication*, 9th Innovations in Theoretical Computer Science, LIPIcs. Leibniz Int. Proc. Inform., vol. 94, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018, pp. Art. No. 25, 15. MR3761761

[3] J. Alman and V. V. Williams, *Limits on all known (and some unknown) approaches to matrix multiplication*, FOCS '18: 59th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, 2018.

[4] J. Alman and R. Williams, *Probabilistic rank and matrix rigidity*, STOC'17—Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, 2017, pp. 641–652. MR3678217

[5] N. Alon and M. Dubiner, *Zero-sum sets of prescribed size*, Combinatorics, Paul Erdős is eighty, Vol. 1, Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 1993, pp. 33–50. MR1249703

[6] N. Alon and M. Dubiner, *A lattice point problem and additive number theory*, Combinatorica **15** (1995), no. 3, 301–309, DOI 10.1007/BF01299737. MR1357277

[7] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, and R. Yuster, *The algorithmic aspects of the regularity lemma*, J. Algorithms **16** (1994), no. 1, 80–109, DOI 10.1006/jagm.1994.1005. MR1251840

[8] N. Alon, A. Shpilka, and C. Umans, *On sunflowers and matrix multiplication*, Comput. Complexity **22** (2013), no. 2, 219–243, DOI 10.1007/s00037-013-0060-1. MR3055780

[9] A. Ambainis, Y. Filmus, and F. Le Gall, *Fast matrix multiplication: limitations of the Coppersmith-Winograd method*, STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing, ACM, New York, 2015, pp. 585–593. MR3388238

[10] L. Bary-Soroker, *Prime tuples in function fields*, Snapshots of modern mathematics from Oberwolfach **10** (2016). DOI 0.14760/SNAP-2016-010-EN.

[11] M. Bateman and N. H. Katz, *New bounds on cap sets*, J. Amer. Math. Soc. **25** (2012), no. 2, 585–613, DOI 10.1090/S0894-0347-2011-00725-X. MR2869028

