Electronic Colloquium on Computational Complexity, Report No. 29 (2012)

An exposition of Sanders quasi-polynomial Freiman-Ruzsa theorem

Shachar Lovett*
Institute for Advanced Study
slovett@math.ias.edu

April 3, 2012

Abstract

The polynomial Freiman-Ruzsa conjecture is one of the important conjectures in additive combinatorics. It asserts than one can switch between combinatorial and algebraic notions of approximate subgroups with only a polynomial loss in the underlying parameters. This conjecture has also already found several applications in theoretical computer science. Recently, Tom Sanders proved a weaker version of the conjecture, with a quasi-polynomial loss in parameters. The aim of this note is to make his proof accessible to the theoretical computer science community, and in particular to people who are less familiar with additive combinatorics.

1 Introduction

Let $A \subset \mathbb{F}_2^n$. Its sumset $A+A$ is defined as $A+A = \{a_1+a_2 | a_1, a_2 \in A\}$. It is straightforward to see that $|A+A| = |A|$ if and only if A is an affine subspace of $\mathbb{F}_2^n$. Thus, one may think of subsets A for which $|A+A| \approx |A|$ as an approximate version of affine subspaces. If $|A+A| \le K|A|$ we say that A has *doubling* K and study the structure of sets of small doubling. For the sake of simplicity of exposition, we focus in this note on subgroups of $\mathbb{F}_2^n$. We note that many of the results discussed here can be extended to vector spaces over larger fields; to general abelian groups; and sometimes even to non-abelian groups.

Ruzsa [Ruz99], following previous work of Freiman [Fre73] who studied similar problems over the integers, showed that sets of small doubling must be contained in subspaces of small dimension. This bounds were later improved in a series of works [GR06, San08, GT09, Kon08, EZ11]. In the following we denote by $\mathrm{Span}(A)$ the linear subspace spanned by A.

Theorem 1.1 (Freiman-Ruzsa Theorem in $\mathbb{F}_2^n$)**.** *Let $A \subseteq \mathbb{F}_2^n$ be a set such that $|A+A| \le K|A|$. Then $|\mathrm{Span}(A)| \le O(2^{2K}/K) \cdot |A|$.*

*Supported by NSF grant DMS-0835373.

ISSN 1433-8092

This bound is sharp, as can be seen from the following example. Let $A = \mathbb{F}_2^m \times \{e_1, \ldots, e_n\} \subset \mathbb{F}_2^{m+n}$. Then $|A| = 2^m n$, $K = |A+A|/|A| \approx n/2$ and $|\mathrm{Span}(A)| = 2^{n+m} \approx (2^{2K}/K)|A|$. This shows that the ratio between $|\mathrm{Span}(A)|$ and $|A|$ must depend exponentially on the doubling of A. However, the above example suggests that maybe a refined question, relating the ratio between the span and the size of large subsets of A, might have better dependence on the doubling of A. This is captured by the Polynomial Freiman-Ruzsa conjecture (PFR).

Conjecture 1.2 (Polynomial Freiman-Ruzsa conjecture)**.** *Let $A \subset \mathbb{F}_2^n$ be a set such that $|A+A| \le K|A|$. Then there exists a subset $A' \subset A$ of size $|A'| \ge K^{-c}|A|$ such that $|\mathrm{Span}(A')| \le K^c|A|$, where $c > 0$ is an absolute constant.*

The PFR conjecture plays a central role in additive combinatorics. The main reason is that it allows one to switch between a combinatorial notion of approximate vector space (that of having small doubling) and an algebraic notion (that of having small linear span) with only a polynomial loss in the parameters. It has many equivalent formulations, we refer the interested reader to a survey of Green [Gre05] which lists many of them. Also, Green and Tao [GT09] and independently Lovett [Lov10] showed the the PFR conjecture is equivalent to a polynomial bound for the inverse Gowers U^3-norm.

The PFR conjecture has found already several diverse applications in computer science as well:

1. Samorodnitsky [Sam07] gave an analysis of linearity testing for maps $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$. If one assumes the PFR conjecture, his result improves to only suffer a polynomial loss in the parameters.

2. Ben-Sasson and Zewi [BSZ11] used the PFR conjecture to construct two-source extractors from affine extractors.

3. Ben-Sasson, Lovett and Zewi [BSLZ11] used it to get the first sub-linear bounds on the deterministic communication complexity of functions in terms of the rank of their associated matrix.

4. Bhowmick, Dvir and Lovett [BDL12] used it to give super-polynomial lower bounds on the block size of locally decodable codes arising from matching vector families.

