Electronic Colloquium on Computational Complexity, Report No. 10 (2010)

Equivalence of polynomial conjectures in additive combinatorics

Shachar Lovett *

January 16, 2010

Abstract

We study two conjectures in additive combinatorics. The first is the polynomial Freiman-Ruzsa conjecture, which relates to the structure of sets with small doubling. The second is the inverse Gowers conjecture for U^3, which relates to functions which locally look like quadratics. In both cases a weak form, with exponential decay of parameters is known, and a strong form with only a polynomial loss of parameters is conjectured. Our main result is that the two conjectures are in fact equivalent.

1 Introduction

Additive combinatorics studies subsets of abelian groups, with the main examples are subsets of the integers and of vector spaces over finite fields. The main problems entail connecting various properties related to the additive structure of the space, to structural properties of the subsets. In a way, additive combinatorics can be viewed as a robust analog of basic linear algebra.

We study in this paper two conjectures relating to objects defined over vector spaces $\mathbb{F}^n$. The first is the polynomial Freiman-Ruzsa conjecture, which relates to subsets $S \subset \mathbb{F}^n$ which are approximately vector spaces. The second is the polynomial inverse Gowers conjecture for the U^3 norm, which relates to functions $f : \mathbb{F}^n \to \mathbb{F}$ which

*Research supported by the Israel Science Foundation (grant 1300/05).

ISSN 1433-8092

are approximately quadratic. Both conjectures aim to give structural properties for these objects.

Our main result is that the two conjectures are equivalent. We focus in the paper on the case of $\mathbb{F} = \mathbb{F}_2$; our results extend easily to any constant finite field $\mathbb{F}_p$.

1.1 Approximate vector spaces

Let $S \subset \mathbb{F}_2^n$. The set S is said to have *doubling* K if $|S+S| \leq K|S|$, where $S+S = \{x+y : x, y \in S\}$. It is clear that S has doubling 1 iff it is an affine space. Thus, a set with small doubling can be viewed as an approximate vector space. Can we infer some structure such sets must have? The following theorem of Ruzsa [Ruz99] claims that any such set is contained in a vector space which is not much larger.

Theorem 1 (Theorem 1 in [Ruz99])**.** *Let $S \subset \mathbb{F}_2^n$ such that $|S+S| \leq K|S|$. Then $|Span(S)| \leq K^2 2^{K^4}|S|$.*

The work of Ruzsa is an analog of a similar result of Freiman [Fre73] for subsets of the integers with small doubling. Theorem 1 was improved in a series of works (Green and Ruzsa [GR06], Sanders [San08] and Green and Tao [GT09c]) to an almost optimal bound.

Theorem 2 (Theorem 1.3 in [GT09c])**.** *Let $S \subset \mathbb{F}_2^n$ such that $|S+S| \leq K|S|$. Then $|Span(S)| \leq 2^{(2+o(1))K}|S|$.*

The bound is tight up to the $o(1)$ term as can be seen by the following example: let $S = \{v_1, \ldots, v_r\}$ where $v_1, \ldots, v_r \in \mathbb{F}_2^n$ are linearly independent. We have $|S+S| \approx \frac{r}{2}|S|$ and $|Span(S)| = 2^r$. We could also have $S = V + \{v_1, \ldots, v_r\}$ where V is a vector space and $v_1, \ldots, v_r \in V^\perp$.

This example shows that the exponential loss of parameters in Theorem 2 is inevitable. It would be beneficial, however, to have some structure theorem for sets with small doubling which have only a polynomial loss of parameters. In general, theorems which involve only a polynomial loss of parameters are useful as they can be applied iteratively several times, resulting again with only a polynomial loss of parameters. The following strengthening of Theorems 1 and 2, known as the *Polynomial Freiman-Ruzsa conjecture* was suggested in several works.

Conjecture 3 (Polynomial Freiman-Ruzsa conjecture)**.** *Let $S \subset \mathbb{F}_2^n$ such that $|S+S| \leq K|S|$. Then there is a subset $S' \subset S$, $|S'| \geq K^{-O(1)}|S|$, such that $|Span(S')| \leq K^{O(1)}|S|$.*

Conjecture 3 was proved by Green and Tao for the special case when S is a downset [GT09c], as well as in the general case with an exponential loss of parameters which is better than that given by Theorem 2 [GT09a].

