Chapter 5: Quasi-polynomial Freiman–Ruzsa

Adam Sheffer

May 3, 2016

1 Introduction

At the end of Chapter 1 we introduced the *polynomial Freiman–Ruzsa* conjecture over $\mathbb{F}_2^n$, which seems to be one of the main open problems in additive combinatorics. The goal of this chapter is to study the current best result for this conjecture — Sanders' *quasi-polynomial* Freiman–Ruzsa theorem [7].

Recall from Chapter 1 that we define $\text{Span}(A) = \bigcup_{k\geq 1} kA$. When working over $\mathbb{F}_2^n$, it suffices to write $\text{Span}(A) = \bigcup_{k=1}^{n} kA$.

Theorem 1.1 (Quasi-polynomial Freiman–Ruzsa in $\mathbb{F}_2^n$). *Let $A \subset \mathbb{F}_2^n$ be a set satisfying $|A+A| \leq k|A|$. Then there exists $A^* \subset A$ such that $|A^*| \geq k^{-O(\log^3 k)}|A|$ and $|\text{Span}(A^*)| \leq 4k^5|A^*|$.*

The proof of Theorem 1.1 that we present in this chapter follows the presentation of Lovett [4]. It is the longest proof in these lecture notes, and is based on Fourier analysis, probabilistic arguments, and studying the translations of a set A that have a large intersection with A.

2 Reducing the problem

Our first step in proving Theorem 1.1 is to reduce it to the following theorem.

Theorem 2.1. *Let $A \subset \mathbb{F}_2^n$ be a set with $|A| \geq \delta 2^n$, for some $0 < \delta < 1$. Then there exists a linear subspace $V \subset 4A$ with $|V| \geq \delta^{O(\lg^3(1/\delta))}|A|$.*

This type of theorem is sometimes called a *Bogolyubov-Ruzsa lemma*. Notice that a set $A \subset \mathbb{F}_2^n$ that satisfies $|A| \geq \delta 2^n$ also satisfies $|A+A| \leq |A|/\delta$. That is, the condition of Theorem 2.1 is in a sense stronger than having small doubling.

Before getting to the reduction, we first introduce another property of Freiman homomorphisms. For a set $A \subset \mathbb{F}_2^n$, a function $\tau : \mathbb{F}_2^n \to \mathbb{F}_2^m$, and a positive integer t, recall that τ is a *Freiman t-homomorphism* of A if the following property holds. If $\tau(a_1) + \tau(a_2) + \cdots + \tau(a_t) = \tau(b_1) + \tau(b_2) + \cdots + \tau(b_t)$ then $a_1 + a_2 + \cdots + a_t = b_1 + b_2 + \cdots + b_t$ (where $a_1, \ldots, a_k, b_1, \ldots, b_k \in A$). We say that a function $\tau : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is *linear* if every $x, y \in \mathbb{F}_2^n$ satisfy $f(x+y) = f(x) + f(y)$.

Lemma 2.2. *Let $A \subset \mathbb{F}_2^n$, let t be a positive integer, and let m be the smallest integer such that there exists a linear Freiman t-homomorphism $\tau : \mathbb{F}_2^n \to \mathbb{F}_2^m$ of A. Then $\tau(2tA) = \mathbb{F}_2^m$.*

Lemma 2.2 is well-defined in the sense that there is always an integer m such that a linear Freiman t-homomorphism $\tau : \mathbb{F}_2^n \to \mathbb{F}_2^m$ of A exists. Indeed, we can always take τ to be the identity map to obtain a Freiman t-homomorphism of A with $m = n$.

Proof of Lemma 2.2. Assume for contradiction that there exists an element $x \in \mathbb{F}_2^m \setminus \tau(2tA)$. Let $\varphi : \mathbb{F}_2^m \to \mathbb{F}_2^{m-1}$ be a surjective linear map that sends x to 0; for example, we may set $\varphi(y) = \pi(y) + \pi(x)$ where $\pi : \mathbb{F}_2^m \to \mathbb{F}_2^{m-1}$ is the projection $\pi(y_1, \ldots, y_m) = (y_1, \ldots, y_{m-1})$. Notice that for every $z \in \mathbb{F}_2^{m-1}$ there are exactly two elements $y \in \mathbb{F}_2^m$ such that $z = \varphi(y)$. If we denote these two elements as y, y' then $y = y' + x$.

Let $\tau' = \varphi \circ \tau$. Since τ' is the composition of two linear functions, it is also linear. Assume that $\tau'(a_1) + \cdots + \tau'(a_t) = \tau'(b_1) + \cdots + \tau'(b_t)$ for $a_1, \ldots, a_t, b_1 \ldots, b_t \in A$. Since τ' is linear, we get that $\tau'(a_1 + \cdots + a_t) = \tau'(b_1 + \cdots + b_t)$. Let $a = a_1 + \cdots + a_t$ and $b = b_1 + \cdots + b_t$ (so $a, b \in tA$). That is, $\tau'(a) = \tau'(b)$, or equivalently $\varphi(\tau(a)) = \varphi(\tau(b))$. For this to happen, we either have $\tau(a) = \tau(b)$ or $\tau(a) = \tau(b) + x$. The latter case cannot happen since it would imply $x = \tau(a) + \tau(b) = \tau(a+b)$ while we assume that $x \in \mathbb{F}_2^m \setminus \tau(2tA)$. Thus $\tau(a) = \tau(b)$, which in turn implies $a = b$ (since $a, b \in tA$ and τ is a Freiman t-homomorphism of A). This establishes that τ' is a linear Freiman t-homomorphism. Since $\tau' : \mathbb{F}_2^n \to \mathbb{F}_2^{m-1}$, we get a contradiction to the minimality of m and τ. □

We also recall Plünnecke's inequality from Chapter 1.