The aim of this note is to give a detailed exposition of the following breakthrough result of Sanders [San10], who proved a weaker version of the Freiman-Ruzsa conjecture with a quasi-polynomial loss in parameters. As noted before, his result extends to more general abelian groups, but we focus on $\mathbb{F}_2^n$ for simplicity of exposition.

Theorem 1.3 (Quasi-polynomial Freiman-Ruzsa theorem [San10])**.** *Let $A \subset \mathbb{F}_2^n$ be a set such that $|A+A| \le K|A|$. Then there exists a subset $A' \subseteq A$ of size $|A'| \ge K^{-O(\log^3 K)}|A|$ such that $|\mathrm{Span}(A')| \le K^{O(1)}|A'|$.*

In fact, Sanders proved an even stronger result. For $t \ge 1$ let $tA = \{a_1 + \ldots + a_t | a_1, \ldots, a_t \in A\}$ denote the t-sumset of A.

Theorem 1.4 (Quasi-polynomial Bogolyubov-Ruzsa theorem [San10])**.** *Let $A \subset \mathbb{F}_2^n$ be a set such that $|A+A| \le K|A|$. Then there exists a linear subspace $V \subset 4A$ such that $|V| \ge K^{-O(\log^3 K)}|A|$.*

The deduction of Theorem 1.3 from Theorem 1.4 is standard given some basic tools and results in additive combinatorics. Given Theorem 1.4, one may conjecture a polynomial version of it, which would in particular imply in a similar way the polynomial Freiman-Ruzsa conjecture.

Conjecture 1.5 (Polynomial Bogolyubov-Ruzsa conjecture)**.** *Let $A \subset \mathbb{F}_2^n$ be a set such that $|A+A| \le K|A|$. Then there exists a linear subspace $V \subset tA$ such that $|V| \ge K^{-c}|A|$, where $t \ge 1, c > 0$ are absolute constants.*

We note that it is not clear whether Conjecture 1.5 is indeed stronger than Conjecture 1.2, or whether one can deduce it assuming Conjecture 1.2.

1.1 Proof overview

We first show, using standard techniques in additive combinatorics, that

1. Theorem 1.3 follows from Theorem 1.4.
2. It suffices to prove Theorem 1.4 for 'large sets' $A \subset \mathbb{F}_2^n$ for which $|A| \ge K^{-1} \cdot 2^n$.

Explicitly, these reductions use a theorem of Ruzsa which bounds the size of $|tA|$ for sets of small doubling, and the notion of a Freiman-homomorphism. We thus assume from now on that $|A| \ge K^{-1} \cdot 2^n$. We then show that there exists a large set $X \subset \mathbb{F}_2^n$ such that $tX \subset 4A$ for $t = O(\log K)$. In fact, we will show a stronger property. For any $x_1, \ldots, x_t \in X$,

$$\Pr_{a_1,a_2 \in A}[a_1 + a_2 + x_1 + \ldots + x_t \in 2A] \ge 0.9 \ . \tag{1}$$

This utilizes an argument of Croot and Sisask [CS10]. The set X allows us to find a large vector space V such that $V \subset 4A$. This is achieved by choosing V to be the subspace orthogonal to the large Fourier coefficients of X. The proof of this latter claim is achieved by applying (1) to randomly chosen $x_1, \ldots, x_t \in X$ and appealing to standard Fourier arguments and Chang's lemma.

Paper organization We give some preliminaries in Section 2. We establish the reductions in Section 3. We prove the existence of the set X in Section 4. We conclude with the Fourier argument in Section 5.

2 Preliminaries

Norms Let $f : \mathbb{F}_2^n \to \mathbb{R}$ be a function. For $1 \le p \le \infty$, its ℓ_p norm is defined as $\|f\|_p = \left(\mathbb{E}_{x \in \mathbb{F}_2^n}[|f(x)|^p]\right)^{1/p}$. Let $f, g : \mathbb{F}_2^n \to \mathbb{R}$ be functions. Their inner product is defined as $\langle f, g \rangle = \mathbb{E}_{x \in \mathbb{F}_2^n}[f(x)g(x)]$. For $1 \le p, q \le \infty$ such that $1/p + 1/q = 1$, the Hölder inequality states that $|\langle f, g \rangle| \le \|f\|_p \|g\|_q$.

