Additive Combinatorics

Summer School, Catalina Island *

August 10th - August 15th 2008

Organizers:

Ciprian Demeter, IAS and Indiana University, Bloomington

Christoph Thiele, University of California, Los Angeles

*supported by NSF grant DMS 0701302

Contents

1 On the Erdös-Volkmann and Katz-Tao Ring Conjectures **5**
Jonas Azzam, UCLA 5
1.1 Introduction 5
1.2 The Ring, Distance, and Furstenburg Conjectures 5
1.3 Main Results 7

2 Quantitative idempotent theorem **10**
Yen Do, UCLA 10
2.1 Introduction 10
2.2 The main argument 11
2.3 Proof of the induction step 12
2.4 Construction of the required Bourgain system 15

3 A Sum-Product Estimate in Finite Fields, and Applications **21**
Jacob Fox, Princeton 21
3.1 Introduction 21
3.2 Preliminaries 23
3.3 Proof outline of Theorem 1 23

4 Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ **26**
S. Zubin Gautam, UCLA 26
4.1 Introduction 26
4.2 Outline of the proof 27
4.3 Part (b) from part (a) 28
4.4 Proof of part (a) 29
4.4.1 A reduction via additive combinatorics 29
4.4.2 Traces and growth 30
4.4.3 A reduction to additive combinatorics 31
4.5 Expander graphs 32
4.6 Recent further progress 33

5 The true complexity of a system of linear equations **35**
Derrick Hart, UCLA 35
5.1 Introduction 35
5.2 Quadratic fourier analysis and initial reductions 37
5.3 Dealing with f_1 39

5.4 Finishing the proof of the main theorem 42

6 On an Argument of Shkredov on Two-Dimensional Corners 43
Vjekoslav Kovač, UCLA 43
6.1 Some history and the main result 43
6.2 Outline of the proof 44
6.3 Some notation 45
6.4 Main ingredients of the proof 46
6.4.1 Generalized von Neumann lemma 46
6.4.2 Density increment lemma 46
6.4.3 Uniformizing a sublattice 47
6.5 Proof of the main theorem 48

7 An inverse theorem for the Gowers $U^3(G)$ norm over the finite field $\mathbb{F}_5^n$ 51
Choongbum Lee, UCLA 51
7.1 Introduction 51
7.2 Preliminaries 51
7.2.1 Notation 52
7.2.2 Gowers uniformity norm $U^d(G)$, $\|\cdot\|_{U^d(G)}$ 52
7.2.3 Local polynomial bias of order d, $\|\cdot\|_{u^d(B)}$ 53
7.3 Main Theorem 55
7.4 Outline of the proof 56
7.5 Application 58

8 New bounds for Szemerédi's Theorem, II: A new bound for $r_4(N)$ 60
Kenneth Maples, UCLA 60
8.1 The Problem 60
8.2 Notation 61
8.3 Strategy and Initial Reductions 61
8.4 Relevant Definitions 63
8.5 Proof Outline 64
8.6 The Gowers U^3-norm and the Quadratic Bohr Sets 65

9 An inverse theorem for the Gowers $U^3(G)$ norm 67
Eyvindur Ari Palsson, Cornell 67
9.1 Introduction 67

9.2 The inverse theorem 71
9.3 Outline of proof for the inverse theorem 72

10 On the Erdös-Volkmann and Katz-Tao ring conjectures 74
Chun-Yen Shen, Indiana 74
10.1 Introduction 74
10.2 Preliminary results 76
10.3 Outline of Proof of Theorem 1 78
10.4 Applications 79
10.4.1 The Falconer distance problem 79
10.4.2 Dimension of sets of Furstenburg type 80

11 Quantitative bounds for Freiman's Theorem 82
Betsy Stovall, UC Berkeley 82
11.1 Introduction 82
11.2 Applications, remaining conjectures 83
11.3 The proof of Theorem 2 84
11.3.1 Reduction to $A \subset \mathbb{Z}_N$. 84
11.3.2 Finding a progression in $2A - 2A$. 85
11.3.3 From $P_0 \subset 2A - 2A$ to $P \supset A$ 86
11.4 Producing a proper progression of small rank. 86

12 Norm Convergence of Multiple Ergodic Averages of Commuting Transformations 88
Zhiren Wang, Princeton University 88
12.1 Introduction 88
12.2 Finitary versions of the main theorem 88
12.2.1 Finite type convergence statement 89
12.2.2 Discretization of the space 89
12.3 Sketch of proof 92
12.3.1 Koopman-von Neumann type decomposition 92
12.3.2 Inductive step 94

1 On the Erdös-Volkmann and Katz-Tao Ring Conjectures

after J. Bourgain [1]
A summary written by Jonas Azzam

Abstract

In this paper, Bourgain solves the long standing conjecture first posed by Erdös and Volkmann on whether or not there exists a Borel subring of the real line of nonintegral dimension. In this summary, we discuss the problem, Bourgain's approach, and outline the preliminaries for the proof of his main result.