Theorem 2.3. *Let A and B be finite subsets of an abelian group $\mathcal{G}$, such that $|A| = |B|$ and $|A + B| \leq k|A|$. Then for any two positive integers m and ℓ, we have $|mA - \ell A| \leq k^{m+\ell}|A|$.*

We are now ready to derive our reduction.

Lemma 2.4. *Theorem 2.1 implies Theorem 1.1.*

Proof. We assume that Theorem 2.1 holds and use it to prove Theorem 1.1. That is, we are given a set $A \subset \mathbb{F}_2^n$ with $|A+A| \leq k|A|$. If $0 \notin A$ we take $x \in A$ and replace A with $A_x = \{x+y : y \in A\}$. Notice that $|A| = |A_x|$, that $A + A = A_x + A_x$. Also, if $B \subset A_x$ then $|\mathrm{Span}(B+x)| \leq 2|\mathrm{Span}(B)|$, so it suffices to find a subset with a bounded span in A_x.

Finding a large subspace in A. Let m be the minimum integer such that there exists a linear Freiman 12-homomorphism $\tau : \mathbb{F}_2^n \to \mathbb{F}_2^m$ of A. Since $0 \in A$, we have that τ is a linear Freiman t-homomorphism of A for every $t \leq 12$; for example, if for $a, b, c, d \in A$ we have that $\tau(a)+\tau(b) = \tau(c)+\tau(d)$ then $\tau(a)+\tau(b)+\tau(0)+\cdots+\tau(0) = \tau(c)+\tau(d)+\tau(0)+\cdots+\tau(0)$, which in turn implies $a+b = c+d$.

Let $A' = \tau(A)$. Since τ is a Freiman 1-homomorphism of A, we have that $|A| = |A'|$. Since τ is a linear Freiman 2-homomorphism of A, we have that $|A+A| = |A'+A'|$. Combining these two observations gives $|A'+A'| \leq k|A'|$. Lemma 2.2 implies that $\tau(24A) = \mathbb{F}_2^m$. By Theorem 2.3, we have $|24A| \leq |A|k^{24}$. That is,

$$|A'| = |A| \geq |24A|k^{-24} \geq 2^m k^{-24}.$$

We can thus apply Theorem 2.1 on A' with $\delta = k^{-24}$, which yields a linear subspace $V' \subset 4A'$ such that $|V'| \geq \delta^{O(\lg^3(1/\delta))}|A| = k^{-O(\lg^3 k)}|A|$. We found a large linear subspace in $4A'$. We now show that $4A$ also contains a large linear subspace.

Since τ is a Freiman 12-homomorphism of A, it is injective on $12A$. Thus, the inverse $\tau^{-1} : 12A' \to 12A$ is well defined. Set $V = \tau^{-1}(V')$. Consider $v' \in V'$ such that $v' = a+b+c+d$ with $a, b, c, d \in A'$. Then $v = \tau^{-1}(v') = \tau^{-1}(a+b+c+d) = \tau^{-1}(a)+\tau^{-1}(b)+\tau^{-1}(c)+\tau^{-1}(d)$. That is, $v \in 4A$, which in turn implies $V \subset 4A$. For two elements $u, w \in V$, let $u' = \tau(u)$ and $w' = \tau(w)$. Notice that $u', w' \in V'$, and since V' is a linear subspace $u'+w' \in V'$. This implies that $u+w \in V$ (this is the only step for which need τ to be a Freiman 12-homomorphism). Since V is closed under addition, it is also a linear subspace of size at least $k^{-O(\lg^3 k)}|A|$.

Using the large subspace. Let $V^\perp$ be the largest linear subspace in $\mathbb{F}_2^m$ that is orthogonal to V. Then the cosets of V are the sets $u+V$ for every $u \in V^\perp$. Let S be a maximum-sized subset of A with the property that no two elements of S are in the same coset of V. Equivalently, for every $s, s' \in S$ we have $s+s' \notin V$. By combining this with $V \subset 4A$ and with Theorem 2.3, we obtain

$$|S||V| = |S+V| = |A+V| \leq |5A| \leq k^5|A|.$$

By rearranging the above we get $|S| \leq k^5|A|/|V|$. For any $s \in S$ we set $A_s = A \cap (s+V)$. We take $s \in S$ that maximizes $|A_s|$. By the pigeonhole principle we have

$|A_s| \geq |A|/|S| \geq |V|k^{-5}$. Since $\mathrm{Span}(s+V) = V \cup (s+V)$, we obtain

$$|\mathrm{Span}(A_s)| \leq |\mathrm{Span}(s+V)| \leq 2|V| \leq 2k^5|A_s|.$$

By setting $A^* = A_s$, we conclude the proof of Theorem 1.1. □

3 Proving Theorem 2.1

In this section we prove Theorem 2.1. In our proof we rely on the following lemma, which we will prove in Section 5.

