Research Problems in Arithmetic Combinatorics

Ernie Croot

July 13, 2006

1. (related to a quesiton of J. Bourgain) Classify all polynomials $f(x,y) \in \mathbb{Z}[x,y]$ which have the following property: There exists $\epsilon > 0$ such that for every prime p sufficiently large if $S \subseteq \mathbb{Z}_p$ satisfies $|S| = \lfloor p^{1/2} \rfloor$, then

$$|f(S,S)| \ > \ p^{1/2+\epsilon},$$

where

$$f(S,S) \ = \ \{f(x,y) \pmod p \ : \ x,y \in S\}.$$

Remark 1. The condition $|S| = \lfloor p^{1/2} \rfloor$ can be replaced with $|S| = p^\theta$, for arbitrary θ, but the question is a little messier to state in this more general case.

Remark 2. No linear polynomials f have the above property. There are degree 2 polynomials that are also pathological; for example, if $f(x,y) = x^2 + y^2$, and if we take

$$S \ = \ \{s \ : \ s^2 \equiv j \pmod p, \text{ where } |j| \lessapprox p^{1/2}\},$$

then $|f(S,S)| = O(p^{1/2})$.

2. (Due to many people, in particular Christian Elsholtz) Suppose S is a subset of the integers $\{1,...,N\}$ with the property that S occupies at most $2p/3$ residue classes modulo p for every prime $p < N^{1/2}$. Naively, we would expect such a set S to be very small, as the following heuristic

demonstrates: Select any collection of $\log^2 N$ primes up to $N^{1/2}$. Since S lies in $2p/3$ residue classes for each of these primes, if these primes were 'independent' in some sense, we would expect that

$$|S| \ < \ N(2/3)^{\log^2 N} \ = \ N^{o(1)}.$$

However, if S is the set of squares up to N, then $|S| \sim N^{1/2}$, and S occupies less than $2p/3$ residue classes modulo p for every prime $p < N^{1/2}$. Thus, the squares 'disprove' our heuristic.

My problem is: Classify all sets $S \subseteq [N] := \{1, ..., N\}$ where $|S| > N^{\epsilon}$, which occupy at most $2p/3$ residue classes modulo p for every prime $p < N^{1/2}$. I conjecture that such sets S are subsets of affine copies of the squares; that is, subsets of $\{an^2 + bn + c \ : \ n \in \mathbb{Z}\}$.

Recently, H. Helfgott and Venkatesh proved a certain "two dimensional" version of this question, which I will not bother to mention here.

3. This is in spirit similar to problem 2, but appears to be much easier, even through it is still unsolved: Suppose that $S \subseteq [N]$, and suppose that for every prime p satisfying

$$p \ < \ N^{\theta}$$

we have that S lies in an arithmetic progression of length $N^{\theta}/2$ modulo p. Must S itself be a subset of an arithmetic progression of size $O(N^{\theta})$?

4. Fix $c, d > 0$ and suppose p is a large prime number. Suppose $S \subseteq \mathbb{Z}_p$, and associate to S its indicator function $S(n)$. Suppose

$$|S| \ > \ p(\log p)^{-c}, \tag{1}$$

and suppose

$$||\hat{S}||_1 \ < \ p(\log p)^{d}. \tag{2}$$

In words, we are supposing S is a 'large' subset of $\mathbb{Z}_p$ such that the L^1 norm of the discrete Fourier transform of its indicator function is 'small'.

It is well-known that sets having a lot of additive structure usually also have the property that the L_1 of the Fourier transforms of their indicator functions is small. For example, if S is an arithmetic progression of length $p(\log p)^{-c}$, then $||\hat{S}||_1 = p(\log p)^{O(1)}$.

My question is: Under the restrictions (1) and (2) must we have that S has a three-term arithmetic progression if p is sufficiently large?

Three-term arithmetic progressions are among the most basic structures that we would expect sets with some additive structure to posess. So, the question is asking whether S has one type of additive structure (small L_1 norm of $\hat{S}$) implies that it has another (presence of 3APs).