1.1 Introduction

The main aim of this paper is to prove the following:

Theorem 1. *A Borel subring of the real line must have dimension 0 or 1.*

This solves a long standing conjecture by Erdös and Volkmann about whether such sets exist with fractional dimension. While this result was proved simultaneously with G. Edgar and C. Miller [2], Bourgain's approach has a wider range of consequences due to the work of Katz and Tao. Here, we will go over their reformulation of the problem, as well as the other closely related conjectures before discussing the preliminaries results to Bourgains main proof.

1.2 The Ring, Distance, and Furstenburg Conjectures

Falconer previously showed that if the ring has dimension greater than 1/2, then its dimension must be 1. This is a corollary of Falconer's theorem that if $A \subseteq \mathbb{R}$ satisfies $\dim A > 1/2$, then

$$D(A) = \{|x - y| : x, y \in A\}$$

has positive Lebesgue measure (let alone dimension 1). With this fact, the proof is short: if A is a ring, then since $D^2(A \times A) = \{|x-y|^2 : x, y \in A \times A\} \subseteq$

A, the square function preserves dimension, and $\dim(A \times A) \geq 2 \dim A > 1$, we have

$$1 \geq \dim(A) \geq \dim D^2(A \times A) \geq D(A \times A) \geq \min\{1, 2 \dim A\} = 1.$$

For more information on dimension geometric techniques employed here, see [4].

Falconer posed a general conjecture that if K is a compact subset of the plane with $\dim K \geq 1$, then $\dim D(K) = 1$. A weaker version of this conjecture is the following:

Conjecture 2. *(Distance Conjecture) There is an absolute constant $c > 0$ such that if $\dim K \geq 1$, then $\dim D(K) \geq \frac{1}{2} + c$.*

So clearly, there is a relationship between the distance and ring conjectures. Katz and Tao explored this relationship in more detail (see [3]). There, they developed discrete analogues of the distance and ring conjectures, in hopes that proving these analogues would lead to a proof of the nondiscrete versions.

Definition 3. *We say $A \subseteq \mathbb{R}^n$ is a $(\delta, \sigma)_n$ set if A is a union of balls of radius δ and*

$$|A \cap B(x, r)| \leq C\delta^{n-\epsilon}(r/\delta)^{\sigma}$$

for any $x \in \mathbb{R}^n$ and $r \in [\delta, 1]$.

This essentially says that A acts like the δ-neighborhood of a σ dimensional set. With this discretized version of a fractal set, Katz and Tao developed a discretized version of the distance conjecture, as well as a discretized version of the Furstenburg conjecture. This latter problem asks whether there is a lower bound $\gamma(\beta)$ for the dimension of sets A such that for any direction $s \in S^1$, there is a line L_s in that direction such that $\dim(A \cap L_s) \geq \beta$.

Conjecture 4. *$\gamma(1/2) \geq 1 + c$ for some constant $c > 0$.*

Finally, Katz and Tao develop the following analog for the ring conjecture[1]:

[1]This actually conjectures that there is no subring of dimension 1/2, however, by replacing the 1/2 with σ, we get the more general conjecture.

Conjecture 5. *(Discretized Ring Conjecture) Let A be a $(\delta, 1/2)_1$ set of measure $\sim \delta^{1/2}$. Then*

$$|A + A| + |A.A| \geq C\delta^{1/2-c}$$

where $c > 0$ is some absolute constant.

As said earlier, these conjectures are all in fact related. In particular, Katz and Tao show:

1. The discritized distance conjecture implies the distance conjecture,
2. the discritized Furstenburg conjecture implies the Furstenberg conjecture, and
3. all three discritized conjectures are equivalent.