The Polynomial Freiman-Ruzsa conjecture can be equivalently restated is several forms. We give below two such forms which relate to approximate homomorphisms. For proofs of the equivalence as well as several other equivalent formulations see [Gre05].

The first formulation relates to testing if a function $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is close to a linear map. A natural way to do so is to sample $x, y \in \mathbb{F}_2^n$ and verify that $f(x+y) = f(x)+f(y)$. The following conjecture states that if this event occurs with polynomial ϵ over the choice of x, y, then f is $poly(\epsilon)$ close to a linear map.

Conjecture 4 (Approximate homomorphism testing)**.** *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be such that $\Pr_{x,y}[f(x+y) = f(x)+f(y)] \geq \epsilon$. Then there is a linear map $\ell : \mathbb{F}_2^n \to \mathbb{F}_2^m$ such that $\Pr_x[f(x) = \ell(x)] \geq \epsilon^{O(1)}$.*

The second formulations relates to structured approximate homomorphisms. For a function $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ define its *difference set* $\Delta f = \{f(x+y) - f(x) - f(y) : x, y \in \mathbb{F}_2^n\}$. The following conjectures claim that if Δf is small then f can be expressed as the sum of a linear function and an error function, where the error function obtains at most $poly(|\Delta f|)$ possible values.

Conjecture 5 (Structured approximate homomorphism)**.** *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ and assume that $|\Delta f| \leq K$. Then there is a linear map $\ell : \mathbb{F}_2^n \to \mathbb{F}_2^m$ such that $f(x) = \ell(x) + e(x)$ where $|\{e(x) : x \in \mathbb{F}_2^n\}| \leq K^{O(1)}$.*

Note that a-priory, it seems that the assumption of Conjecture 5 is much stronger than that of Conjecture 4; nevertheless, the conjectures are equivalent. We also note that analogs of Conjectures 4 and 5 with exponential loss of parameters follow from Theorem 2.

There is another natural definition for an approximate vector space; $S \subset \mathbb{F}_2$ is an approximate vector space if for many pairs $x, y \in S$ we have $x + y \in S$. The following theorem due to Balog, Szemerèdi and Gowers [BS94, Gow98] shows this property is polynomially related to the case of small doubling.

Theorem 6 (Balog-Szemerèdi-Gowers). *Let $S \subset \mathbb{F}_2^n$. If $\Pr_{x,y\in S}[x+y \in S] \geq \epsilon$ then there is a subset $S' \subset S$, $|S'| \geq \epsilon^{O(1)}|S|$ such that S' has small doubling, $|S'+S'| \leq \epsilon^{-O(1)}|S'|$.*

1.2 Approximate polynomials

Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be a function. Define the derivative of f in direction $y \in \mathbb{F}_2^n$ as $f_y(x) = f(x+y)+f(x)$ [1]. If f is a degree d polynomial then f_y is a polynomial of degree at most $d-1$. Define iterated derivatives as

$$f_{y_1,\ldots,y_d}(x) = (f_{y_1,\ldots,y_{d-1}})_{y_d}(x) = \sum_{I\subseteq[d]} f(x+\sum_{i\in I} y_i)$$

and observe that f is a polynomial of degree at most $d-1$ iff $f_{y_1,\ldots,y_d}(x) \equiv 0$ for all $y_1,\ldots,y_d \in \mathbb{F}_2^n$. On the other hand, if f is a random boolean function then $f_{y_1,\ldots,y_d}(x)$ is distributed close to uniform over $\mathbb{F}_2$. Thus, a plausible definition for an approximate polynomial is a function f for which $\Pr_{x,y_1,\ldots,y_d}[f_{y_1,\ldots,y_d}(x) = 0] \geq 1/2+\epsilon$. This is captured by the *Gowers norm*, defined originally by Gowers [Gow98] in his seminal work on a new proof for Szemerédi's theorem. The Gowers norm is defined over complex functions $F : \mathbb{F}_2^n \to \mathbb{C}$ (think of $F(x) = (-1)^{f(x)}$). Define the derivative of F in direction $y \in \mathbb{F}_2^n$ as $F_y(x) = F(x+y)\overline{F(x)}$, and iterated derivatives analogously.