5. (posed by Ben Green) Suppose that S is a subset of $\mathbb{Z}_p$, $|S| = (p-1)/2$, and S has the minimal number of three-term arithmetic progressions modulo p, among all subsets with $(p-1)/2$ elements. How many three-term progressions must S have? It is known that this number is $\sim cp^2$. The problem is to determine c.

Remark 1. I have some structural results on this problem, but at present do not know how to find this value of c.

Remark 2. One can ask the same question for $|S| \sim \theta p$ for arbitrary $0 < \theta \leq 1$, instead of $\theta = 1/2$ as we have asked above.

6. (related to a result of D. Coppersmith) Suppose that $f(x) \in \mathbb{Z}[x]$ is a monic, degree d polynomial. Show that for every $0 \leq \theta < 1$ there exists a constant $C = C(d, \theta)$ such that for every square-free integer N,

$$|\{r \ : \ |r| \leq N^{\theta}, \ f(r) \equiv 0 \pmod{N}\}| \ < \ C. \tag{3}$$

Remark 1. What makes this problem non-trivial is the fact that the C does not depend on the number of prime factors of N. Also note that (3) would be false if N is allowed to have repeated prime factors.

Remark 2. There is a simple proof that the claim is true for $\theta < 1/d$. Proving it for all $\theta \geq 1/d$ appears profoundly difficult.

Remark 3. Here is a naive heuristic: It can be shown that N has at most $O(\log N/\log\log N)$ prime factors; and so, since f has degree d, it can have at most $d^{O(\log N/\log\log N)} = N^{o(1)}$ roots mod N. The number of these roots we expect to have absolute value at most N^{θ} is therefore $N^{o(1)}/(2N^{\theta}) = o(1)$. Thus, we expect there to be no roots at all of this size. Obviously this heurisitc does not reflect the reality for some polynomials, such as $f(x) = x(x-1)\cdots(x-d+1)$, which has d roots of size at most N^{θ} for N sufficiently

large. Nonetheless, perhaps the heuristic is close to being true, and this is just what we are conjecturing in (3).

7. (asked by J. Solymosi) It is known that for every $k \geq 1$ there exists $\epsilon = \epsilon(k) > 0$ such that if N is sufficiently large, and if $S \subseteq [1, N] \cap \mathbb{Z}$ satisfies $|S| > N^{1-\epsilon}$, then $S + S$ contains k-term arithmetic progressions.

Is the same true of any positive density subset T of $S + S$? That is, is it true that for every $0 < \theta \leq 1$ and every $k \geq 1$, there exists $\epsilon = \epsilon(\theta, k)$, such that for all N sufficiently large, if $S \subset [1, N] \cap \mathbb{Z}$ satisfies $|S| > N^{1-\epsilon}$, then any set $T \subseteq S + S$ with $|T| \geq \theta |S + S|$ must have k-term progressions?

Remark. I imagine that there might be a Roth-type argument, where, for example, if $T \subseteq S + S$ has no three-term progressions (so we are looking at the case $k = 3$), then there is a subset $A \subset S$, with $|A| \geq |S|^{1-\delta}$, such that $T \cap (A + A)$ has higher density inside $A + A$ then $T \cap (S + S)$ does inside $S + S$.

8. For a group G define $r_3(G)$ to be the size of the maximal subset of G containing no three-term arithmetic progressions. In this context, a three term progression is a triple of points $a, a + d, a + 2d$. For abelian groups G of odd order our upper and lower bounds for $r_3(G)$ are appalingly far apart. However, perhaps there are non-abelian groups for which we can more easily deduce good upper and lower bounds. Here are two questions related to this

Question 1. Does there exist an infinite family of non-abelian groups G of odd order for which $r_3(G) > |G|/\log^T |G|$?

Question 2. Does there exist an infinite family of non-abelian groups G of odd order for which one can prove good upper and lower bounds for $r_3(G)$ – bounds that differ by a constant factor, or even asymptotic bounds?

Remark. The following may not make any sense, but I thought I would mention it anyway: Perhaps Behrend's construction can be made to answer question 1, if it is suitably re-interpreted. Say S is a subset of the box $\{0, 1, 2\}^n$ without three-term progressions. As it stands, the natural addition on this box is not closed. One way to make it closed is to change the topology, and work with a discrete torus $(\mathbb{Z}_3)^n$. But then, S suddenly has lots of three-term progressions. However, perhaps with a different topology, one

can preserve the fact that S has no three-term progressions, while making addition on the set closed, though non-commutative.