[12] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U. S. A. **32** (1946), 331–332, DOI 10.1073/pnas.32.12.331. MR0018694
[13] M. Bennett, *Bounds on sizes of caps in $AG(n, q)$ via the Croot–Lev–Pach polynomial method*, arXiv:1806.05303 [math.CO] (2018).
[14] J. Blasiak, T. Church, H. Cohn, J. A. Grochow, E. Naslund, W. F. Sawin, and C. Umans, *On cap sets and the group-theoretic approach to matrix multiplication*, Discrete Anal. (2017), Paper No. 3, 27. MR3631613
[15] J. Blasiak, T. Church, H. Cohn, J. A. Grochow, and C. Umans, *Which groups are amenable to proving exponent two for matrix multiplication?*, arXiv:1712.02302 [math.GR] (2017).
[16] T. F. Bloom, *A quantitative improvement for Roth's theorem on arithmetic progressions*, J. Lond. Math. Soc. (2) **93** (2016), no. 3, 643–663, DOI 10.1112/jlms/jdw010. MR3509957
[17] R. C. Bose, *Mathematical theory of the symmetrical factorial design*, Sankhyā **8** (1947), 107–166. MR0026781
[18] J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. **9** (1999), no. 5, 968–984, DOI 10.1007/s000390050105. MR1726234
[19] J. Bourgain, *Roth's theorem on progressions revisited*, J. Anal. Math. **104** (2008), 155–192, DOI 10.1007/s11854-008-0020-x. MR2403433
[20] T. C. Brown and J. P. Buhler, *A density version of a geometric Ramsey theorem*, J. Combin. Theory Ser. A **32** (1982), no. 1, 20–34, DOI 10.1016/0097-3165(82)90062-0. MR640624
[21] J. P. Buhler, H. W. Lenstra Jr., and C. Pomerance, *Factoring integers with the number field sieve*, The development of the number field sieve, Lecture Notes in Math., vol. 1554, Springer, Berlin, 1993, pp. 50–94, DOI 10.1007/BFb0091539. MR1321221
[22] A. R. Calderbank and P. C. Fishburn, *Maximal three-independent subsets of $\{0, 1, 2\}^n$*, Des. Codes Cryptogr. **4** (1994), no. 3, 203–211, DOI 10.1007/BF01388452. MR1277940
[23] S. Chowla, *There exists an infinity of 3-combinations of primes in A. P*, Proc. Lahore Philos. Soc. **6** (1944), no. 2, 15–16. MR0014125
[24] H. Cohn, R. Kleinberg, B. Szegedy, and C. Umans, *Group-theoretic algorithms for matrix multiplication*, FOCS '05: 46th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, 2005, arXiv:math/0511460 [math.GR], 379–388 (2005).
[25] H. Cohn and C. Umans, *A group-theoretic approach to fast matrix multiplication*, FOCS '03: 44th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, 2003, arXiv:math/0307321 [math.GR], 438–449 (2003).
[26] D. Coppersmith and S. Winograd, *Matrix multiplication via arithmetic progressions*, J. Symbolic Comput. **9** (1990), no. 3, 251–280, DOI 10.1016/S0747-7171(08)80013-2. MR1056627
[27] E. S. Croot III and V. F. Lev, *Open problems in additive combinatorics*, Additive combinatorics, CRM Proc. Lecture Notes, vol. 43, Amer. Math. Soc., Providence, RI, 2007, pp. 207–233. MR2359473
[28] E. Croot, V. F. Lev, and P. P. Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*, Ann. of Math. (2) **185** (2017), no. 1, 331–337, DOI 10.4007/annals.2017.185.1.7. MR3583357
[29] B. L. Davis and D. Maclagan, *The card game SET*, Math. Intelligencer **25** (2003), no. 3, 33–40, DOI 10.1007/BF02984846. MR2005098
[30] M. S. Dousti and K. Ghasemloo, *Answers to "Adding integers represented by their factorization is as hard as factoring?"*, https://cstheory.stackexchange.com/q/7491/129 (2011).
[31] Z. Dvir, *On the size of Kakeya sets in finite fields*, J. Amer. Math. Soc. **22** (2009), no. 4, 1093–1097, DOI 10.1090/S0894-0347-08-00607-3. MR2525780
[32] Z. Dvir and B. Edelman, *Matrix rigidity and the Croot–Lev–Pach lemma*, arXiv:1708.01646 [cs.CC] (2017).
[33] Z. Dvir, S. Kopparty, S. Saraf, and M. Sudan, *Extensions to the method of multiplicities, with applications to Kakeya sets and mergers*, SIAM J. Comput. **42** (2013), no. 6, 2305–2328, DOI 10.1137/100783704. MR3143848
[34] Y. Edel, *Extensions of generalized product caps*, Des. Codes Cryptogr. **31** (2004), no. 1, 5–14, DOI 10.1023/A:1027365901231. MR2031694
[35] Y. Edel and J. Bierbrauer, *Large caps in small spaces*, Des. Codes Cryptogr. **23** (2001), no. 2, 197–212, DOI 10.1023/A:1011216716700. MR1830941
[36] Y. Edel, S. Ferret, I. Landjev, and L. Storme, *The classification of the largest caps in* AG$(5, 3)$, J. Combin. Theory Ser. A **99** (2002), no. 1, 95–110, DOI 10.1006/jcta.2002.3261. MR1911459

[37] M. Elkin, *An improved construction of progression-free sets*, Israel J. Math. **184** (2011), 93–128, DOI 10.1007/s11856-011-0061-1. MR2823971

[38] J. S. Ellenberg, *Sumsets as unions of sumsets of subsets*, Discrete Anal. (2017), Paper No. 14, 5. MR3695477

[39] J. S. Ellenberg and D. Gijswijt, *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression*, Ann. of Math. (2) **185** (2017), no. 1, 339–343, DOI 10.4007/annals.2017.185.1.8. MR3583358