Definition 1 (Gowers norm). Let $F : \mathbb{F}_2^n \to \mathbb{C}$. The d-th Gowers norm of F is defined as

$$\|F\|_{U^d} = \left(\mathbb{E}_{x,y_1,\ldots,y_d\in\mathbb{F}_2^n}[F_{y_1,\ldots,y_d}(x)]\right)^{1/2^d}.$$

The following summarize some simple facts regarding the Gowers norm.

Fact 7 (Simple facts regarding the Gowers norm). *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$.*

1. *$\|\cdot\|_{U^d}$ is a norm of complex functions (for $d = 1$ it is a semi-norm).*
2. *$0 \leq \|(-1)^f\|_{U^d} \leq 1$.*
3. *$\|(-1)^f\|_{U^d} = 1$ iff f is a polynomial of degree at most $d-1$.*
4. *If f is a random boolean function then $\|(-1)^f\|_{U^d} \approx 0$.*
5. *Assume there is a polynomial $p(x)$ of degree at most $d-1$ such that $\Pr_x[f(x) = p(x)] \geq \frac{1+\epsilon}{2}$. Then $\|(-1)^f\|_{U^d} \geq \epsilon$.*

[1] Over odd fields define $f_y(x) = f(x+y) - f(x)$.

The hard direction is proving structure theorems for functions with noticeable Gowers norm. This is known as the *inverse Gowers conjecture*. Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be a function for which $\|(-1)^f\|_{U^d} \geq \epsilon$. The conjecture speculates there exists a polynomial $p(x)$ of degree at most $d-1$ for which $\Pr_x[f(x) = p(x)] \geq \frac{1}{2} + \epsilon'$, where ϵ' may depend on ϵ and d, but crucially it does not depend on the number of variables n. Much is known today about the inverse Gowers conjecture. We summarize below the current state of affairs.

Fact 8 (Inverse Gowers conjecture)**.** *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ such that $\|f\|_{U^d} \geq \epsilon$ for $d \leq 3$. Then there exists a polynomial $p(x)$ of degree at most $d-1$ such that $\Pr[f(x) = p(x)] \geq \frac{1}{2} + \epsilon'$, where*

- *$d = 1$: It is easy to verify that $\|(-1)^f\|_{U^1} = |\mathbb{E}[(-1)^f]|$. This gives $\epsilon' = \epsilon/2$.*
- *$d = 2$: It can be shown by simple Fourier analysis that the U^2 norm of $(-1)^f$ is equal to the L_4 norm of the Fourier coefficients of f, that is $\|f\|_{U^2} = \|\hat{f}\|_4$. This gives $\epsilon' \geq \Omega(\epsilon^2)$.*
- *$d = 3$: This case is more involved. Results of Green and Tao [GT08] and Samorodnitsky [Sam07] give that in this case $\epsilon' \geq exp(-1/\epsilon)$.*

When $d \geq 4$ things become trickier. It is no longer true that if $\|f\|_{U^d} \geq \epsilon$ there must exist a polynomial $p : \mathbb{F}_2^n \to \mathbb{F}_2$ of degree at most $d-1$ which approximates f with probability noticeably larger than $1/2$ [LMS08, GT07]. Nevertheless, a refined inverse conjecture holds: there exists a "non-classical" polynomial $F : \mathbb{F}_2^n \to \mathbb{C}$ which approximates f. The notion of approximation is the natural generalization of our previous definition, $\mathbb{E}_x[(-1)^{f(x)}F(x)] \geq \epsilon'$. A "non-classical" polynomial is a function $F : \mathbb{F}_2^n \to \mathbb{C}$ for which $F_{y_1,\ldots,y_d}(x) \equiv 1$ (for example, $F(x) = i^{x_1+\ldots+x_n}$ is a "non-classical" quadratic). This was proved by Bergelson, Tao and Ziegler [BTZ09, TZ09] using Ergodic theory. A major caveat of this approach is that currently no explicit bound on ϵ' in terms of ϵ and d is known. All that is known is that ϵ' is some constant depending only on ϵ, d.