Lemma 3.1. *Consider a set $A \subset \mathbb{F}_2^n$ with $|A| \geq \delta 2^n$ and a positive integer $t = O(\lg(1/\delta))$. Then there exists $X \subset \mathbb{F}_2^n$ such that $|X| \geq 2^n \delta^{O(\lg^3(1/\delta))}$ and at least $0.9|A|^2|X|^t$ tuples $(a_1, a_2, x_1, \ldots, x_t) \in A^2 \times X^t$ satisfy $a_1 + a_2 + x_1 + \cdots + x_t \in 2A$.*

Intuitively, the lemma states that for any sufficiently large set A, there exist many translations $2A + x$ that have a large intersection with $2A$.

As with the proofs of the previous chapter, the proof of Theorem 2.1 is based on finding large Fourier coefficients. For a set $A \subset \mathbb{F}_2^n$ and a constant $0 < \gamma < 1$, we define the γ-*spectrum* of A as

$$\mathrm{Spec}_\gamma(A) = \left\{ \alpha \in \mathbb{F}_2^n : \, |\widehat{1_A}(\alpha)| \geq \gamma \frac{|A|}{2^n} \right\}.$$

We recall Parseval's theorem from Chapter 4.

Theorem 3.2. *For any $f : \mathbb{F}_p^n \to \mathbb{C}$, we have*

$$\sum_{\alpha \in \mathbb{F}_p^n} |\hat{f}(\alpha)|^2 = p^{-n} \sum_{x \in \mathbb{F}_p^n} |f(x)|^2.$$

Applying this theorem for 1_A implies $\sum_{\alpha \in \mathbb{F}_2^n} |\widehat{1_A}(\alpha)|^2 = 2^{-n} \sum_{x \in \mathbb{F}_p^n} |1_A(x)|^2 = 2^{-n}|A|$. This in turn implies that

$$|\mathrm{Spec}_\gamma(A)| \leq \frac{2^{-n}|A|}{(\gamma|A|/2^n)^2} = \frac{2^n}{\gamma^2|A|}.$$

This in turn implies that the dimension of $\mathrm{Span}(\mathrm{Spec}_\gamma(A))$ is at most $\frac{2^n}{\gamma^2|A|}$. It might seem naive to bound this dimension by $|\mathrm{Spec}_\gamma(A)|$. Indeed, Chang [2] proved the following stronger result (we only present the special case of $\mathbb{F}_2^n$).

Lemma 3.3. *For any $A \subset \mathbb{F}_2^n$ we have* $\dim \mathrm{Span}(\mathrm{Spec}_\gamma(A)) \leq \dfrac{8\lg(2^n/|A|)}{\gamma^2}$.

We do not prove Lemma 3.3 in these lecture notes. This lemma is also a main ingredient in deriving recent bounds for Roth's theorem (e.g., see [1, 6]).

Proof of Theorem 2.1. We set $t = \lg(10/\delta)$. By Lemma 3.1, there exists $X \subset \mathbb{F}_2^n$ with $X \geq 2^n \delta^{O(\lg^3(1/\delta))}$ such that there are at least $0.9|A|^2|X|^t$ tuples $(a_1, a_2, x_1, \ldots, x_t) \in A^2 \times X^t$ that satisfy $a_1 + a_2 + x_1 + \cdots + x_t \in 2A$.

We consider the vector space

$$V = \left\{ v \in \mathbb{F}_2^n : \, v \cdot \alpha = 0 \text{ for every } \alpha \in \mathrm{Spec}_{1/2}(X) \right\}.$$

That is, V is the linear subspace of the vectors that are orthogonal to every large Fourier coefficient of 1_X. By Lemma 3.3, the dimension of $\mathrm{Span}(\mathrm{Spec}_{1/2}(X))$ is at most $32\lg(2^n/|X|)$. This implies that the dimension of V is at least $n - 32\lg(2^n/|X|)$, so

$$|V| \geq \frac{2^n}{(2^n/|X|)^{32}} = 2^n \delta^{O(\lg^3(1/\delta))}. \tag{1}$$

To complete the proof of the theorem, it remains to show that $V \subset 4A$.

By the definition of a Fourier coefficient, we have

$$\begin{aligned}
&\sum_{\alpha \in \mathbb{F}_2^n} \widehat{1_A}(\alpha)^2 \widehat{1_X}(\alpha)^t \widehat{1_{2A}}(\alpha) \\
&\quad = 2^{-(t+3)n} \sum_{\substack{\alpha, a_1, a_2, \\ x_1, \cdots, x_t, z \in \mathbb{F}_2^n}} 1_A(a_1) 1_A(a_2) 1_X(x_1) \cdots 1_X(x_t) 1_{2A}(z) e^{-\pi i(\alpha \cdot (a_1 + a_2 + x_1 + \cdots + x_t + z))} \\
&\quad = 2^{-(t+3)n} \sum_{\substack{a_1 + a_2 + x_1 \\ + \cdots + x_t + z = 0}} 1_A(a_1) 1_A(a_2) 1_X(x_1) \cdots 1_X(x_t) 1_{2A}(z) \cdot 2^n.
\end{aligned}$$