[40] P. Erdős, *Résultats et problèmes en théorie des nombres*, Séminaire Delange–Pisot–Poitou (14e année: 1972/73), Théorie des nombres, Fasc. 2, Exp. No. 24, Secrétariat Mathématique, Paris, 1973, p. 7. MR0396376

[41] P. Erdős, *Problems in number theory and combinatorics*, Proceedings of the Sixth Manitoba Conference on Numerical Mathematics (Univ. Manitoba, Winnipeg, Man., 1976), Congress. Numer., XVIII, Utilitas Math., Winnipeg, Man., 1977, pp. 35–58. MR532690

[42] P. Erdős, *On the combinatorial problems which I would most like to see solved*, Combinatorica **1** (1981), no. 1, 25–42, DOI 10.1007/BF02579174. MR602413

[43] P. Erdős, *Some of my favorite problems and results*, The mathematics of Paul Erdős, I, Algorithms Combin., vol. 13, Springer, Berlin, 1997, pp. 47–67, DOI 10.1007/978-3-642-60408-9_3. MR1425174

[44] P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), no. 2, 113–121, DOI 10.1007/BF01788085. MR932119

[45] P. Erdős and C. Pomerance, *On the largest prime factors of n and $n+1$*, Aequationes Math. **17** (1978), no. 2-3, 311–321, DOI 10.1007/BF01818569. MR0480303

[46] P. Erdős and R. Rado, *Intersection theorems for systems of sets*, J. London Math. Soc. **35** (1960), 85–90, DOI 10.1112/jlms/s1-35.1.85. MR0111692

[47] P. Erdős and E. Szemerédi, *Combinatorial properties of systems of sets*, J. Combinatorial Theory Ser. A **24** (1978), no. 3, 308–313. MR0491202

[48] P. Erdös and P. Turán, *On Some Sequences of Integers*, J. London Math. Soc. **11** (1936), no. 4, 261–264, DOI 10.1112/jlms/s1-11.4.261. MR1574918

[49] J. Fox, *A new proof of the graph removal lemma*, Ann. of Math. (2) **174** (2011), no. 1, 561–579, DOI 10.4007/annals.2011.174.1.17. MR2811609

[50] J. Fox and L. M. Lovász, *A tight bound for Green's arithmetic triangle removal lemma in vector spaces*, Adv. Math. **321** (2017), 287–297, DOI 10.1016/j.aim.2017.09.037. MR3715712

[51] J. Fox and L. Sauermann, *Erdös-Ginzburg-Ziv constants by avoiding three-term arithmetic progressions*, Electron. J. Combin. **25** (2018), no. 2, Paper 2.14, 9. MR3799432

[52] P. Frankl, R. L. Graham, and V. Rödl, *On subsets of abelian groups with no 3-term arithmetic progression*, J. Combin. Theory Ser. A **45** (1987), no. 1, 157–161, DOI 10.1016/0097-3165(87)90053-7. MR883900

[53] G. A. Freĭman, *Foundations of a structural theory of set addition*, Translations of Mathematical Monographs, vol. 37, American Mathematical Society, Providence, RI, 1973. Translated from the Russian. MR0360496

[54] J. Friedman, *A note on matrix rigidity*, Combinatorica **13** (1993), no. 2, 235–239, DOI 10.1007/BF01303207. MR1237045

[55] Z. Füredi, *Extremal hypergraphs and combinatorial geometry*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, pp. 1343–1352, DOI 10.1007/978-3-0348-9078-6_65. MR1404036

[56] P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), no. 1, 4–9, DOI 10.1112/S0025579300016442. MR0409385

[57] G. Ge and C. Shangguan, *Rank counting and maximum subsets of $\mathbb{F}_q^n$ containing no right angles*, arXiv:1612.08255 [math.CO] (2016).

