An easy combinatorial proof of an intermediate value property for sumsets

Ernie Croot

October 8, 2009

1 Introduction

Here I prove the following theorem by a simple combinatorial argument, instead of by the usual Fourier methods, related to a "continuity" property of sumsets.

Theorem. Suppose that A is a subset of $\mathbb{F}_p$ of density α, and that A fails to have the "intermediate value property", by which we mean that if $x \in A+A$, then x has at least βp representations as $x = a + b$, with $a, b \in A$. Then, $A + A$ must be almost translation-invariant (which we would expect, since if $(A * A)(x)$ is large everywhere, then $A + A$ is translation invariant), by which I mean that there exists t such that

$$|(A + A + t)\Delta(A + A)| \ < \ p^{1-\varepsilon}, \ \varepsilon = \varepsilon(\alpha, \beta).$$

(Here, $U\Delta V$ means symmetric difference.) Note the strength of the conclusion – we get a p^{ε} savings over the trivial upper bound.

The fact that I got such a strong upper bound is encouraging, because even Fourier methods cannot prove such a claim (without substantial extra effort).

Also, as you know, generalizing Fourier methods to handle multiple linear forms involves higher Gowers norms, nilsequences, and so on. But often simple probabilistic arguments generalize much more easily. So it may be that the argument can be used to drastically simplify certain proofs about multiple linear forms.

2 Proof of the Theorem.

Given A satisfying the hypotheses of the theorem, it is easy to see that with high probability if we choose a random subset $B \subset A$ of size about

$$z \ := \ (\log p)/\log(1/\alpha),$$

then $B + A$ is essentially the same set as $A + A$, in the sense that

$$|(B + A)\Delta(A + A)| \ < \ p^{1-\varepsilon}, \ \varepsilon = \varepsilon(\alpha, \beta). \qquad (1)$$

To see this, note that if x satisfies $(A * A)(x) \neq 0$, then we know that there are at least βp pairs (a, b) such that $a + b = x$, under the "discontinuity assumption". Note that β/α fraction of all the elements of A appear here as a first coordinate. Now, the probability that among the z randomly-selected elements making up B we have that none appear as one of these first coordinates a is something like

$$(1 - \beta/\alpha)^z \ = \ p^{\log(1-\beta/\alpha)/\log(1/\alpha)} \ = \ p^{-\varepsilon_0}, \ \varepsilon_0 = \varepsilon_0(\alpha, \beta).$$

Since "most" B we choose have this property, it means that there are roughly

$$\binom{\alpha p}{z} \ \sim \ \alpha^z \binom{p}{z} \ > \ \frac{1}{p}\binom{p}{z}$$

subsets B have the property (1).

But, now, the subsets of $\mathbb{F}_p$ of size z can be broken down into conjugacy classes induced by taking translations; that is, the $\binom{p}{z}$ subsets of $\mathbb{F}_p$ of size z break down into $(1/p)\binom{p}{z}$ classes, where the subsets of each class are all translates of one another.

Since the number of sets B we produced above, exceeds the number of translation classes, two of our such sets B must lie in the same class. This means, then, that there are sets B and $B + t$ such that

$$|(B + A)\Delta(A + A)| \ = \ \text{"small"}, \text{ and } |(B + A + t)\Delta(A + A)| \ = \ \text{"small"}.$$

It is easy to see now that $A + A$ is close to being translation-invariant.

■