9. This problem is motivated by a hope to give alternate proofs of sum-product inequalities over the integers. Recall that sum-product inequalties state that if A is a set of integers, then either $A+A$ or AA has $C|A|^{1+\epsilon}$ elements. It was conjectured by Erdos that this holds for any $\epsilon < 1$, though the closer to 1 we take ϵ, the larger C would need to be. My hope is to move the set of integers A to a short interval via Freiman isomorphisms, where the machinery of analytic number theory comes into play, in the hopes of proving good sum-product bounds (I should emphasize the word *hopes*, because I don't yet know how to prove sum-product inequalities, even given that A lies in a short interval). Enough motivation – here are my questions:

Question 1. Given a set of integers A, does there exist a map $\varphi : \mathbb{Z} \to [-N, N]$, where $N < \exp(|A|^c)$ (for some $c > 0$) such that for $a, b, c, d \in A$ we have

$$\varphi(a) + \varphi(b) \;=\; \varphi(c) + \varphi(d) \quad \Longleftrightarrow \quad a + b = c + d; \tag{4}$$

and

$$\varphi(a)\varphi(b) \;=\; \varphi(c)\varphi(d) \quad \Longleftrightarrow \quad ab = cd\ ? \tag{5}$$

Let us call such a map φ a "double Freiman isomorphism". Note that if A is a geometric progression then we would need that N is at least $\exp(|A|^d)$ for some $d \geq 1$.

Question 2. Perhaps it is too difficult to prove that such a map φ exists for all of A, though perhaps it holds for a largish subset of A. Let me be a bit vague on this final question: For every set of integers A, does there exist a largish subset $A' \subseteq A$, and a map $\varphi : A' \to [-N, N]$, with $N < \exp(|A'|^c)$, such that both (4) and (5) both hold ? Is it possible to get a result with a much smaller value of N?

10. There is a beautiful and useful result due to Mei Chu Chang on the structure of large Fourier coefficients, now known as Chang's structure theorem. Before I get to my question, let me state it:

Theorem. Let $\rho, \alpha \in [0, 1]$, let $A \subseteq \mathbb{Z}_N$ be a set of size αN and let $R \subseteq \mathbb{Z}_N$ be the set of all r for which $|\hat{A}(r)| \geq \rho|A|$. Then, R is contained in a cube of dimension at most $2\rho^{-2}\log(1/\alpha)$. (In the sense that there is a set of

$n \sim 2\rho^{-2}\log(1/\alpha)$ numbers $x_1, ..., x_n \in \mathbb{Z}_N$, such that if x a place where $|\hat{A}(x)| \geq \rho|A|$, then $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$, where the λ_i are all $0, 1$ or -1.)

It would be interesting, and very useful, if there were finer structure theorems. Of particular interest to me would be the following:

Question. Suppose that $A \subseteq \mathbb{Z}_N$, $|A| = \alpha N$, $|\hat{A}(r)| \leq c_0\alpha^2 N$ for all $r \in \mathbb{Z}_N$, and that there are at least $c_1\alpha^{-3}$ values of x for which $|\hat{A}(x)| \geq c_2\alpha^2 N$. What sort of structure must these places x have? For example, are they unions of many long arithmetic progressions? (If so, can you show that A is very dense on the complement of the union of several small Bohr neighborhoods?)

11. (due to various people, implicit in a problem of Bourgain and Tao) Given a finite abelian group G, let $r_3(G)$ denote the size of the largest subset of G having no three-term arithmetic progressions. In this context, a three-term progression is a solution to $x + y = 2z$, where x, y, z are distinct elements of G. A consequence of a theorem of Roth and Meshulam is that if

$$G \;=\; \mathbb{Z}_3 \times \mathbb{Z}_3 \times \cdots \times \mathbb{Z}_3, \;\; (n \text{ copies of } \mathbb{Z}_3 \text{ in all})$$

then

$$r_3(G) \;\ll\; \frac{3^n}{n}.$$

The problem is to show that

$$r_3(G) \;=\; o\left(\frac{3^n}{n}\right).$$