[58] F. Gesmundo, J. D. Hauenstein, C. Ikenmeyer, and J. M. Landsberg, *Complexity of linear circuits and geometry*, Found. Comput. Math. **16** (2016), no. 3, 599–635, DOI 10.1007/s10208-015-9258-8. MR3494506

[59] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. **8** (1998), no. 3, 529–551, DOI 10.1007/s000390050065. MR1631259

[60] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588, DOI 10.1007/s00039-001-0332-9. MR1844079

[61] W. T. Gowers, *What is difficult about the cap-set problem?*, Gowers's Weblog, https://gowers.wordpress.com/2011/01/11/what-is-difficult-about-the-cap-set-problem, (2011).

[62] W. T. Gowers, *Erdős and arithmetic progressions*, Erdös centennial, Bolyai Soc. Math. Stud., vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 265–287, DOI 10.1007/978-3-642-39286-3_8. MR3203599

[63] W. T. Gowers, *Reflections on the recent solution of the cap-set problem I*, Gowers's Weblog, `https://gowers.wordpress.com/2016/05/19/reflections-on-the-recent-solution-of-the-cap-set-problem-i/` (2016).

[64] B. Green, *Finite field models in additive combinatorics*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 1–27, DOI 10.1017/CBO9780511734885.002. MR2187732

[65] B. Green, *Roth's theorem in the primes*, Ann. of Math. (2) **161** (2005), no. 3, 1609–1636, DOI 10.4007/annals.2005.161.1609. MR2180408

[66] B. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), no. 2, 340–376, DOI 10.1007/s00039-005-0509-8. MR2153903

[67] B. Green, *Sárközy's theorem in function fields*, Q. J. Math. **68** (2017), no. 1, 237–242, DOI 10.1093/qmath/haw044. MR3658291

[68] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), no. 2, 481–547, DOI 10.4007/annals.2008.167.481. MR2415379

[69] B. Green, T. Tao, and T. Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$-norm*, Ann. of Math. (2) **176** (2012), no. 2, 1231–1372, DOI 10.4007/annals.2012.176.2.11. MR2950773

[70] B. Green and J. Wolf, *A note on Elkin's improvement of Behrend's construction*, Additive number theory, Springer, New York, 2010, pp. 141–144, DOI 10.1007/978-0-387-68361-4_9. MR2744752

[71] D. J. Grynkiewicz, *Structural additive theory*, Developments in Mathematics, vol. 30, Springer, Cham, 2013. MR3097619

[72] R. K. Guy, *Unsolved problems in number theory*, 3rd ed., Problem Books in Mathematics, Springer-Verlag, New York, 2004. MR2076335

[73] H. Harborth, *Ein Extremalproblem für Gitterpunkte* (German), J. Reine Angew. Math. **262/263** (1973), 356–360, DOI 10.1515/crll.1973.262-263.356. Collection of articles dedicated to Helmut Hasse on his seventy-fifth birthday. MR0327666

[74] D. R. Heath-Brown, *Integer sets containing no arithmetic progressions*, J. London Math. Soc. (2) **35** (1987), no. 3, 385–394, DOI 10.1112/jlms/s2-35.3.385. MR889362

[75] H. A. Helfgott and A. de Roton, *Improving Roth's theorem in the primes*, Int. Math. Res. Not. IMRN **4** (2011), 767–783, DOI 10.1093/imrn/rnq108. MR2773330

[76] S. Jukna, *Extremal combinatorics*: *With applications in computer science*, 2nd ed., Texts in Theoretical Computer Science. An EATCS Series, Springer, Heidelberg, 2011. MR2865719

[77] B. S. Kashin and A. A. Razborov, *New lower bounds for the stability of Hadamard matrices* (Russian, with Russian summary), Mat. Zametki **63** (1998), no. 4, 535–540, DOI 10.1007/BF02311250; English transl., Math. Notes **63** (1998), no. 3-4, 471–475. MR1680943

[78] T. Kim and S. Oum, *An upper bound on tricolored ordered sum-free sets*, arXiv:1708.07263 [math.CO] (2017).