Fourier analysis Let $f : \mathbb{F}_2^n \to \mathbb{R}$ be a function. Its Fourier coefficients are $\widehat{f}(\alpha) = \mathbb{E}_{x \in \mathbb{F}_2^n}[f(x)(-1)^{\langle x, \alpha \rangle}]$ where $\alpha \in \mathbb{F}_2^n$. Parseval identity asserts that $\|f\|_2^2 = \sum_{\alpha \in \mathbb{F}_2^n} \widehat{f}(\alpha)^2$. For functions $f, g : \mathbb{F}_2^n \to \mathbb{R}$ their convolution $f * g : \mathbb{F}_2^n \to \mathbb{R}$ is defined as $(f * g)(x) = \mathbb{E}_{y \in \mathbb{F}_2^n} f(y) g(x+y)$. The Fourier coefficients of the convolution obey $\widehat{f * g}(\alpha) = \widehat{f}(\alpha)\widehat{g}(\alpha)$.

3 Reductions

We show in this section that

1. Theorem 1.3 follows from Theorem 1.4.

2. It suffices to prove Theorem 1.4 for 'large sets' $A \subset \mathbb{F}_2^n$ for which $|A| \geq K^{-1} \cdot 2^n$.

3.1 First reduction

We first show how Theorem 1.3 follows from Theorem 1.4. This requires the following theorem of Plünnecke [Plu69] and Ruzsa [Ruz99], showing that if A has small doubling then tA cannot be too large.

Theorem 3.1 ([Plu69, Ruz99]). *Let $A \subset \mathbb{F}_2^n$ be a set such that $|A + A| \leq K|A|$. Then for any $t \geq 1$ we have that $|tA| \leq K^t |A|$.*

Let $A \subset \mathbb{F}_2^n$ be a set such that $|A + A| \leq K|A|$. Theorem 1.4 asserts that there exists a linear subspace $V \subset 4A$ of size $|V| \geq \delta |A|$ where $\delta = K^{-O(\log^3 K)}$. Let $S \subset A$ be maximal such that elements of S fall in different cosets of V; that is, $s + s' \notin V$ for all $s, s' \in S$. We claim that $|S| \leq K^5/\delta$, since

$$|S||V| = |S + V| = |A + V| \subset |A + 4A| = |5A| \leq K^5 |A|,$$

where the last inequality follows from Theorem 3.1. Let $A' = A \cap (V + s)$ where $s \in S$ is chosen to maximize $|A'|$. We have that $|A'| \geq |A|/|S| = K^{-O(\log^3)K}|A|$, and that $|\mathrm{Span}(A')| \leq 2|V| \leq 2K^5|A'|$.

3.2 Second reduction

We next show it suffices to prove Theorem 1.4 for large sets. This requires the notion of a *Freiman homomorphism.* Let $A \subset \mathbb{F}_2^n$. A linear map $\phi : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is said to be a Freiman homomorphism of A of order t if ϕ is injective on tA. That is, for any $a_1, \ldots, a_t, b_1, \ldots, b_t \in A$,

$$\phi(a_1) + \ldots + \phi(a_t) = \phi(b_1) + \ldots + \phi(b_t) \quad \Rightarrow \quad a_1 + \ldots + a_t = b_1 + \ldots + b_t.$$

The following claim is very useful.

Claim 3.2. *Let $A \subset \mathbb{F}_2^n$. Let m be minimal such that a Freiman homomorphism $\phi : \mathbb{F}_2^n \to \mathbb{F}_2^m$ of A of order t exists. Then $\phi(2tA) = \mathbb{F}_2^m$.*

Proof. We first note that m is well defined since for $m = n$ the identity map is a Freiman homomorphism of all orders. Assume by contradiction that $\phi(2tA) \subsetneq \mathbb{F}_2^m$. We will show that m cannot be minimal. Indeed, there must exist $x \in \mathbb{F}_2^m \setminus \phi(2tA)$. Let $\psi : \mathbb{F}_2^m \to \mathbb{F}_2^{m-1}$ be a surjective linear map which sends x to zero, and define $\phi' = \psi \circ \phi$. We claim that ϕ' is also a Freiman homomorphism of A of order t, which contradicts the minimality of m. To show that, we need to show that ϕ' is injective on tA. If this is not the case, then there exist distinct $a, b \in tA$ such that $\phi'(a) = \phi'(b)$, that is $\psi(\phi(a)) = \psi(\phi(b))$. Now, by definition of ψ this can only occur if $\phi(a) = \phi(b)$ or $\phi(a) = \phi(b) + x$. The first case is ruled out since we assumed ϕ is injective on tA, hence by the linearity of ϕ we have that $x = \phi(a+b) \in \phi(2tA)$, violating our initial assumption. □

We now show it suffices to prove Theorem 1.4 for large sets. We will assume throughout that $0 \in A$, which can be assumed without loss of generality by replacing A with $A + a$ for some $a \in A$. Let $A \subseteq \mathbb{F}_2^n$ be such that $|A + A| \leq K|A|$. Let $\phi : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be a minimal Freiman homomorphism of A of order 12 and define $A' = \phi(A)$. We note that by the assumption that $0 \in A$, we have that ϕ is injective on tA for all $t \leq 12$.