That is, $2^{(t+2)n} \sum_{\alpha \in \mathbb{F}_2^n} \widehat{1_A}(\alpha)^2 \widehat{1_X}(\alpha)^t \widehat{1_{2A}}(\alpha)$ is the number of tuples $(a_1, a_2, x_1, \ldots, x_t) \in A^2 \times X^t$ that satisfy $a_1 + a_2 + x_1 + \cdots + x_t \in 2A$. We denote this number as N_{2A+tX}. By the definition of X we have $N_{2A+tX} \geq 0.9|A|^2|X|^t$. A similar argument implies that $2^{(t+3)n} \sum_{\alpha \in \mathbb{F}_2^n} \widehat{1_A}(\alpha)^2 \widehat{1_X}(\alpha)^t \widehat{1_{2A}}(\alpha) \widehat{1_V}(\alpha)$ is the number of tuples $(a_1, a_2, x_1, \ldots, x_t, v) \in A^2 \times X^t \times V$ for which $a_1 + a_2 + x_1 + \cdots + x_t + v \in 2A$. We denote this number as $N_{2A+tX+V}$.

For a vector $u \in \mathbb{F}_2^n$, we write $u \perp V$ if v is orthogonal to V; that is, if for every $v \in V$ we have $v \cdot u = 0$. Since V is a linear subspace, we notice that

$$\widehat{1_V}(\alpha) = 2^{-n} \sum_{x \in \mathbb{F}_2^n} 1_V(x) e^{-\pi i(\alpha \cdot x)} = 2^{-n} \sum_{x \in V} e^{-\pi i(\alpha \cdot x)} = \begin{cases} 2^{-n}|V|, & \alpha \perp V, \\ 0, & \text{otherwise.} \end{cases}$$

That is, if $\alpha \perp V$ then $\widehat{1_V}(\alpha)$ is the density of V in $\mathbb{F}_2^n$, and otherwise it is zero (this is obtained by applying the cancellation property inside of the subspace V).

By combining the above we obtain

$$\begin{aligned}
N_{2A+tX+v} &= 2^{(t+3)n} \sum_{\alpha \in \mathbb{F}_2^n} \widehat{1_A}(\alpha)^2 \widehat{1_X}(\alpha)^t \widehat{1_{2A}}(\alpha) \widehat{1_V}(\alpha) \\
&= \frac{|V|}{2^n} \left(2^{(t+3)n} \sum_{\alpha \in \mathbb{F}_2^n} \widehat{1_A}(\alpha)^2 \widehat{1_X}(\alpha)^t \widehat{1_{2A}}(\alpha) - 2^{(t+3)n} \sum_{\alpha \not\perp V} \widehat{1_A}(\alpha)^2 \widehat{1_X}(\alpha)^t \widehat{1_{2A}}(\alpha) \right) \\
&= |V| \left(N_{2A+tX} - 2^{(t+2)n} \sum_{\alpha \not\perp V} \widehat{1_A}(\alpha)^2 \widehat{1_X}(\alpha)^t \widehat{1_{2A}}(\alpha) \right). \qquad (2)
\end{aligned}$$

Notice that the imaginary parts of the Fourier coefficients must cancel themselves out, since $N_{2A+tX+v}$ is an integer. By the definition of V, for any $\alpha \not\perp V$ we get that $\alpha \notin \mathrm{Spec}_{1/2}(X)$ and thus $|\widehat{1_X}(\alpha)| < \frac{|X|}{2^{n+1}}$. We also have $|\widehat{1_{2A}}(\alpha)| \leq \frac{|2A|}{2^n} \leq 1$, and (by Theorem 3.2)

$$\sum_{\alpha \not\perp V} |\widehat{1_A}(\alpha)|^2 \leq \sum_{\alpha \in \mathbb{F}_2^n} |\widehat{1_A}(\alpha)|^2 = 2^{-n} \sum_{x \in \mathbb{F}_2^n} |1_A(x)|^2 = \frac{|A|}{2^n}.$$

Combining these bounds with (2) yields

$$\begin{aligned}
N_{2A+tX+v} &\geq |V| \left(N_{2A+tX} - 2^{(t+2)n} \left(\frac{|X|}{2^{n+1}} \right)^t \sum_{\substack{\alpha \not\perp V \\ \alpha \in \mathbb{F}_2^n}} |\widehat{1_A}(\alpha)|^2 \right) \\
&\geq |V| \left(N_{2A+tX} - 2^{(t+2)n} \left(\frac{|X|}{2^{n+1}} \right)^t \frac{|A|}{2^n} \right).
\end{aligned}$$

Recalling that $N_{2A+tX} \geq 0.9|A|^2|X|^t$ and that $|A| \geq \delta 2^n$ gives

$$N_{2A+tX+v} \geq |V| \left(0.9|A|^2|X|^t - 2^{n-t}|X|^t|A| \right) \geq |V||A||X|^t \left(0.9|A| - |A|/(\delta 2^t) \right).$$

Since $t = \lg(10/\delta)$, we have $N_{2A+tX+v} \geq 0.8|V||A|^2|X|^t$. That is, there are at least $0.8|V||A|^2|X|^t$ tuples $(a_1, a_2, x_1, \ldots, x_t, v) \in A^2 \times X^t \times V$ with $a_1 + a_2 + x_1 + \cdots + x_t + v \in 2A$. By the pigeonhole principle, there exists a choice of $a_1, a_2, x_1, \ldots, x_t$ that forms a valid tuple with at least $0.8|V|$ elements $v \in V$. Equivalently, there exists $b = a_1 + a_2 + x_1 + \cdots + x_t$ such that $|V \cap (b + 2A)| \geq 0.8|V|$.