[79] R. Kleinberg, W. F. Sawin, and D. E. Speyer, *The growth rate of tri-colored sum-free sets*, Discrete Anal. (2018), Paper No. 12.

[80] S. V. Konyagin and V. F. Lev, *Combinatorics and linear algebra of Freiman's isomorphism*, Mathematika **47** (2000), no. 1-2, 39–51 (2002), DOI 10.1112/S0025579300015709. MR1924486

[81] A. Kumar, S. V. Lokam, V. M. Patankar, and M. N. J. Sarma, *Using elimination theory to construct rigid matrices*, Comput. Complexity **23** (2014), no. 4, 531–563, DOI 10.1007/s00037-013-0061-0. MR3274825

[82] G. Kuperberg, *Answer to question "Open problems with monetary rewards" on MathOverflow*, https://mathoverflow.net/a/66219/38434 (2011).

[83] J. C. Lagarias, *The $3x + 1$ problem: an annotated bibliography (1963–1999)*, The ultimate challenge: the $3x + 1$ problem, Amer. Math. Soc., Providence, RI, 2010, pp. 267–341. MR2560718

[84] J. M. Landsberg, *Geometry and the complexity of matrix multiplication*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 2, 247–284, DOI 10.1090/S0273-0979-08-01176-2. MR2383305

[85] J. M. Landsberg, *New lower bounds for the rank of matrix multiplication*, SIAM J. Comput. **43** (2014), no. 1, 144–149, DOI 10.1137/120880276. MR3162411

[86] F. Le Gall, *Powers of tensors and fast matrix multiplication*, ISSAC 2014—Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2014, pp. 296–303, DOI 10.1145/2608628.2608664. MR3239939

[87] A. K. Lenstra, *Integer factoring*: *Towards a quarter-century of public key cryptography*, Des. Codes Cryptogr. **19** (2000), no. 2-3, 101–128, DOI 10.1023/A:1008397921377. MR1758972

[88] H. W. Lenstra Jr., *Factoring integers with elliptic curves*, Ann. of Math. (2) **126** (1987), no. 3, 649–673, DOI 10.2307/1971363. MR916721

[89] V. F. Lev, *Progression-free sets in finite abelian groups*, J. Number Theory **104** (2004), no. 1, 162–169, DOI 10.1016/S0022-314X(03)00148-3. MR2021632

[90] V. F. Lev, *Character-free approach to progression-free sets*, Finite Fields Appl. **18** (2012), no. 2, 378–383, DOI 10.1016/j.ffa.2011.09.006. MR2890558

[91] S. V. Lokam, *On the rigidity of Vandermonde matrices*, Theoret. Comput. Sci. **237** (2000), no. 1-2, 477–483, DOI 10.1016/S0304-3975(00)00008-6. MR1756225

[92] S. V. Lokam, *Quadratic lower bounds on matrix rigidity*, Theory and applications of models of computation, Lecture Notes in Comput. Sci., vol. 3959, Springer, Berlin, 2006, pp. 295–307, DOI 10.1007/11750321_28. MR2277251

[93] S. V. Lokam, *Complexity lower bounds using linear algebra*, Found. Trends Theor. Comput. Sci. **4** (2008), no. 1-2, front matter, 1–155 (2009), DOI 10.1561/0300000020. MR2539154

[94] L. M. Lovász and Lisa Sauermann, *A lower bound for the k-multicored sum-free problem in $\mathbb{Z}_m^n$*, arXiv:1804.08837 [math.CO] (2018).