We first note that A' also has doubling K, since $|A'| = |A|$ and $|A' + A'| = |A + A|$ because by assumption ϕ is injective on both A and $2A$. This implies that A' is large in $\mathbb{F}_2^m$ since

$$|\mathbb{F}_2^m| = |24A'| \leq K^{24}|A'|,$$

where the equality follows from Claim 3.2 and the inequality from Theorem 3.1. We can thus apply the assumed Theorem 1.4 for large sets on A'. The theorem asserts the existence of a linear subspace $V' \subset 4A'$ of size $V' \geq \delta|A'|$ where $\delta = \exp(-O(\log^4(K^{24})))$. Since ϕ is injective on $12A$ we can define a local inverse $\phi^{-1} : 12A' \to 12A$. In particular, set $V = \phi^{-1}(4A') \subset 4A$. We will show that V is also a linear subspace, thus establishing the theorem for A.

We will use the fact that the property of being a linear subspace can be verified by local tests. Specifically, we need to show that for any $x, y \in V$ we have that $x + y \in V$. Let $x' = \phi(x), y' = \phi(y)$. Then $x', y' \in V'$ and hence $z' = x' + y' \in V'$ since V' is a linear subspace. Let $z = \phi^{-1}(z') \in V$. We need to show that $x + y = z$. Note that since $x, y, z \in V \subset 4A$ then $x+y+z \in 12A$. However, $\phi(x+y+z) = x'+y'+z' = 0$ and since ϕ is injective on $12A$ and since $0 \in 12A$ is mapped by ϕ to zero, we must have that $x+y+z = 0$.

4 Existence of a large near-invariant set

We establish the following lemma in this section

Lemma 4.1. *Let $A \subset \mathbb{F}_2^n$ be such that $|A| \geq K^{-1} \cdot 2^n$. Set $t = O(\log K)$. Then there exist $X \subset \mathbb{F}_2^n$ of size $|X| \geq K^{-O(\log^3(K))} \cdot 2^n$ such that for any $x_1, \ldots, x_t \in X$,*

$$\Pr_{a_1,a_2 \in A}[a_1 + a_2 + x_1 + \ldots + x_t \in 2A] \geq 0.9\ .$$

We first fix some notations. For a set $A \subset \mathbb{F}_2^n$ let $\mathbf{1}_A : \mathbb{F}_2^n \to \{0,1\}$ denote the indicator function for A, and $\varphi_A(x) = \frac{2^n}{|A|}\mathbf{1}_A(x)$ denote the normalized indicator for which $\mathbb{E}_{x \in \mathbb{F}_2^n}[\varphi_A(x)] = 1$. Given two functions $f, g : \mathbb{F}_2^n \to \mathbb{R}$ define their convolution by $(f * g)(x) = \mathbb{E}_{y \in \mathbb{F}_2^n}[f(y)g(x+y)]$ and their inner product by $\langle f, g \rangle = \mathbb{E}_{x \in \mathbb{F}_2^n}[f(x)g(x)]$. Note that $(\varphi_A * f)(x) = \mathbb{E}_{a \in A}[f(x+a)]$ is a 'smoothing' of the function f by a random shift from a set A. For an element $x \in X$ we shorthand $\varphi_x = \varphi_{\{x\}}$ and note that $(\varphi_x * f)(y) = f(x+y)$ is a shift of f by x. In these notations, for any $x \in \mathbb{F}_2^n$ we have that

$$\Pr_{a_1,a_2 \in A}[a_1 + a_2 + x \in 2A] = \langle \varphi_x * \varphi_A * \varphi_A, \mathbf{1}_{2A} \rangle = \langle \varphi_x * \varphi_A * \mathbf{1}_{2A}, \varphi_A \rangle . \tag{2}$$

Note that for $x = 0$,

$$\langle \varphi_A * \mathbf{1}_{2A}, \varphi_A \rangle = \Pr_{a_1,a_2 \in A}[a_1 + a_2 \in 2A] = 1. \tag{3}$$

We will show that there exists a large set $X \subset \mathbb{F}_2^n$ so that for all $x \in tX$, $\varphi_x * \varphi_A * \mathbf{1}_{2A} \approx \varphi_A * \mathbf{1}_{2A}$. In particular, this will show that (2) $\approx$ (3) and would imply Lemma 4.1. In order to do so, we will use the following lemma of Croot and Sisask [CS10]. The lemma shows that if we take a bounded function f and smooth it by a random shift from a large set A, then the resulting function will be nearly invariant to many shifts.