1.3 Main Results

In Bourgain's paper, he proves the discretized ring conjecture and further shows that this implies the original distance conjecture, and by the results of Katz and Tao, this gives positive results for the other two conjectures. The two main results are the following:

Theorem 6. *If A is a $(\delta, \sigma)_1$ set, $\sigma \in (0, 1)$, such that $|A| > \delta^{\sigma+\epsilon}$, then*

$$|A + A| + |A.A| > \delta^{\sigma-c}$$

for some absolute constant $c = c(\sigma) > 0$.

Theorem 7. *The discretized ring conjecture (i.e. the previous theorem) implies the ring conjecture.*

He first proves the latter theorem for dimension 2, that is, the discretized ring conjecture implies there does not exist a Borel subring of dimension 2 (which we will demonstrate in lecture) and later shows how one can adapt this to include all dimensions between 0 and 1.

Next, he introduces and develops some preliminary lemmas for the main proof of the discretized ring conjecture.

First, he develops a partition theorem that says for sets A whose sumsets are not too large, then the sets have some regular spacing.

Lemma 8. *Let $A \subseteq \mathbb{R}$ such that*

$$N(A+A,\delta) < KN(A,\delta)$$

and that

$$\sigma \sim K(\delta N(A,\delta))^{K^{-3}}$$

is small. Then there is $q \in \mathbb{N}$ such that $\frac{\sigma}{q} > \delta$ and

$$A \subseteq \bigcup_{a\in\mathbb{Z}} B(\xi_0 + \frac{a}{q}, \frac{\sigma}{q}).$$

Next, he'd like to prove a weaker version of this theorem but with a weaker restriction on the scale σ of the size of the intervals with respect of their spacing of each other. It uses the previous lemma the following basic lemma:

Lemma 9. *Assume*

$$\frac{N(2A,\delta)}{N(A,\delta)} < K.$$

Let $\eta > \delta$ and $A = \bigcup_{j\in S} A_j$ where $A_j = A \cap I_j \neq \emptyset$ and $\{I_j\}$ is a partition of the real line into intervals of length η. Let j_* satisfy

$$N(A_{j^*},\delta) = \max_j N(A_j,\delta)$$

and let

$$S_1 = \{j \in S : N(A_j,\delta) > TK^{-1}N(A_j + A_{j^*},\delta)\}.$$

Then

$$\sum_{j\in S_1^c} N(A_j,\delta) < 4TN(A,\delta)$$

and if $j \in S_1$,

$$\frac{N(2A_j,\delta)}{N(A_j,\delta)} < (\frac{K}{T})^3.$$

Lemma 10 (Main Lemma)**.** *Let $A \subseteq \mathbb{R}$ be a bounded set, $\delta > 0$, and assume*

$$\frac{N(2A,\delta)}{N(A,\delta)} < K \text{ and } \delta N(A,\delta) < \kappa.$$

Then there exist $\sigma, \delta' > 0$ with $\sigma\delta' > \delta$ and $A' \subseteq A$ such that

1. $\sigma < (\log K)^C \kappa^{(\log K)^{-C}}$,

2. *A' is contained in a union of intervals of length $\sigma\delta'$ and spaced apart by at least δ', and*

3. $N(A', \delta) > \frac{N(A,\delta)}{K \log(1/\delta)}$.

So we can't guarantee that the entire set A is evenly spaced apart, but we can find a sufficiently large subset which is spaced by δ', that is to say that it's porous. In the proof of the discretized ring conjecture, one assumes $|A| > \delta^{1/2+\epsilon}$ and $|A + A| + |A.A| < \delta^{1/2-\epsilon}$. According to Katz and Tao, one may assume this implies $|A.A - A.A| < \delta^{1/2-\epsilon}$. This lemma is used in finding a large subset C of A which is sufficiently porous such that on average multiplicative translates xC have small intersection and hence $|x_0C + xC|$ will be large on average, and this will help show that $|A.A - A.A|$ is large for a desired contradiction.

References

[1] J. Bourgain, *On the Erdös-Vokmann and Katz-Tao Ring Conjectures*, GAFA **13** (2003), 334-365.

[2] G. A. Edgar, C. Miller, *Borel subrings of the reals*, Proc. Amer. Math. Soc. **131**, no. 4 (2003), 1121-1129.

[3] N. Katz, T. Tao, *Some connections between Falconer's distance set conjecture, and sets of Furstenberg type*, New York J. Math. **7** (2001), 149-187 (electronic).

[4] P. Mattila, *Geometry of Sets and Measures in Euclidean Space.* Cambridge Press, New York, NY, 1999.