[95] J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), no. 1, 383–413, DOI 10.4007/annals.2015.181.1.7. MR3272929

[96] L. McMahon, G. Gordon, H. Gordon, and R. Gordon, *The joy of SET*: *The many mathematical dimensions of a seemingly simple card game*, Princeton University Press, Princeton, NJ; National Museum of Mathematics, New York, 2017. MR3559496

[97] R. Meshulam, *On subsets of finite abelian groups with no 3-term arithmetic progressions*, J. Combin. Theory Ser. A **71** (1995), no. 1, 168–172, DOI 10.1016/0097-3165(95)90024-1. MR1335785

[98] P. Michel, *Busy beaver competition and Collatz-like problems*, Arch. Math. Logic **32** (1993), no. 5, 351–367, DOI 10.1007/BF01409968. MR1223396

[99] M. A. Morrison and J. Brillhart, *A method of factoring and the factorization of F_7*, Math. Comp. **29** (1975), 183–205, DOI 10.2307/2005475. Collection of articles dedicated to Derrick Henry Lehmer on the occasion of his seventieth birthday. MR0371800

[100] G. Muller, *Answer to "What is the importance of the Collatz conjecture?"*, `https://math.stackexchange.com/a/10608/224688` (2010).

[101] S. Norin, *A distribution on triples with maximum entropy marginal*, arXiv:1608.00243 [math.CO] (2016).

[102] L. Pebody, *Proof of a conjecture of Kleinberg–Sawin–Speyer*, Discrete Anal. (2018), Paper No. 13.

[103] G. Pellegrino, *Sul massimo ordine delle calotte in $S_{4,3}$* (Italian), Matematiche (Catania) **25** (1970), 149–157 (1971). MR0363952

[104] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci. **1** (2014), Art. 12, 83, DOI 10.1186/s40687-014-0012-7. MR3373710

[105] C. Pomerance, *Analysis and comparison of some integer factoring algorithms*, Computational methods in number theory, Part I, Math. Centre Tracts, vol. 154, Math. Centrum, Amsterdam, 1982, pp. 89–139. MR700260

[106] A. Potechin, *Maximal caps in* AG$(6, 3)$, Des. Codes Cryptogr. **46** (2008), no. 3, 243–259, DOI 10.1007/s10623-007-9132-z. MR2372838

[107] H. Riesel, *Prime numbers and computer methods for factorization*, 2nd ed., Progress in Mathematics, vol. 126, Birkhäuser Boston, Inc., Boston, MA, 1994. MR1292250

[108] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109, DOI 10.1112/jlms/s1-28.1.104. MR0051853

[109] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, Amsterdam-New York, 1978, pp. 939–945. MR519318

[110] R. Salem and D. C. Spencer, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U. S. A. **28** (1942), 561–563, DOI 10.1073/pnas.28.12.561. MR0007405

[111] T. Sanders, *Roth's theorem in* $\mathbb{Z}_4^n$, Anal. PDE **2** (2009), no. 2, 211–234, DOI 10.2140/apde.2009.2.211. MR2560257

[112] T. Sanders, *On Roth's theorem on progressions*, Ann. of Math. (2) **174** (2011), no. 1, 619–636, DOI 10.4007/annals.2011.174.1.20. MR2811612

[113] T. Sanders, *On certain other sets of integers*, J. Anal. Math. **116** (2012), 53–82, DOI 10.1007/s11854-012-0003-9. MR2892617

[114] W. F. Sawin, *Bounds for matchings in nonabelian groups*, arXiv:1702.00905 [math.CO] (2017).

[115] T. Schoen and I. D. Shkredov, *Roth's theorem in many variables*, Israel J. Math. **199** (2014), no. 1, 287–308, DOI 10.1007/s11856-013-0049-0. MR3219538

[116] T. Schoen and O. Sisask, *Roth's theorem for four variables and additive structures in sums of sparse sets*, Forum Math. Sigma **4** (2016), e5, 28, DOI 10.1017/fms.2016.2. MR3482282

[117] B. Segre, *Le geometrie di Galois* (Italian), Ann. Mat. Pura Appl. (4) **48** (1959), 1–96, DOI 10.1007/BF02410658. MR0116259

[118] B. Segre, *Introduction to Galois geometries* (English, with Italian summary), Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. I (8) **8** (1967), 133–236. MR0238846

[119] Set Enterprises, Inc., *Marsha Jean Falco—the creative genius behind SET®*, https://puzzles.setgame.com/set/history.htm.