Lemma 4.2. *Let $A \subset \mathbb{F}_2^n$ be a set such that $|A| \geq K^{-1} \cdot 2^n$. Let $f : \mathbb{F}_2^n \to [0,1]$ be a function. Let $p \geq 1$ and $\varepsilon > 0$ be parameters. Then there exists a set $X \subset \mathbb{F}_2^n$ of size $|X| \geq K^{-O(p/\varepsilon^2)} \cdot 2^n$ such that for any $x \in X$,*

$$\|\varphi_x * \varphi_A * f - \varphi_A * f\|_p \leq \varepsilon.$$

We first show how Lemma 4.1 follows from 4.2. Set $f = \mathbf{1}_{2A}, p = \log K, \varepsilon = 1/(20t) = O(1/\log K)$ in Lemma 4.2 so that $|X| \geq K^{-O(\log^3(K))} \cdot 2^n$ as claimed. We first claim that for any $x \in tX$ (where $t = O(\log K)$) we have that

$$\|\varphi_x * \varphi_A * \mathbf{1}_{2A} - \varphi_A * \mathbf{1}_{2A}\|_p \leq t\varepsilon. \tag{4}$$

In order to establish (4) let $x = x_1 + \ldots + x_t$ where $x_1, \ldots, x_t \in X$ and expand it as a telescopic sum. Then

$$\begin{aligned}
&\|\varphi_{x_1+\ldots+x_t} * \varphi_A * \mathbf{1}_{2A} - \varphi_A * \mathbf{1}_{2A}\|_p \\
&\leq \sum_{i=1}^{t} \|\varphi_{x_1+\ldots+x_i} * \varphi_A * \mathbf{1}_{2A} - \varphi_{x_1+\ldots+x_{i-1}} * \varphi_A * \mathbf{1}_{2A}\|_p \\
&= \sum_{i=1}^{t} \|\varphi_{x_i} * \varphi_A * \mathbf{1}_{2A} - \varphi_A * \mathbf{1}_{2A}\|_p \leq t\varepsilon,
\end{aligned}$$

where we used the fact that the ℓ_p norm is invariant under shifts, that is $\|\varphi_x * g\|_p = \|g\|_p$ for all elements $x \in \mathbb{F}_2^n$ and functions $g : \mathbb{F}_2^n \to \mathbb{R}$. By our setting of $\varepsilon = 1/(20t) = O(1/\log K)$, we have that for all $x \in tX$

$$\|\varphi_x * \varphi_A * \mathbf{1}_{2A} - \varphi_A * \mathbf{1}_{2A}\|_p \leq t\varepsilon \leq 1/20. \tag{5}$$

We next apply Hölder inequality. We have that

$$|\langle \varphi_x * \varphi_A * \mathbf{1}_{2A} - \varphi_A * \mathbf{1}_{2A}, \varphi_A \rangle| \leq \|\varphi_x * \varphi_A * \mathbf{1}_{2A} - \varphi_A * \mathbf{1}_{2A}\|_p \|\varphi_A\|_q \tag{6}$$

where $q = p/(p-1)$ is the dual of p. By the choice of $p = \log K$ we have that

$$\|\varphi_A\|_q = (2^n/|A|)^{1-1/q} \leq K^{1/(\log K - 1)} \leq 2. \tag{7}$$

Combining (2), (3), (6) and (7) we conclude that for any $x \in tX$,

$$\begin{aligned} \Pr_{a_1, a_2 \in A}[a_1 + a_2 + x \in 2A] &= \langle \varphi_x * \varphi_A * \mathbf{1}_{2A}, \varphi_A \rangle \\ &= 1 - \langle \varphi_A * \mathbf{1}_{2A} - \varphi_x * \varphi_A * \mathbf{1}_{2A}, \varphi_A \rangle \geq 0.9 \end{aligned} \tag{8}$$

which concludes the proof of Lemma 4.1. We now move to prove Lemma 4.2. The proof will use the Marcinkiewicz-Zygmund inequality [MZ37], which is a generalization of the Khintchine inequality.

Theorem 4.3 (Marcinkiewicz-Zygmund inequality)**.** *Let $X_1, \ldots, X_\ell$ be independent, mean zero random variables with $\mathbb{E}|X_i|^p < \infty$. Then for any $p \geq 1$,*

$$\mathbb{E}\left[|X_1 + \ldots + X_\ell|^p\right] \leq (Cp)^{p/2} \cdot \mathbb{E}\left[\left(|X_1|^2 + \ldots + |X_\ell|^2\right)^{p/2}\right],$$

where $C > 0$ is an absolute constant.