Since V is a linear subspace, for every $v \in V$ there exist $|V|/2$ pairs $v_1, v_2 \in V$ such that $v = v_1 + v_2$. Since $|V \cap (b+2A)| \geq 0.8|V|$, there must be such pairs $v_1, v_2 \in V$ with $v_1, v_2 \in 2A+b$. Fix such a pair and notice that $v = v_1 + v_2 \in (2A+b)+(2A+b) = 4A$. Since this holds for every $v \in V$, we obtain $V \subset 4A$, as asserted. □

4 Convolutions

The *convolution* of two functions $f, g : \mathbb{F}_p^n \to \mathbb{R}$ is defined as

$$(f * g)(x) = \sum_{y \in \mathbb{F}_p^n} f(y) g(x-y).$$

We will not rely on the following claim. It is given here only to practice the concept of convolutions.

Claim 4.1. *For any two functions $f, g : \mathbb{F}_p^n \to \mathbb{C}$, we have:*
(i) $(\widehat{f} * \widehat{g})(\alpha) = \widehat{fg}(\alpha)$ (where fg stands for the function $f(x) \cdot g(x)$).
(ii) $\widehat{f * g}(\alpha) = p^n \cdot \widehat{f}(\alpha)\widehat{g}(\alpha)$.

Proof. We do not explain steps that were described in detail in Chapter 4. For (i), we observe that

$$\begin{aligned}
(\widehat{f} * \widehat{g})(\alpha) &= \sum_{\beta \in \mathbb{F}_p^n} \widehat{f}(\beta)\widehat{g}(\alpha-\beta) \\
&= \sum_{\beta \in \mathbb{F}_p^n} \left(p^{-n} \sum_{x \in \mathbb{F}_p^n} f(x) e^{-2\pi i (x \cdot \beta)/p} \right) \left(p^{-n} \sum_{y \in \mathbb{F}_p^n} g(y) e^{-2\pi i (y \cdot (\alpha-\beta))/p} \right) \\
&= p^{-2n} \sum_{x,y \in \mathbb{F}_p^n} f(x) g(y) e^{-2\pi i (y \cdot \alpha)/p} \sum_{\beta \in \mathbb{F}_p^n} e^{-2\pi i ((x-y) \cdot \beta)/p} \\
&= p^{-n} \sum_{x \in \mathbb{F}_p^n} f(x) g(x) e^{-2\pi i (x \cdot \alpha)/p} = \widehat{fg}(\alpha).
\end{aligned}$$

Similarly, for (ii) we have

$$\begin{aligned}
\widehat{f * g}(\alpha) &= p^{-n} \sum_{x \in \mathbb{F}_p^n} (f * g)(x) e^{-2\pi i (x \cdot \beta)/p} = p^{-n} \sum_{x \in \mathbb{F}_p^n} e^{-2\pi i (x \cdot \beta)/p} \sum_{y \in \mathbb{F}_p^n} f(y) g(x-y) \\
&= p^{-n} \sum_{y \in \mathbb{F}_p^n} f(y) e^{-2\pi i (y \cdot \beta)/p} \sum_{x \in \mathbb{F}_p^n} g(x-y) e^{-2\pi i ((x-y) \cdot \beta)/p} = p^n \widehat{f}(\alpha)\widehat{g}(\alpha).
\end{aligned}$$

In the last transition we relied on the simple observation $\sum_{x\in\mathbb{F}_p^n} g(x-y)e^{-2\pi((x-y)\cdot\beta)/p} = \sum_{x\in\mathbb{F}_p^n} g(x)e^{-2\pi(x\cdot\beta)/p}$. □

Convolutions are strongly connected to sum sets and to additive energy. Let $A \subset \mathbb{F}_p^n$. For any $x \in \mathbb{F}_p^n$ we have

$$(1_A * 1_A)(x) = \sum_{y\in\mathbb{F}_p^n} 1_A(y)1_A(x-y) = r_A^+(x).$$

We can thus express the additive energy of A as

$$E^+(A) = \sum_{x\in\mathbb{F}_p^n} r_A^+(x)^2 = \sum_{x\in\mathbb{F}_p^n} (1_A * 1_A)(x)^2.$$

We now return to work in $\mathbb{F}_2^n$. In this case, we can equivalently write a convolution as $(f*g)(x) = \sum_{y\in\mathbb{F}_p^n} f(y)g(x+y)$. For a function $f : \mathbb{F}_2^n \to \mathbb{R}$ and a set $A \subset \mathbb{F}_2^n$, we have

$$(1_A * f)(x) = \sum_{y\in\mathbb{F}_2^n} 1_A(y)f(x+y) = \sum_{y\in A} f(x+y).$$

That is, $(1_A * f)(x)$ is the sum of f applied on translations of y by elements of A. Given an element $x \in \mathbb{F}_2^n$, for brevity we write 1_x instead of $1_{\{x\}}$. Note that $(1_x * f)(y) = f(x+y)$.