Jonas Azzam
email: **jonasazzam@math.ucla.edu**

2 Quantitative idempotent theorem

after B. Green and T. Sanders [1]
A summary written by Yen Do

Abstract

The classical theorem proved by P. Cohen [2] states that any idempotent measure on locally compact abelian group G lies in the coset ring of $\widehat{G}$, i.e.

$$\widehat{\mu} = \sum_{j=1}^{L} \pm 1_{\gamma_j + \Gamma_j}$$

where the Γ_j are open subgroups of $\widehat{G}$. This theorem however does not give any information about L. We will prove that L can be bounded by $e^{e^{C\|\mu\|^4}}$ and the number of distinct open subgroups Γ_j may be bounded above by $\|\mu\| + \frac{1}{100}$.

2.1 Introduction

Let G be a locally compact abelian group with dual group $\widehat{G}$. Let $M(G)$ be the algebra of bounded regular Borel measures on G. A measure $\mu \in M(G)$ is idempotent iff $\mu * \mu = \mu$, i.e. $\widehat{\mu}$ is a characteristics function on $\widehat{G}$.

Theorem 1 (Cohen's idempotent theorem)**.** *μ is idempotent iff $\{\gamma \in \widehat{G} : \widehat{\mu}(\gamma) = 1\}$ lies in the coset ring of $\widehat{G}$, i.e.*

$$\widehat{\mu} = \sum_{j=1}^{L} \pm 1_{\gamma_j + \Gamma_j} \tag{1}$$

where the Γ_j are open subgroups of $\widehat{G}$.

This was proved by Paul Cohen [2]. Partial results was previously obtained for $G = \mathbb{T}$ by Helson [3] and $G = \mathbb{T}^d$ by Rudin [4]. This theorem however does not give any information about L; for instance, it is trivial when G is finite. We will prove the following quantitative version of Cohen's theorem:

Theorem 2 (Quantitative idempotent theorem)**.** *Suppose that $\mu \in M(G)$ is idempotent. Then in (1) we can bound L by $e^{e^{C\|\mu\|^4}}$ and the number of distinct open subgroups Γ_j by $\|\mu\| + \frac{1}{100}$.*

When G is finite, theorem 2 describes characteristics functions of the finite abelian group $\widehat{G}$. These results, indeed, hold for $\mathbb{Z}$-valued functions on $\widehat{G}$:

Theorem 3. *Suppose that G is a finite abelian group. For any $f : G \to \mathbb{Z}$ with $\|f\|_A := \|\widehat{f}\|_{L^1(\widehat{G})} \leq M$, we may write*

$$f = \sum_{j=1}^{L} \pm 1_{x_j + H_j}$$

where all $x_j \in G$, $H_j \leq G$ subgroups and $L \leq e^{e^{CM^4}}$. The number of distinct subgroups H_j may be bounded above by $M + \frac{1}{100}$

Notes. In theorem 3, G and $\|f\|_A$ play the roles of $\widehat{G}$ and $\|\mu\|$ in the idempotent theorem. The number $\frac{1}{100}$ in the estimates of theorem 2 and 3 can be improved to any given positive δ.

Theorem 2 can be deduced from theorem 3 by a standard argument using finite group approximation. In this exposition, we will prove theorem 3.

2.2 The main argument

The proof of theorem 3 is essentially by inducting on the algebra norm $\|f\|_A$. Ideally we would like to decompose f into $f_1 + f_2$ where f_1 and f_2 are integer-valued functions, with either smaller algebra norm or nice behavior (to be specified below). It turns out to be easier to work with almost integer-valued functions, i.e. functions that take values in $\mathbb{Z} + [-\epsilon, \epsilon]$ for some ϵ small. If f is an ϵ-almost $\mathbb{Z}$-valued function, we define $f_{\mathbb{Z}}$ to be its integer-valued approximation and write $d(f, \mathbb{Z}) \leq \epsilon$.