The only case where there is an explicit relation between ϵ' and ϵ which is not polynomial is the case of $\|\cdot\|_{U^3}$, where it is believed to be suboptimal. The following polynomial relation is conjectured.

Conjecture 9 (Polynomial Inverse Gowers conjecture for U^3)**.** *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$. If $\|f\|_{U^3} \geq \epsilon$ then there exists a quadratic polynomial $p(x)$ such that $\Pr[f(x) = p(x)] \geq \frac{1}{2} + \epsilon^{O(1)}$.*

Our main result is that the Polynomial Freiman-Ruzsa conjecture and the Polynomial Inverse Gowers conjecture for U^3 are equivalent.

Theorem 10. *Conjecture 3 and Conjecture 9 are equivalent.*

One direction is simple. The only place in the proof of the inverse conjecture for U^3 where a super-polynomial loss occurs is in the use of the Ruzsa theorem. Assuming the polynomial Freiman-Ruzsa conjecture this loss can be avoided. We sketch the required change in the proof in Section 3.

The main innovation of this work is a proof of the polynomial Freiman-Ruzsa conjecture assuming a polynomial inverse theorem for U^3. We prove this in Section 2.

We note that this result was also independently discovered by Green and Tao [GT09b].

2 Deducing the polynomial Freiman-Ruzsa conjecture, assuming a polynomial inverse conjecture for U^3

We will prove Conjecture 5, which is equivalent to the polynomial Freiman-Ruzsa conjecture. Let $f : \mathbb{F}_2^n \to \mathbb{F}_n^m$ be a function and let $\Delta f = \{f(x+y) - f(x) - f(y) : x, y \in \mathbb{F}_2^n\}$. We assume $|\Delta f| \leq K$, and wish to prove that there exists a linear map $\ell : \mathbb{F}_2^n \to \mathbb{F}_2^m$ such that $|\{f(x) - \ell(x) : x \in \mathbb{F}_2^n\}| \leq K^{O(1)}$.

Define a function $F : \mathbb{F}_2^{n+m} \to \mathbb{F}_2$ by $F(x, z) = \langle f(x), z\rangle$ for $x \in \mathbb{F}_2^n, z \in \mathbb{F}_2^m$, where $\langle \cdot, \cdot \rangle$ denotes inner product. The proof will proceed in the following steps.

1. Show that $\|F\|_{U^3} \geq K^{-O(1)}$.
2. By the polynomial inverse conjecture for U^3, there is a quadratic polynomial $Q(x, z)$ such that $\Pr[F(x, z) = Q(x, z)] \geq \frac{1}{2} + K^{-O(1)}$.
3. Deduce there is a linear map $\ell : \mathbb{F}_2^n \to \mathbb{F}_2^m$ such that $\Pr[f(x) = \ell(x) + c] \geq K^{-O(1)}$ for some $c \in \mathbb{F}_2^m$.
4. Conclude by showing that $|\{f(x) - \ell(x) : x \in \mathbb{F}_2^n\}| \leq K^{O(1)}$.

Lemma 11. $\|F\|_{U^3} \geq K^{-7/8}$.

Proof. We compute $\|F\|_{U^3}^8$. Let $x, y_1, y_2, y_3 \in \mathbb{F}_2^n$ and $z, w_1, w_2, w_3 \in \mathbb{F}_2^m$ be chosen uniformly. We have

$$\begin{aligned}\|F\|_{U^3}^8 &= \mathbb{E}[(-1)^{\sum_{I\subseteq[3]} F(x+\sum_{i\in I} y_i, z+\sum_{i\in I} w_i)}] \\ &= \mathbb{E}[(-1)^{\sum_{I\subseteq[3]} \langle f(x+\sum_{i\in I} y_i), z+\sum_{i\in I} w_i\rangle}] \\ &= \mathbb{E}[(-1)^{\langle z, A_0\rangle + \langle w_1, A_1\rangle + \langle w_2, A_2\rangle + \langle w_3, A_3\rangle}]\end{aligned}$$

where

$$\begin{aligned}A_0 &= \sum_{I\subseteq\{1,2,3\}} f(x + \sum_{i\in I} y_i) \\ A_1 &= \sum_{I\subseteq\{2,3\}} f(x + y_1 + \sum_{i\in I} y_i) \\ A_2 &= \sum_{I\subseteq\{1,3\}} f(x + y_2 + \sum_{i\in I} y_i) \\ A_3 &= \sum_{I\subseteq\{1,2\}} f(x + y_3 + \sum_{i\in I} y_i)\end{aligned}$$