We will actually only need the following corollary for bounded random variables.

Corollary 4.4. *Let $X_1, \ldots, X_\ell$ be independent, mean zero random variables with $|X_i| \leq 1$. Then for any $p \geq 1$,*

$$\mathbb{E}\left[\left|\frac{1}{\ell}(X_1 + \ldots + X_\ell)\right|^p\right] \leq (Cp/\ell)^{p/2}.$$

We now turn to prove Lemma 4.2.

Proof of Lemma 4.2. Let $A \subset \mathbb{F}_2^n$ be a set of size $|A| \geq K^{-1} \cdot 2^n$ and let $f : \mathbb{F}_2^n \to [0,1]$ be a function. For ℓ to be determined later let $a_1, \ldots, a_\ell$ be uniformly chosen elements from A. We first claim if ℓ is chosen large enough, then $\varphi_A * f$ can be approximated by $\frac{1}{\ell}\sum_{i=1}^{\ell} \varphi_{a_i} * f$. That is, we approximate the 'smoothing' of f with a random shift from A, by a random shift from the empirical sample $a_1, \ldots, a_\ell$. Explicitly, we will show that for $\ell = O(p/\varepsilon^2)$ we have that

$$\Pr_{a_1, \ldots, a_\ell \in A}\left[\|\varphi_A * f - \frac{1}{\ell}\sum_{i=1}^{\ell} \varphi_{a_i} * f\|_p \leq \varepsilon/2\right] \geq 1/2. \tag{9}$$

In order to show (9), we will establish that

$$\mathbb{E}_{a_1, \ldots, a_\ell \in A}\left[\|\varphi_A * f - \frac{1}{\ell}\sum_{i=1}^{\ell} \varphi_{a_i} * f\|_p^p\right] \leq (Cp/\ell)^{p/2}, \tag{10}$$

where $C > 0$ is an absolute constant, and then apply the Markov bound. Now, (10) follows from Corollary 4.4. Define $X_i = \varphi_A * f - \varphi_{a_i} * f$ so that $X_i(x) = \mathbb{E}_{a \in A}[f(x+a)] - f(x+a_i)$. Then

$$\|\varphi_A * f - \frac{1}{\ell}\sum_{i=1}^{\ell} \varphi_{a_i} * f\|_p^p = \mathbb{E}_{x \in \mathbb{F}_2^n}\left[\left|\frac{1}{\ell}(X_1(x) + \ldots + X_\ell(x))\right|^p\right],$$

and the claim follows by averaging over $a_1, \ldots, a_\ell$ and applying Corollary 4.4.

Let $S(A) \subset (\mathbb{F}_2^n)^\ell$ denote the set of $(a_1, \ldots, a_\ell)$ for which $\|\varphi_A * f - \frac{1}{\ell}\sum_{i=1}^{\ell} \varphi_{a_i} * f\|_p \le \varepsilon/2$. We have just shown that by our choice of ℓ, at least half the sequences $(\alpha_1, \ldots, \alpha_\ell) \in A^\ell$ have this property. Hence

$$|S(A)| \ge 0.5|A|^\ell \ge 0.5K^{-\ell} \cdot 2^{n\ell} \,. \tag{11}$$

Applying the same argument to any shift $A + x$ of A we deduce that $S(A+x) \ge 0.5K^{-\ell} \cdot 2^{n\ell}$ as well. Hence, by an averaging argument there must exist a subset $X' \subset \mathbb{F}_2^n$ of size $|X'| \ge 0.5K^{-\ell} \cdot 2^n$ and a sequence $(a_1, \ldots, a_\ell) \in (\mathbb{F}_2^n)^\ell$ such that $(a_1, \ldots, a_\ell) \in S(A+x)$ for all $x \in X'$. But then we get that for all $x', x'' \in X'$ we have that

$$\|\varphi_{A+x'} * f - \varphi_{A+x''} * f\|_p \le \|\varphi_{A+x'} * f - \frac{1}{\ell}\sum_{i=1}^{\ell} \varphi_{a_i} * f\|_p + \|\varphi_{A+x''} * f - \frac{1}{\ell}\sum_{i=1}^{\ell} \varphi_{a_i} * f\|_p \le \varepsilon.$$

Let $x' \in X'$ be arbitrary and set $X = X' + x'$. We conclude that for any $x \in X$,

$$\|\varphi_{A+x} * f - \varphi_A * f\|_p = \|\varphi_{A+x+x'} * f - \varphi_{A+x'} * f\|_p \le \varepsilon.$$