Lemma 4 (Induction step)**.** *Suppose $f : G \to \mathbb{R}$ has $f_{\mathbb{Z}} \not\equiv 0$, $\|f\|_A \leq M$ for some $M \geq 1$, and $d(f, \mathbb{Z}) \leq \delta$, for some very small δ (says $\delta = e^{e^{-C_1 M^4}}$, C_1 large given). For any $\epsilon = e^{e^{-C_0 M^4}}$, we may write $f = f_1 + f_2$, such that:*

(i) The algebra norm is preserved: $\|f\|_A = \|f_1\|_A + \|f_2\|_A$;

(ii) the components are still almost $\mathbb{Z}$-valued: $d(f_1, \mathbb{Z}) \leq d(f, \mathbb{Z}) + \epsilon$, and (consequently) $d(f_2, \mathbb{Z}) \leq 2d(f, \mathbb{Z}) + \epsilon$;

(iii) f_2 has a significantly smaller algebra norm: $\|f_2\|_A \leq \|f\|_A - 1/2$;

(iv) either f_1 has smaller algebra norm ($\|f_2\|_A \leq \|f\|_A - 1/2$), or $(f_1)_{\mathbb{Z}}$ is a sum of at most $e^{e^{CM^4}}$ functions of the form $\pm 1_{x_i+H}$ (in which case we say it is finished).

Proof of Theorem 3 using Lemma 4. . Let $f : G \to \mathbb{Z}$ be a nonzero function, with $\|f\|_A \leq M$. Let $\epsilon = e^{e^{-C_0 M^4}}$ be a small parameter. If at every step we can apply lemma 4 on all unfinished components of f, the number of steps can not exceed $2M$, because otherwise $\|f\|_A \geq \frac{1}{2}(2M+1) > M$. Since $d(f, \mathbb{Z}) = 0$, estimate (ii) in lemma 4 shows that all the obtained components are δ-almost $\mathbb{Z}$-valued functions, with

$$\delta \leq 2(2(\dots 2d(f, \mathbb{Z}) + \epsilon) \dots) + \epsilon) + \epsilon < 2^{2M}\epsilon$$

Thus, if C_0 was chosen to be very large at the beginning, all intermediate unfinished components of f are admissible candidates for lemma 4. Thus, after at most $2M$ steps, all functions will be finished, i.e. we have a decomposition $f = \sum_{k=1}^{L} f_k$ where:

(a) $L \leq 2^{2M}$;

(b) each $(f_k)_{\mathbb{Z}}$ can be written as a sum of at most $e^{e^{CM^4}}$ functions of the form $\pm 1_{x_{j,k}+H_k}$, where $H_k \leq G$ is a subgroup;

(c) $d(f_k, \mathbb{Z}) \leq 2^{2M}\epsilon$ for all k.

From here it is not hard to see that $f = \sum_{j=1}^{L} (f_k)_{\mathbb{Z}}$. Theorem 3 will be proved if we can show that $L \leq M + \frac{1}{100}$. To see this, just notice that the total algebra norm is preserved and if $(f_k)_{\mathbb{Z}} \not\equiv 0$ for some k, we'll have:

$$\|f_k\|_A \geq \|f_k\|_\infty \geq \|(f_k)_{\mathbb{Z}}\|_\infty - d(f_k, \mathbb{Z}) \geq 1 - 2^{2M}\epsilon \geq \frac{M}{M + \frac{1}{100}}$$

□

2.3 Proof of the induction step

¿From now on, we will always assume that G is finite abelian. The decomposition in lemma 4 will be of the form:

$$f = \psi_S f + (f - \psi_S f)$$

here ψ_S is the Fourier multiplier operator with multiplier $\widehat{\beta}_1$ and β_1 has the form $\frac{1_{X_1}}{|X_1|} * \frac{1_{X_1}}{|X_1|}$ for some $X_1 \subset G$ will be a nice subset of G:

$$\widehat{\psi_S f}(\gamma) = \widehat{f}(\gamma)\widehat{\beta}_1(\gamma)$$

It is not hard to see that $0 \leq \widehat{\beta}_1(\gamma) \leq 1$ for all γ, so the required condition of lemma 4 on algebra norm preservation is automatically satisfied. To meet the other requirements, we will construct X_1 as part of a nice *regular Bourgain system*:

Definition 5 (Bourgain system)**.** *Let G be a finite abelian group and $d \geq 1$ be an integer. A Bourgain system S of dimension d is a collection $(X_\rho)_{\rho\in[0,4]}$ of subsets of G satisfying:*

BS1 *(Nesting) If $\rho' \leq \rho$ we have $X_{\rho'} \subseteq X_\rho$;*

BS2 *(Zero) $0 \in X_0$;*

BS3 *(Symmetry) If $x \in X_\rho$ then $-x \in X_\rho$;*

BS4 *(Addition) For all ρ, ρ' such that $\rho + \rho' \leq 4$ we have $X_\rho + X_{\rho'} \subseteq X_{\rho+\rho'}$;*