Hence we have

$$\begin{aligned}\|F\|_{U^3}^8 &= \mathbb{E}[(-1)^{\langle z, A_0\rangle + \langle w_1, A_1\rangle + \langle w_2, A_2\rangle + \langle w_3, A_3\rangle}] \\ &= Pr_{x,y_1,y_2,y_3\in\mathbb{F}_2^n}[A_0 = 0, A_1 = 0, A_2 = 0, A_3 = 0].\end{aligned}$$

The proof will follow from the following general claim.

Claim 12. *For any $k \geq 1$ there exist values $c_1, \ldots, c_k \in \mathbb{F}_2^m$ such that the set*

$$S_k = \{(x, y_1, \ldots, y_k) \in (\mathbb{F}_2^n)^{k+1} : \forall I \subseteq [k],\ f(x+\sum_{i\in I} y_i) = f(x)+\sum_{i\in I} f(y_i)+\sum_{i\in I} c_i\}$$

has relative size at least $\frac{|S_k|}{2^{n(k+1)}} \geq (1/K)^{2^k-1}$.

Before proving the claim we show how it can be applied to conclude the proof of Lemma 11. Let $c_1, c_2, c_3 \in \mathbb{F}_2^m$ be values and let

$$S_3 = \{(x, y_1, y_2, y_3) \in (\mathbb{F}_2^n)^4 : \forall I \subseteq [3],\ f(x+\sum_{i\in I} y_i) = f(x)+\sum_{i\in I} f(y_i)+\sum_{i\in I} c_i\}.$$

such that its relative size is $\frac{|S_3|}{2^{4n}} \geq (1/K)^7$. Notice that if $(x, y_1, y_2, y_3) \in S_3$ then $A_0 = A_1 = A_2 = A_3 = 0$ as each variable appears an even number of times in each of A_0, A_1, A_2, A_3. Thus we conclude that

$$\|F\|_{U^3}^8 = \Pr[A_0 = A_1 = A_2 = A_3 = 0] \geq (1/K)^7.$$

□

We now turn to prove the claim.

Proof of Claim 12. The proof will be by induction on k. For $k = 1$ this follows since $f(x + y) - f(x) - f(y) \in \Delta$ for all $x, y \in \mathbb{F}_2^n$ and $|\Delta| \leq K$. Assume the claim holds for k, and we will prove it for $k+1$. Let $c_1, \ldots, c_k \in \mathbb{F}_2^m$ be such that

$$S_k = \{(x, y_1, \ldots, y_k) \in (\mathbb{F}_2^n)^{k+1} : \forall I \subseteq [k],\ f(x+\sum_{i\in I} y_i) = f(x)+\sum_{i\in I} f(y_i)+\sum_{i\in I} c_i\}$$

has relative size at least $\frac{|S_k|}{2^{n(k+1)}} \geq (1/K)^{2^k-1}$. Consider the set

$$S' = \{(x, y_1, \ldots, y_k, y_{k+1}) \in (\mathbb{F}_2^n)^{k+2} : (x, y_1, \ldots, y_k) \in S_k \text{ and } (x+y_{k+1}, y_1, \ldots, y_k) \in S_k\}$$