For a function $f : \mathbb{F}_2^n \to \mathbb{R}$ and a real number $1 \le p \le \infty$, recall that the ℓ_p *norm* of f is defined as

$$\|f\|_p = \left(\sum_{x\in\mathbb{F}_2^n} |f(x)|^p\right)^{1/p}.$$

The triangle inequality states that for every two functions $f, g : \mathbb{F}_2^n \to \mathbb{R}$ and $p \ge 1$, we have $\|f+g\|_p \le \|f\|_p + \|g\|_p$.

5 Rich translations

In this section we consider large sets $A \subset \mathbb{F}_2^n$, and show that there exist many translations $x + A$ (where $x \in \mathbb{F}_2^n$) that have a large intersection with the original set A. Specifically, our goal is to prove Lemma 3.1. We begin by presenting a special case of a result by Croot and Sisask [3].

Lemma 5.1. *Consider a set $A \subset \mathbb{F}_2^n$ with $|A| \geq \delta 2^n$, a function $f : \mathbb{F}_2^n \to [0,1]$, $p \geq 1$, and $\varepsilon > 0$. Then there exists $X \subset \mathbb{F}_2^n$ such that $|X| \geq \delta^{O(p/\varepsilon^2)} 2^n$ and for any $x \in X$ we have*

$$\|1_x * 1_A * f - 1_A * f\|_p \leq 2^n \varepsilon. \tag{3}$$

More intuitively, for every bounded function f, if we consider the sum of the translations of f by elements of some large set A, then the resulting function will be nearly invariant to many translations. We prove this lemma by using a probabilistic argument. We first recall *Markov's inequality*.

Lemma 5.2. *If X is a nonnegative random variable and $a > 0$, then* $\Pr[X \geq a] \leq \mathbb{E}[X]/a$.

We will also rely on the following straightforward corollary of the *Marcinkiewicz–Zygmund inequality* [5].

Lemma 5.3. *There exists a constant c such that the following holds for any $p \geq 1$. Let $X_1, \ldots, X_\ell$ be independent real-valued random variables, each with with $|X_j| \leq 1$ and $\mathbb{E}[X_j] = 0$. Then*

$$\mathbb{E}\left[\left|\frac{1}{\ell}\sum_{j=1}^{\ell} X_j\right|^p\right] \leq (cp/\ell)^{p/2}.$$

Proof of Lemma 5.1. For a positive integer ℓ whose value will be determined below, let $a_1, \ldots, a_\ell$ be elements that were uniformly and independently chosen from A. For $1 \leq j \leq \ell$, we set $f_j = \frac{1}{|A|}(1_A * f) - 1_{a_j} * f$ and notice that

$$f_j(x) = \frac{1}{|A|}(1_A * f)(x) - (1_{a_j} * f)(x) = \frac{1}{|A|}\sum_{a \in A} f(x+a) - f(x+a_j).$$

That is, $f_j(x)$ is the difference between the average value of f_j taken over every translation of x by an element of A, and the translation by the specific element a_j.

Since $f : \mathbb{F}_2^n \to [0,1]$, we have $|f_j(x)| \leq 1$. For a fixed $x \in \mathbb{F}_2^n$, we can treat $f_j(x)$ as a random variable with

$$\begin{aligned}\mathbb{E}[f_j(x)] &= \sum_{a_j \in A} \frac{1}{|A|}\left(\frac{1}{|A|}\sum_{a \in A} f(x+a) - f(x+a_j)\right) \\ &= \frac{1}{|A|}\sum_{a \in A} f(x+a) - \frac{1}{|A|}\sum_{a_j \in A} f(x+a_j) = 0.\end{aligned}$$

By the above, for any fixed $x \in \mathbb{F}_2^n$ we may apply Lemma 5.3 to the random variables $f_1(x), \ldots, f_\ell(x)$. That is,

$$\mathbb{E}\left[\left|\frac{1}{\ell}\sum_{j=1}^{\ell} f_1(x)\right|^p\right] \leq (cp/\ell)^{p/2}.$$

Summing this up for every $x \in \mathbb{F}_2^n$ and noting that $\sum_{j=1}^{\ell} f_j = \frac{\ell}{|A|}(1_A * f) - \sum_{j=1}^{\ell} 1_{a_j} * f$ implies

$$\mathbb{E}_{a_1,\ldots,a_\ell}\left[\left\|\frac{1}{|A|}(1_A * f) - \frac{1}{\ell}\sum_{j=1}^{\ell} 1_{a_j} * f\right\|_p^p\right] = \mathbb{E}_{a_1,\ldots,a_\ell}\left[\sum_{x\in\mathbb{F}_2^n}\left|\frac{1}{\ell}\sum_{j=1}^{\ell} f_j(x)\right|^p\right] \leq 2^n(cp/\ell)^{p/2}.$$

We set $\ell = cp \cdot (4/\varepsilon)^{2/p}$. By Markov's inequality (Lemma 5.2) we have

$$\Pr_{a_1,\ldots,a_\ell}\left[\left\|\frac{1}{|A|}(1_A * f) - \frac{1}{\ell}\sum_{j=1}^{\ell} 1_{a_j} * f\right\|_p^p \geq 2^{n-1}\varepsilon\right] \leq \frac{2^n(cp/\ell)^{p/2}}{2^{n-1}\varepsilon} = 1/2.$$

That is, with probability at least $1/2$ we have $\left\|\frac{1}{|A|}(1_A * f) - \frac{1}{\ell}\sum_{j=1}^{\ell} 1_{a_j} * f\right\|_p^p < 2^{n-1}\varepsilon$.