BS5 *(Doubling) If $\rho \leq 1$ we have $|X_{2\rho}| \leq 2^d|X_\rho|$.*

For a Bourgain system S, we define $|S| := |X_1|$ to be its *size* and $\mu(S) := |S|/|G|$ to be its *density* in G. Notice that there is a canonical probability measure β_ρ on any $X_{2\rho}$, defined by $\beta_\rho := \frac{1_{X_\rho}}{|X_\rho|} * \frac{1_{X_\rho}}{|X_\rho|}$. Examples of a nontrivial Bourgain system include Bohr systems, and the collection of closed balls at the 0 in a translation-invariant pseudometrics space with doubling properties.

If $\lambda > 0$, we define the dilated of S by $\lambda S := (X_{\lambda\rho})_{\rho\in[0,4]}$. The following simple lemma is useful in the sequel:

Lemma 6 (Dilated Bourgain systems)**.** *For any Bourgain system S and for $\lambda \in (0, 1]$, we have $\dim(\lambda S) = \dim(S)$ and $|\lambda S| \geq (\lambda/2)^d|S|$.*

Definition 7 (Regular Bourgain system)**.** *A Bourgain system $S = (X_\rho)_{\rho\in[0,4]}$ with dimention d is said to be regular if*

$$\left|1 - \frac{|X_1|}{|X_{1+\kappa}|}\right| \leq 10d\kappa$$

whenever $|\kappa| \leq \frac{1}{10d}$.

Remark. Really, this is the definition of a Bourgain system being regular at $\rho = 1$. The number 10 in the definition can be replaced by any positive number, it is chosen only for simplicity of the argument.

It is not hard to show that:

Lemma 8 (Regular dilated Bourgain systems)**.** *For any Bourgain system S, $\exists \lambda \in [1/2, 1]$ such that λS is regular.*

Regular Bourgain systems has the following important property:

Lemma 9 (Almost invariance)**.** *If f is any complex-valued function on G and S is a regular Bourgain system of dimension d on G, then for every $\kappa \in (0, 1)$ and $y \in X_\kappa$ we have:*

$$|\psi_S f(x + y) - \psi_S f(x)| \leq 20 d\kappa \|f\|_\infty$$

Now, to prove the induction step (i.e. lemma 4), we will use the following proposition:

Proposition 10. *Under the same assumptions as in the main lemma (Lemma 4), we can find a Bourgain system S such that:*

(A) Bounded dimension: $d = \dim(S) \leq e^{CM^4}$;

(B) Positive density $\mu(S) \geq e^{e^{-CM^4}} \|f_{\mathbb{Z}}\|_1$;

(C) ψ_S does not kill f: $\|\psi_S f\|_\infty \geq 2e^{-CM^4}$ (to avoid a trivial decomposition of f);

(D) $\psi_S f$ remains almost $\mathbb{Z}$-valued: $d(\psi_S f, \mathbb{Z}) \leq d(f, \mathbb{Z}) + \epsilon$

Proof that such a system gives a desired decomposition. First, recall that $\epsilon = e^{-C_0 M^4}$ was given, and we can choose C_0 to be very large compared to the constant C in property (C). Now, as previously indicated, we define $f_1 := \psi_S f$ and $f_2 := f - \psi_S f$. Property (D) immediately gives us the desired control on $d(f_1, \mathbb{Z})$ and $d(f_2, \mathbb{Z})$.

Now, since C_0 is very large compared to C we will have $\|\psi_S f\|_\infty > d(\psi_S f, \mathbb{Z})$ therefore $(\psi_S f)_{\mathbb{Z}} \not\equiv 0$. This easily implies $\|f_1|_A \equiv \|\psi_S f\|_A \geq 1 - \epsilon > 1/2$ and gives the desired bound for $\|f_2\|_A$.