By the Cauchy-Schwartz inequality its relative size is lower bounded by $(1/K)^{2^k-2}$, as

$$\begin{aligned}
\frac{|S'|}{2^{n(k+2)}} &= \mathbb{E}_{x,y_1,\ldots,y_k,y_{k+1}}[\mathbf{1}_{(x,y_1,\ldots,y_k)\in S_k} \cdot \mathbf{1}_{(x+y_{k+1},y_1,\ldots,y_k)\in S_k}] \\
&= \mathbb{E}_{x,z,y_1,\ldots,y_k}[\mathbf{1}_{(x,y_1,\ldots,y_k)\in S_k} \cdot \mathbf{1}_{(z,y_1,\ldots,y_k)\in S_k}] \\
&= \mathbb{E}_{y_1,\ldots,y_k}[\mathbb{E}_x[\mathbf{1}_{(x,y_1,\ldots,y_k)\in S_k}]^2] \\
&\geq (\mathbb{E}_{x,y_1,\ldots,y_k}[\mathbf{1}_{(x,y_1,\ldots,y_k)\in S_k}])^2 \\
&= \left(\frac{|S_k|}{2^{n(k+1)}}\right)^2 \geq (1/K)^{2^k-2}.
\end{aligned}$$

Fix $c_{k+1} \in \Delta$ such that $\Pr_{(x,y_1,\ldots,y_{k+1})\in S'}[f(x+y_{k+1})-f(x)-f(y_{k+1}) = c_{k+1}] \geq 1/K$. The required set S_{k+1} is chosen to be

$$\begin{aligned}
S_{k+1} = \{(x, y_1, \ldots, y_{k+1}) \in (\mathbb{F}_2^n)^{k+2} :&(x, y_1, \ldots, y_{k+1}) \in S' \text{ and} \\
&f(x + y_{k+1}) - f(x) - f(y_{k+1}) = c_{k+1}\}.
\end{aligned}$$

Observe that the relative size of S_{k+1} is as required,

$$\frac{|S_{k+1}|}{2^{n(k+2)}} \geq \frac{1}{K} \cdot \frac{|S'|}{2^{n(k+2)}} \geq (1/K)^{2^{k+1}-1}.$$

Let $(x, y_1, \ldots, y_{k+1}) \in S_{k+1}$. We need to show that $f(x + \sum_{i\in I} y_i) = f(x) + \sum_{i\in I} y_i + \sum_{i\in I} c_i$. If $(k+1) \notin I$ this follow by induction from the assumption on S_k. Otherwise, using the fact that $(x + y_{k+1}, y_1, \ldots, y_k) \in S_k$ we get that $f(x + y_{k+1} + \sum_{i\in I} y_i) = f(x + y_{k+1}) + \sum_{i\in I\setminus\{k+1\}} f(y_i) + \sum_{i\in I\setminus\{k+1\}} c_i$, and since $f(x + y_{k+1}) = f(x) + f(y_{k+1}) + c_{k+1}$ we conclude the proof. □

Using Lemma 11 and the polynomial inverse conjecture for the Gowers U^3 norm, we get there is a quadratic polynomial such that $\Pr[\langle f(x), z\rangle = Q(x, z)] \geq \frac{1}{2} + K^{-O(1)}$.

Lemma 13. *Let $Q(x, z)$ be a quadratic polynomial such that $\Pr[\langle f(x), z\rangle = Q(x, z)] \geq \frac{1}{2} + \epsilon$. Then there exist a linear mapping $\ell : \mathbb{F}_2^n \to \mathbb{F}_2^m$ and a constant $c \in \mathbb{F}_2^m$ such that*

$$\Pr[f(x) = \ell(x) + c] \geq \epsilon^4/K.$$

Proof. Let $Q(x, z) = x^T Az + Q_1(x) + Q_2(z)$, where A is an $n \times m$ matrix and Q_1, Q_2 are quadratics just in the x and z variables. We have

$$|\mathbb{E}[(-1)^{\langle f(x),z\rangle + Q(x,z)}]| = 2\Pr[\langle f(x), z\rangle = Q(x, z)] - 1 \geq \epsilon.$$

We will use the following simple claim, which follows by applying the Cauchy-Schwartz inequality twice: for any function $G(x, z)$ the following holds

$$G(x, z)^4 \leq \mathbb{E}_{x',x'',z',z''} G(x', z')G(x'', z')G(x', z'')G(x'', z'').$$

Let $G(x, z) = (-1)^{\langle f(x),z\rangle + Q(x,z)}$. Note that

$$G(x', z')G(x'', z')G(x', z'')G(x'', z'') = (-1)^{\langle f(x')+f(x''),z'+z''\rangle + (x'+x'')^T A(z'+z'')}$$

hence we deduce, by setting $w = z' + z''$ that

$$\mathbb{E}_{x',x'',w}[(-1)^{\langle f(x')+f(x''),w\rangle + (x'+x'')^T Aw}] \geq \epsilon^4.$$