□

5 A Fourier analytic argument

Let $A \subset \mathbb{F}_2^n$ be a set such that $|A| \ge K^{-1} \cdot 2^n$. We showed in Lemma 4.1 that there exists a set $X \subset \mathbb{F}_2^n$ of size $|X| \ge K^{-O(\log^3 K)} \cdot 2^n$ such that for any $x \in tX$, where $t = O(\log K)$, we have that

$$\Pr_{a_1, a_2 \in A}[a_1 + a_2 + x \in 2A] \ge 0.9 \,. \tag{12}$$

We now show that the linear subspace $V \subset \mathbb{F}_2^n$ which is orthogonal to the large Fourier coefficients of X is contained in $4A$. In order to show that, we apply (12) for $x = x_1 + \ldots + x_t$ where $x_1, \ldots, x_t \in X$ are chosen uniformly, and deduce that

$$\Pr_{a_1, a_2 \in A, x_1, \ldots, x_t \in X}[a_1 + a_2 + x_1 + \ldots + x_t \in 2A] \ge 0.9 \,. \tag{13}$$

We shorthand for a set $X \subset \mathbb{F}_2^n$ by $\widehat{X}(\alpha)$ the Fourier coefficients of φ_X,

$$\widehat{X}(\alpha) = \widehat{\varphi_X}(\alpha) = \mathbb{E}_{x \in X}[(-1)^{\langle \alpha, x \rangle}].$$

Note that $\widehat{X}(0) = 1$. The spectrum of a set X is the set of its large Fourier coefficients. Explicitly, its γ-spectrum for $0 < \gamma < 1$ is defined as

$$\mathrm{Spec}_\gamma(X) = \{\alpha \in \mathbb{F}_2^n : |\widehat{X}(\alpha)| \geq \gamma\}.$$

Parseval's identity allows one to bound $|\mathrm{Spec}_\gamma(X)| \leq (2^n/|X|) \cdot (1/\gamma)^2$. A better bound on the dimension of $\mathrm{Spec}_\gamma(X)$ is given by Chang's theorem [Cha02].

Theorem 5.1 (Chang)**.** *Let* $X \subseteq \mathbb{F}_2^n$*. Then* $\dim(\mathrm{Spec}_\gamma(X)) \leq 8\log(2^n/|X|) \cdot (1/\gamma)^2$.

Define the vector space $V \subseteq \mathbb{F}_2^n$ as the orthogonal space to $\mathrm{Spec}_{1/2}(X)$.

$$V = \mathrm{Spec}_{1/2}(X)^\perp = \{v \in \mathbb{F}_2^n : \langle v, \alpha \rangle = 0 \ \forall \alpha \in \mathrm{Spec}_{1/2}(X)\}.$$

Theorem 5.1 implies that $|V| \geq (|X|/2^n)^{32} \cdot 2^n = K^{-O(\log^3 K)} \cdot 2^n$. We next show that $V \subset 4A$. We will do so by showing that

$$\Pr[a_1 + a_2 + x_1 + \ldots + x_t + v \in 2A] \approx \Pr[a_1 + a_2 + x_1 + \ldots + x_t \in 2A] \geq 0.9,$$

where $a_1, a_2 \in A$, $x_1, \ldots, x_t \in X$ and $v \in V$ are chosen uniformly. In particular we will show that

$$\Pr[a_1 + a_2 + x_1 + \ldots + x_t + v \in 2A] \geq 0.8 \ .$$

Hence, there exists a fixing for $b = a_1 + a_2 + x_1 + \ldots + x_t$ such that $|V \cap (2A + b)| \geq 0.8|V|$. This implies that $V \subset 4A$: every element $v \in V$ can be written in $|V|/2$ disjoint ways as $v = v_1 + v_2$ where $v_1, v_2 \in V$, and at least for one of these it must hold that $v_1, v_2 \in 2A + b$ and hence $v = v_1 + v_2 \in 4A$.