We next define

$$S(A) = \left\{(a_1, \ldots, a_\ell) \in (\mathbb{F}_2^n)^\ell : \left\|\frac{1}{|A|}(1_A * f) - \frac{1}{\ell}\sum_{j=1}^{\ell} 1_{a_j} * f\right\|_p^p < 2^{n-1}\varepsilon\right\}.$$

By the above, we know that at least half of the choices $(a_1, \ldots, a_\ell) \in A^\ell$ satisfy this condition. Thus, $|S(A)| \geq |A|^\ell/2 \geq 2^{n\ell-1}\delta^\ell$. For any $x \in \mathbb{F}_2^n$, the set $S(A+x)$ is a translation by x of the set $S(A)$, so $|S(A+x)| \geq 2^{n\ell-1}\delta^\ell$. By the pigeonhole principle, there must exist an ℓ-tuple $(a_1, \ldots, a_\ell) \in (\mathbb{F}_2^n)^\ell$ that is in $S(A+x)$ for at least $2^n \cdot 2^{n\ell-1}\delta^\ell/2^{n\ell} = 2^{n-1}\delta^\ell$ different elements $x \in \mathbb{F}_2^n$.

For the popular ℓ-tuple $(a_1, \ldots, a_\ell)$ from the previous paragraph, we set $X' = \{x \in \mathbb{F}_2^n : (a_1, \ldots, a_\ell) \in S(A+x)\}$. By the triangle inequality, for any $x, x' \in X'$ we have

$$\begin{aligned}\|1_{A+x} * f - 1_{A+x'} * f\|_p &\leq \left\|1_{A+x} * f - \frac{1}{\ell}\sum_{j=1}^{\ell} 1_{a_j} * f\right\|_p + \left\|\frac{1}{\ell}\sum_{j=1}^{\ell} 1_{a_j} * f - 1_{A+x} * f\right\|_p \\ &\leq 2^{n-1}\varepsilon + 2^{n-1}\varepsilon = 2^n\varepsilon. \qquad (4)\end{aligned}$$

For an arbitrary $x' \in X'$ we set $X = x' + X'$. Notice that $|X| = |X'| \geq 2^{n-1}\delta^{\ell} \geq \delta^{O(p/\varepsilon^2)}2^n$. Moreover, for any $x \in X$ we have $x + x' \in X'$. Thus, by (4) (and recalling that $\|\cdot\|_p$ is not affected by translations) we get

$$\|1_{A+x} * f - 1_A * f\|_p = \|1_{A+x+x'} * f - 1_{A+x'} * f\|_p \leq 2^n\varepsilon.$$

□

Recall that the *inner product* of two functions $f, g : \mathbb{F}_2^n \to \mathbb{R}$ is defined as

$$\langle f, g\rangle = \sum_{x\in\mathbb{F}_2^n} f(x)g(x).$$

We also recall *Hölder's inequality*. Let p and q be two positive real numbers that satisfy $1/p + 1/q = 1$. Then for any two sequences of real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, we have

$$\sum_{j=1}^{n} |a_j b_j| \leq \left(\sum_{j=1}^{n} |a_j|^p\right)^{1/p} \left(\sum_{j=1}^{n} |b_j|^q\right)^{1/q}.$$

Specifically, for p and q as above and functions $f, g : \mathbb{F}_2^n \to \mathbb{R}$, we have

$$|\langle f, g\rangle| \leq \sum_{x\in\mathbb{F}_2^n} |f(x)g(x)| \leq \left(\sum_{x\in\mathbb{F}_2^n} |f(x)|^p\right)^{1/p} \left(\sum_{x\in\mathbb{F}_2^n} |g(x)|^q\right)^{1/q} = \|f\|_p \cdot \|g\|_q. \quad (5)$$

We are now ready to prove Lemma 3.1. We first repeat the statement of the lemma.

Lemma 3.1. *Consider a set $A \subset \mathbb{F}_2^n$ with $|A| \geq \delta 2^n$ and a positive integer $t = O(\lg(1/\delta))$. Then there exists $X \subset \mathbb{F}_2^n$ such that $|X| \geq 2^n\delta^{O(\lg^3(1/\delta))}$ and at least $0.9|A|^2|X|^t$ tuples $(a_1, a_2, x_1, \ldots, x_t) \in A^2 \times X^t$ satisfy $a_1 + a_2 + x_1 + \cdots + x_t \in 2A$.*

Proof. We apply Lemma 5.1 with $f = 1_{2A}$, $p = \lg(1/\delta)$, and $\varepsilon = 1/(20t)$. The lemma implies that there exists a set $X \subset \mathbb{F}_2^n$ such that X satisfies (3) and $|X| \geq \delta^{O(p/\varepsilon^2)}2^n \geq \delta^{O(\lg^3(1/\delta))}2^n$. We consider a tuple $(x_1, \ldots, x_t) \in X^t$, set $x = x_1 + \cdots + x_t$, and denote by N_{2A+x} the number of pairs $(a_1, a_2) \in A^2$ for which $a_1 + a_2 + x \in 2A$. To prove the lemma, we will show that for every such x we have $N_{2A+x} \geq 0.9|A|^2$. For any such x,

notice that

$$\begin{aligned}\langle 1_x * 1_A * 1_{2A}, 1_A\rangle &= \sum_{y\in\mathbb{F}_2^n} 1_x * 1_A * 1_{2A}(y)\cdot 1_A(y) = \sum_{y\in A} 1_x * 1_A * 1_{2A}(y)\\ &= \sum_{y\in A} 1_A * 1_{2A}(x+y) = \sum_{y\in A}\sum_{z\in\mathbb{F}_2^n} 1_A(z)1_{2A}(x+y+z)\\ &= \sum_{y,z\in A} 1_{2A}(x+y+z) = N_{2A+x}.\end{aligned}$$