Finally, assume $\|f_1\|_A > \|f\|_A - 1/2$ and we will show that f_1 is finished by showing $(f_1)_{\mathbb{Z}} \equiv f_{\mathbb{Z}}$ and is constant on $H :=< X_\rho >$ for $\rho = \frac{\epsilon}{20dM}$. If this is done, we clearly can write:

$$(f_1)_{\mathbb{Z}} = f_{\mathbb{Z}} = \sum_{j=1}^{L} \pm 1_{x_j+H}$$

where we may take

$$L \leq \frac{M\|f_{\mathbb{Z}}\|}{\|1_H\|_1}$$

and the desired bound $L \leq e^{e^{C''M^4}}$ can be obtain by observing that:

$$\|1_H\|_1 \geq \frac{|X_\rho|}{|G|} \geq (\rho/2)^d \mu(S)$$

To prove the above claims, first notice that $\|f_2\|_A = \|f\|_A - \|f_1\|_A < 1/2$, this easily implies $(f_2)_{\mathbb{Z}} \equiv 0$ and so $f_{\mathbb{Z}} = (\psi_S f)_{\mathbb{Z}}$. Now using the almost invariance of regular Bourgain system and the bound (A) on $\dim(S)$, for any $x - x' \in X_\rho$ we easily have:

$$|(\psi_S f)_{\mathbb{Z}}(x) - (\psi_S f)_{\mathbb{Z}}(x')| \leq 2d(f, \mathbb{Z}) + 3\epsilon < 1$$

□

2.4 Construction of the required Bourgain system

In this section we sketch the main ideas used prove proposition 10. The construction is done incrementally. It is not hard to construct a system satisfying given bounded dimension and given positive density; it is however harder to meet (C). After a Bourgain system with (A,B,C) have been constructed, an averaging argument will be used to refine it to a system with additional property (D):

Theorem 11. *Suppose that $f : G \to M$ satisfies $\|f\|_A \leq M$ for some $M \geq 1$, and also assume that $d(f, \mathbb{Z}) < 1/4$. For any positive $\epsilon \leq 1/4$ and any Bourgain system S with dimension $d \geq 2$, there is a regular Bourgain system S' with*

(i) $\dim(S') \leq 4d + \frac{64M^2}{\epsilon^2}$;

(ii) $|S'| \geq e^{-\frac{CdM^4}{\epsilon^4}\log(dM/\epsilon)}|S|$;

(iii) $\|\psi_{S'}f\|_\infty \geq \|\psi_S f\| - \epsilon$;

(iv) $d(\psi_{S'}f, \mathbb{Z}) \leq d(f, \mathbb{Z}) + \epsilon$.

Sketch of proof. We'll construct a nested sequence of Bourgain systems $S^{(0)} = S, S^{(1)}, S^{(2)}, \ldots$ until a system satisfying all (i-iv) is found. In this sequence, for every $S^{(j)}$ not having property (iv) (this is true for everyone, except for the last one - if the sequence terminates), there is a subset $\Gamma^{(j)} \subset \widehat{G}$ such that

$$\sum_{\gamma\in\Gamma^{(j)}} |\widehat{f}(\gamma)| > \frac{\epsilon^2}{16M} \tag{2}$$

These $\Gamma^{(j)}$ will be disjointly constructed, and this immediately implies that the sequence must stop at some $j = J \leq \frac{16M^2}{\epsilon^2}$. Now the bound on j can be used to show that every member $S^{(j)}$ of the sequence (including the last one, which we will take as S') will satisfy (i,ii,iii). The construction of $\Gamma^{(j)}$ and $S^{(j+1)}$ out of an $S^{(j)}$ without property (iv) is done through the following steps:

Step 1: If property (iv) failed for S then for any $x_0 \in G$ and any ρ sufficiently small (says, $\rho \leq \frac{\epsilon}{80d'M}$) that makes $\rho S'$ regular, we can use the almost invariance property (lemma 9) to show:

$$E_{x\in G}(f - \psi_{S'})(x)^2\beta'_\rho(x - x_0) > \epsilon^2/4$$

Step 2: By lemma 6 on dilated Bourgain systems, we can take $\rho \sim \frac{\epsilon}{d'M}$ in the previous step. Plancherel's theorem then implies the existence of $\gamma_0^{(j+1)} \in \widehat{G}$ such that

$$|\sum_{\gamma\in\widehat{G}} \widehat{f}(\gamma)(1 - \beta_1^{(j)}(\gamma))\widehat{\beta}_\rho^{(j)}(\gamma_0^{(j+1)} - \gamma)| > \frac{\epsilon^2}{8M}$$