[120] M. A. Shokrollahi, D. A. Spielman, and V. Stemann, *A remark on matrix rigidity*, Inform. Process. Lett. **64** (1997), no. 6, 283–285, DOI 10.1016/S0020-0190(97)00190-7. MR1608240

[121] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comput. **26** (1997), no. 5, 1484–1509, DOI 10.1137/S0097539795293172. MR1471990

[122] A. Soifer, *The mathematical coloring book*, Springer, New York, 2009. Mathematics of coloring and the colorful life of its creators; With forewords by Branko Grünbaum, Peter D. Johnson, Jr. and Cecil Rousseau. MR2458293

[123] A. Stothers, *On the complexity of matrix multiplication*, Ph.D. thesis, University of Edinburgh, Edinburgh, UK (2010).

[124] V. Strassen, *Gaussian elimination is not optimal*, Numer. Math. **13** (1969), 354–356, DOI 10.1007/BF02165411. MR0248973

[125] E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Acad. Sci. Hungar. **20** (1969), 89–104, DOI 10.1007/BF01894569. MR0245555

[126] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245, DOI 10.4064/aa-27-1-199-245. Collection of articles in memory of Juriĭ Vladimirovič Linnik. MR0369312

[127] E. Szemerédi, *Integer sets containing no arithmetic progressions*, Acta Math. Hungar. **56** (1990), no. 1-2, 155–158, DOI 10.1007/BF01903717. MR1100788

[128] T. Tao, *Open question: best bounds for cap sets*, What's new, `https://terrytao.wordpress.com/2007/02/23/open-question-best-bounds-for-cap-sets/` (2007).

[129] T. Tao, *Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory*, EMS Surv. Math. Sci. **1** (2014), no. 1, 1–46, DOI 10.4171/EMSS/1. MR3200226

[130] T. Tao, *A symmetric formulation of the Croot–Lev–Pach–Ellenberg–Gijswijt capset bound*, What's new, `https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/` (2016).

[131] T. Tao and W. F. Sawin, *Notes on the "slice rank" of tensors*, What's new, `https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/` (2016).

[132] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006. MR2289012

[133] L. G. Valiant, *Graph-theoretic arguments in low-level complexity*, Mathematical foundations of computer science (Proc. Sixth Sympos., Tatranská Lomnica, 1977), Springer, Berlin, 1977, pp. 162–176. Lecture Notes in Comput. Sci., Vol. 53. MR0660702

[134] J. G. van der Corput, *Über Summen von Primzahlen und Primzahlquadraten* (German), Math. Ann. **116** (1939), no. 1, 1–50, DOI 10.1007/BF01597346. MR1513216

[135] V. V. Williams, *Multiplying matrices faster than Coppersmith-Winograd*, STOC'12—Proceedings of the 2012 ACM Symposium on Theory of Computing, ACM, New York, 2012, pp. 887–898, DOI 10.1145/2213977.2214056. MR2961552

[136] S. S. Wagstaff Jr., *The joy of factoring*, Student Mathematical Library, vol. 68, American Mathematical Society, Providence, RI, 2013. MR3135977

[137] J. Wolf, *Finite field models in arithmetic combinatorics—ten years on*, Finite Fields Appl. **32** (2015), 233–274, DOI 10.1016/j.ffa.2014.11.003. MR3293412

[138] D. Zeilberger, *A motivated rendition of the Ellenberg–Gijswijt gorgeous proof that the largest subset of* F_3^n *with no three-term arithmetic progression is* $O(c^n)$*, with* $c = \sqrt[3]{(5589 + 891\sqrt{33}}/8 = 2.755104613023633000221 27\ldots$, arXiv:1607.01804 [math.CO] (2016).

[139] Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), no. 3, 1121–1174, DOI 10.4007/annals.2014.179.3.7. MR3171761

Department of Computer Science and Department of Mathematics, University of Colorado Boulder, Boulder, Colorado

Email address: jgrochow@colorado.edu