Thus, it suffices to prove that $\langle 1_x * 1_A * 1_{2A}, 1_A\rangle \geq 0.9|A|^2$ for every t-tuple $(x_1,\ldots,x_t)\in X^t$. To show this, we notice that

$$\langle 1_x * 1_A * 1_{2A}, 1_A\rangle = \langle 1_A * 1_{2A}, 1_A\rangle + \langle 1_x * 1_A * 1_{2A} - 1_A * 1_{2A}, 1_A\rangle. \tag{6}$$

We have

$$\langle 1_A * 1_{2A}, 1_A\rangle = \sum_{y\in\mathbb{F}_2^n} (1_A * 1_{2A}(y))\cdot 1_A(y) = \sum_{y,z\in A} 1_{2A}(y+z) = |A|^2. \tag{7}$$

By (6), it remains to show that $\langle 1_x * 1_A * 1_{2A} - 1_A * 1_{2A}, 1_A\rangle \geq -0.1|A|^2$. Set $q = p/(p-1)$, so that $1/p + 1/q = 1$. By applying Hölder's inequality as in (5), we have

$$\left|\langle 1_x * 1_A * 1_{2A} - 1_A * 1_{2A}, 1_A\rangle\right| \leq \|1_x * 1_A * 1_{2A} - 1_A * 1_{2A}\|_p \|1_A\|_q. \tag{8}$$

Since $p = \lg(1/\delta)$, we obtain

$$\|1_A\|_q = \left(\sum_{x\in\mathbb{F}_2^n} 1_A(x)\right)^{1/q} = |A|^{(p-1)/p} \leq |A|\,(2^n\delta)^{-1/p} = |A|2^{1-n/\lg(1/\delta)}. \tag{9}$$

For a t-tupe $(x_1,\ldots,x_t)\in X^t$ and $x = x_1+\cdots+x_t$, by repeatedly applying the triangle inequality we get

$$\|1_x * 1_A * 1_{2A} - 1_A * 1_{2A}\|_p \leq \sum_{j=1}^{t} \|1_{x_1+\cdots+x_j} * 1_A * 1_{2A} - 1_{x_1+\cdots+x_{j-1}} * 1_A * 1_{2A}\|_p.$$

Since any ℓ_p norm is invariant under translations, we translate the j'th term of the sum by $x_1+\cdots+x_{j-1}$, to obtain

$$\|1_x * 1_A * 1_{2A} - 1_A * 1_{2A}\|_p \leq \sum_{j=1}^{t} \|1_{x_j} * 1_A * 1_{2A} - 1_A * 1_{2A}\|_p.$$

By the definition of X and recalling that $\varepsilon = 1/(20t)$, we obtain

$$\|1_x * 1_A * 1_{2A} - 1_A * 1_{2A}\|_p \leq 2^n \varepsilon t = 2^n/20 \leq |A|/(20\delta) = |A|2^{\lg(1/\delta)}/20.$$

Combining this with (8) and (9) gives

$$\left|\langle 1_x * 1_A * 1_{2A} - 1_A * 1_{2A}, 1_A\rangle\right| \leq |A|2^{1-n/\lg(1/\delta)} \cdot |A|2^{\lg(1/\delta)}/20 = |A|^2 2^{\lg(1/\delta)-n/\lg(1/\delta)}/10.$$

We obviously have $|A|^2 2^{\lg(1/\delta)-n/\lg(1/\delta)}/10 < |A|^2/10$. Combining this with (6) and (7) yields

$$\langle 1_x * 1_A * 1_{2A}, 1_A\rangle > |A|^2 - |A|^2/10 = 0.9|A|^2,$$

which completes the proof of the lemma. □

References

[1] T. F. Bloom, A quantitative improvement for Roth's theorem on arithmetic progressions, arXiv:1405.5800.

[2] M.-C. Chang, A polynomial bound in Freiman's theorem, *Duke mathematical journal* **113** (2002), 399–420.

[3] E. Croot and O. Sisask, A probabilistic technique for finding almost-periods of convolutions, *Geometric and functional analysis* **20** (2010), 1367–1396.

[4] S. Lovett, An exposition of Sanders quasi-polynomial Freiman-Ruzsa theorem, *Electronic Colloquium on Computational Complexity* **19** (2012).

[5] J. Marcinkiewicz and A. Zygmund, Quelques théoremes sur les fonctions indépendantes, *Studia Mathematica* **7** (1938): 104–120.

[6] T. Sanders, On Roths theorem on progressions, *Ann. of Math.* **174** (2011), 619–636.

[7] T. Sanders, On the Bogolyubov–Ruzsa lemma, *Analysis & PDE* **5** (2012), 627–655.