To conclude the proof, we will show that

$$|\Pr[a_1 + a_2 + x_1 + \ldots + x_t + v \in 2A] - \Pr[a_1 + a_2 + x_1 + \ldots + x_t \in 2A]| \leq 0.1,$$

where again $a_1, a_2 \in A$, $x_1, \ldots, x_t \in X$ and $v \in V$ are chosen uniformly. We now apply Fourier analysis. We can rewrite

$$\Pr_{a_1,a_2 \in A, x_1,\ldots,x_t \in X}[a_1 + a_2 + x_1 + \ldots + x_t \in 2A] = \sum_{\alpha \in \mathbb{F}_2^n} \widehat{A}(\alpha)^2 \widehat{X}(\alpha)^t \widehat{1_{2A}}(\alpha). \tag{14}$$

and

$$\Pr_{a_1,a_2 \in A, x_1,\ldots,x_t \in X, v \in V}[a_1 + a_2 + x_1 + \ldots + x_t + v \in 2A] = \sum_{\alpha \in \mathbb{F}_2^n} \widehat{A}(\alpha)^2 \widehat{X}(\alpha)^t \widehat{V}(\alpha) \widehat{1_{2A}}(\alpha). \tag{15}$$

The Fourier coefficients of V are simple to describe since it is a linear subspace. We have that $\widehat{V}(\alpha) = 1$ if $\alpha \in V^\perp$ and that $\widehat{V}(\alpha) = 0$ otherwise. Thus

$$\begin{aligned} &\Pr[a_1 + a_2 + x_1 + \ldots + x_t \in 2A] - \Pr[a_1 + a_2 + x_1 + \ldots + x_t + v \in 2A] \\ &= \sum_{\alpha \notin V^\perp} \widehat{A}(\alpha)^2 \widehat{X}(\alpha)^t \widehat{1_{2A}}(\alpha). \end{aligned} \tag{16}$$

We now bound (16). By the definition of V, we have that if $\alpha \notin V^{\perp}$ then $\alpha \notin \mathrm{Spec}_{1/2}(X)$, and hence

$$|\widehat{X}(\alpha)|^t \le 2^{-t}.$$

Moreover, $|\widehat{1_{2A}}(\alpha)| \le 1$ and

$$\sum_{\alpha \notin V^{\perp}} \widehat{A}(\alpha)^2 \le \sum_{\alpha \in \mathbb{F}_2^n} \widehat{A}(\alpha)^2 = \mathbb{E}_{x \in \mathbb{F}_2^n}[\varphi_A(x)^2] = K,$$

Thus we conclude since

$$|\Pr[a_1 + a_2 + x_1 + \ldots + x_t \in 2A] - \Pr[a_1 + a_2 + x_1 + \ldots + x_t + v \in 2A]| \le 2^{-t}K \le 0.1$$

by choosing $t = \log(10K)$.

6 Acknowledgements

I thank Eli Ben-Sasson, Zeev Dvir, Hamed Hatami, Amir Shpilka, Avi Wigderson, Noga Zewi and David Zuckerman for the encouragement to write this exposition and for useful comments and suggestions on an earlier version of this note.

References

[BDL12] Abhishek Bhowmick, Zeev Dvir, and Shachar Lovett. New bounds for matching vector codes. submitted, 2012.

[BSLZ11] Eli Ben-Sasson, Shachar Lovett, and Noga Zewi. An additive combinatorics approach to the log-rank conjecture in communication complexity. 2011.

[BSZ11] Eli Ben-Sasson and Noga Zewi. From affine to two-source extractors via approximate duality. In *STOC*, pages 177–186, 2011.

[Cha02] Mei-Chu Chang. A polynomial bound in freiman's theorem. *Duke Mathematical Journal*, 113(3):399–419, 2002.

[CS10] Ernie Croot and Olof Sisask. A probabilistic technique for finding almost-periods of convolutions. *Geometric And Functional Analysis*, 20:1367–1396, 2010. 10.1007/s00039-010-0101-8.

[EZ11] Chaim Even-Zohar. On sums of generating sets in $\mathbb{Z}_2^n$. 2011.

[Fre73] G.A Freiman. *Foundations of a structural theory of set addition*, volume 37. American Mathematical Society, Translations of Mathematical Monographs, 1973.

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

[Gre05] Ben Green. The polynomial freiman-ruzsa conjecture, 2005. http://www.dpmms.cam.ac.uk/ bjg23/papers/PFR.pdf.

[GT09] Ben Green and Terence Tao. Freiman's theorem in finite fields via extremal set theory. *Comb. Probab. Comput.*, 18:335–355, May 2009.

[Kon08] S. V. Konyagin. On the Freiman theorem in finite fields. *Mathematical Notes*, 84:435–438, 2008.

[Lov10] Shachar Lovett. Equivalence of polynomial conjectures in additive combinatorics. 2010.

[MZ37] J. Marcinkiewicz and A. Zygmund. Sur les fonction independantes. *Func. Math. 29*, 1937.

[Plu69] H. Plunneke. Eigenschaften und abschätzungen von wirkingsfunktionen. *Gesellschaft für Mathematik und Datenverarbeitung*, 1969.

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

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

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

[San10] Tom Sanders. On the bogolyubov-ruzsa lemma. Submitted, 2010.