Let $\ell : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be the linear mapping defined by A, that is $\ell(x) = x^T A$. Note that we have

$$\mathbb{E}_{x',x'',w}[(-1)^{\langle f(x')+f(x'')+\ell(x'+x''),w\rangle}] \geq \epsilon^4,$$

hence

$$\Pr_{x',x''}[f(x') + f(x'') = \ell(x' + x'')] \geq \epsilon^4.$$

We are nearly done. To complete the proof we use the fact that $f(x'+x'')-f(x')-f(x'') \in \Delta f$ to deduce that there is some $c \in \Delta f$ such that

$$\Pr_{x',x''}[f(x'+x'') = f(x')+f(x'')+c | f(x')+f(x'') = \ell(x'+x'')] \geq 1/K,$$

hence we got the required result,

$$\Pr_x[f(x) = \ell(x)+c] \geq \epsilon^4/K.$$

□

We finish the proof by a standard covering argument.

Lemma 14. *Assume there is a linear map $\ell : \mathbb{F}_2^n \to \mathbb{F}_2^m$ and $c \in \mathbb{F}_2^n$ such that $\Pr[f(x) = \ell(x)+c] \geq \epsilon$. Then $|\{f(x)-\ell(x) : x \in \mathbb{F}_2^n\}| \leq K^2/\epsilon$.*

Proof. Let $T = \{x \in \mathbb{F}_2^n : f(x) = \ell(x)+c\}$. Let $B \subset \mathbb{F}_2^n$ be maximal such that for any distinct $b', b'' \in B$ the sets $T+b'$ and $T+b''$ are disjoint. Clearly $|B| \leq 1/\epsilon$. Let $x \in \mathbb{F}_2^n$ be arbitrary. By the maximality of B we have that $(T+x) \cap (T+B) \neq \emptyset$, hence we get that $x \in T+T+B$. Let $x = t'+t''+b$ for $t', t'' \in T$ and $b \in B$. We have $f(x) = f(t')+f(t'')+f(b)+r$ for $r \in \Delta f + \Delta f$. Thus we have

$$\begin{aligned} f(x)-\ell(x) &= f(t'+t''+b) - \ell(t'+t''+b) \\ &= (f(t')-\ell(t')) + (f(t'')-\ell(t'')) + (f(b)-b)) + r \\ &= c+c+(f(b)-b)+r. \end{aligned}$$

Let $B' = \{f(b)-b : b \in B\}$. We got that $\{f(x)-\ell(x) : x \in \mathbb{F}_2^n\} \subset \Delta f + \Delta f + B'$, and an obvious upper bound is $|\Delta f + \Delta f + B'| \leq |\Delta f|^2 |B'| \leq K^2/\epsilon$. □

3 Deducing a polynomial inverse theorem for U^3, assuming the polynomial Freiman-Ruzsa conjecture

We follow the proof of the inverse theorem for U^3 of Samorodnitsky [Sam07]. Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be a function such $\|(-1)^f\|_{U^3} \geq \epsilon$. The proof proceeds as follows.

1. For $\epsilon^{O(1)}$ fraction of $y \in \mathbb{F}_2^n$ we have that $\|(-1)^{f_y}\|_{U^2} \geq \epsilon^{O(1)}$ (Corollary 6.2).
2. Using the inverse theorem for $\|\cdot\|_{U^2}$, there exist linear maps $\ell^{(y)} : \mathbb{F}_2^n \to \mathbb{F}_2$ such that $\Pr_{x,y}[f_y(x) = \ell^{(y)}(x)] \geq \epsilon^{O(1)}$.
3. Crucially, one can choose the linear maps to behave linearly in y, $\Pr_{y,z}[\ell^{(y+z)} \equiv \ell^{(y)} + \ell^{(z)}] \geq \epsilon^{O(1)}$ (Lemma 6.7).
4. Define $S = \{(y, \ell^{(y)}) : y \in \mathbb{F}_2^n\} \subset \mathbb{F}_2^{2n}$. We have $\Pr_{a,b \in S}[a + b \in S] \geq \epsilon^{O(1)}$.
5. By the Balog-Szemerèdi-Gowers theorem there exists $S' \subset S$ such that $|S'| \geq \epsilon^{O(1)}|S|$ and $|S' + S'| \leq \epsilon^{-O(1)}$.
6. Originally, Ruzsa's theorem was used to deduce that $Span(S') \leq exp(1/\epsilon^4)|S'|$. We replace it by the polynomial Freiman-Ruzsa conjecture to deduce there is $S'' \subset S'$ such that $|S''| \geq \epsilon^{O(1)}|S'|$ and $|Span(S'')| \leq \epsilon^{-O(1)}|S''|$.
7. S'' can be used to construct a global linear map $L : \mathbb{F}_2^n \to \mathbb{F}_2^n$ such that $\Pr_y[\ell^{(y)} = L(y)] \geq \epsilon^{O(1)}$ (Discussion following Theorem 6.9).
8. The remainder of the proof shows how to integrate L to find a quadratic $q(x)$ such that $\Pr[q(x) = f(x)] \geq \frac{1}{2} + \epsilon^{O(1)}$ (Lemmas 6.10 and 6.11).

Acknowledgement. I would like to thank Amir Shpilka and Partha Mukhopadhyay for invaluable discussions. I would also like to thank Alex Samorodnitsky and Seva Lev for early discussions on this problem.

References

[BS94] Antal Balog and Endre Szemerédi. A statistical theorem of set addition. *Combinatorica*, 14(3):263–268, 1994.

[BTZ09] Vitaly Bergelson, Terence Tao, and Tamar Ziegler. An inverse theorem for the uniformity seminorms associated with the action of $\mathcal{F}^\omega$. Submitted, 2009.

[Fre73] G. A. Freiman. *Foundations of a structural theory of set addition, by G. A. Freiman.* American Mathematical Society, Providence, R.I.,, 1973.

[Gow98] W. T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. 8(3):529–551, 1998.

[GR06] Ben Green and Imre Z. Rusza. Sets with small sumset and rectification. *Bulletin of the London Mathematical Society*, 38(01):43–52, 2006.

[Gre05] Ben Green. Finite field models in additive combinatorics. *London Math. Soc. Lecture Note Ser.*, 327:1–27, 2005.

[GT07] Ben Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms. Submitted, 2007.

[GT08] Ben Green and Terence Tao. An inverse theorem for the Gowers U^3 norm. 2008.

[GT09a] Ben Green and Terence Tao. A note on the Freiman and Balog-Szemerédi-Gowers theorems in finite fields. *J. Aust. Math. Soc.*, 86(1):61–74, 2009.

[GT09b] Ben Green and Terence Tao. An equivalence between inverse sumset theorems and inverse conjectures for the U^3-norm. to appear in Math. Proc. Camb. Phil. Soc., 2009.

[GT09c] Ben Green and Terence Tao. Freiman's theorem in finite fields via extremal set theory. *Comb. Probab. Comput.*, 18(3):335–355, 2009.

[LMS08] Shachar Lovett, Roy Meshulam, and Alex Samorodnitsky. Inverse conjecture for the Gowers norm is false. In *Proceedings of the* 40^{th} *annual ACM symposium on Theory of computing (STOC '08)*, pages 547–556, New York, NY, USA, 2008. ACM.

[Ruz99] Imre Z. Ruzsa. An analog of Freiman's theorem in groups. *Structure Theory of Set-Addition*, Astérique 258:323–326, 1999.

[Sam07] Alex Samorodnitsky. Low-degree tests at large distances. In *Proceedings of the* 39^{th} *annual ACM symposium on Theory of computing (STOC '07)*, pages 506–515, New York, NY, USA, 2007. ACM.

[San08] T. Sanders. A note on freĭman's theorem in vector spaces. *Comb. Probab. Comput.*, 17(2):297–305, 2008.

[TZ09] Terence Tao and Tamar Ziegler. The inverse conjecture for the Gowers norm over finite fields via the correspondence principle. Submitted, 2009.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de