Now, (2) is satisfied if we define $\Gamma^{(j)}$ by truncating from $\widehat{G}$ the tails in the above sum, i.e. avoiding places where $1 - \beta_1^{(j)}(\gamma)$ and $\widehat{\beta}_\rho^{(j)}(\gamma_0^{(j+1)} - \gamma)$ are too small:

$$\begin{aligned}\Gamma^{(j)} &:= \{1-\beta_1^{(j)}(\gamma) > \frac{\epsilon^2}{32M^2}\} \cap \{|\widehat{\beta}_\rho^{(j)}(\gamma_0^{(j+1)}-\gamma)| > \frac{\epsilon^2}{64M^2}\} \\ &\equiv \Big(\gamma_0^{(j+1)} + \textsc{Spec}_{\epsilon^2/64M^2}(\beta_\rho^{(j)})\Big) \setminus \textsc{Spec}_{1-\epsilon^2/32M^2}(\beta_1^{(j)})\end{aligned}$$

Step 3: The disjointness of $\Gamma^{(j)}$'s will be ensured if we choose a regular Bourgain system $S^{(j+1)}$ where

$$\gamma_0^{(j+1)} + \textsc{Spec}_{\epsilon^2/64M^2}(\beta_\rho^{(j)}) \subset \textsc{Spec}_{1-\epsilon^2/32M^2}(\beta_1^{(j+1)})$$

and for every $k \leq j$,

$$\textsc{Spec}_{1-\epsilon^2/32M^2}(\beta_1^{(k)}) \subset \textsc{Spec}_{1-\epsilon^2/32M^2}(\beta_1^{(j+1)})$$

This could be done by joining a suitable dilated of $S^{(j)}$ and a small Bohr system of the character $\gamma_0^{(j+1)}$ (needed to control the behavior of $\gamma_0^{(j+1)}$). The resulting system might need to be dilated by some $\lambda \in [1/2, 1]$ to ensure regularity. More explicitly, we can choose:

$$S^{(j+1)} := \lambda\Big(\kappa\rho S^{(j)} \wedge \textsc{Bohr}_{\kappa'}(\{\lambda_0^{(j+1)}\})\Big)$$

where $\kappa \sim \epsilon^4/d_j M^4, \kappa' \sim \epsilon^2/M^2$.

Step 4: Standard bounds on dimension and size of Bohr systems then easily give the required bounds (i,ii) on the size and dimension for $S^{(j+1)}$. Also, regularity and the nesting property of $S^{(j)}$ gives

$$|\psi_S f - \psi_{S'} f| = |f * (\beta_1 * \beta_1' - \beta_1)|$$

which is small if β_1' is supported on a very small neighbourhood of 0. The latter is however guaranteed by the construction of the sequence $S^{(j)}$ (notice that if the sequence actually stops at $S^{(0)} = S$ then we wouldn't have this small support property, but clearly if this was the case nothing need to be proved). □

We now sketch the main steps in constructing a Bourgain system with properties (A,B,C) from given function f satisfying the condition of lemma 4. This is essentially based on the following result:

Proposition 12 (Weak Freiman)**.** *If $A \subset G$ is a nonempty finite subset of a finite abelian group G, with $|A+A| \leq K|A|$ then there is a regular Bourgain system $S = (X_\rho)_{\rho \in [0,4]}$ such that*

$$\dim(S) \leq CK^C, \ |S| \geq e^{-CK^C}|A|, \textit{ and } \|\psi_S 1_A\|_\infty \geq cK^{-C}$$

When A is a large subset of G, a Bogolyubov-Chang's argument [6, 7] can be used to proved the above proposition; furthermore in that case it is possible to show that $X_4 \subset 2A - 2A$. The general case can then be reduced to this special case by applying an analogue of Freiman's theorem for arbitrary abelian groups [5].

To continue, we need the notion of *arithmetic connectedness*. First, we say that a set $\{a_1, \ldots, a_k\}$ is *dissociated* if the only solution to $\epsilon_1 a_1 + \cdots + \epsilon_k a_k = 0$ with $\epsilon_k \in \{0, 1, -1\}$ is the trivial solution.

Definition 13 (Arithmetic connectedness)**.** *A subset $A \subset G$ with $0 \notin A$ is called m-arithmetically connected for some $m \geq 1$ if for any subset A' of A with size $|A'| = m$ we will have either:*

(i) A' is not dissociated, or

(ii) A' is dissociated, and there is some $x \in A \setminus A'$ with $x \in \langle A \rangle$.

Essentially, if A is an m-arithmetically connected then for any distinct $x_0, x_1, \ldots, x_m$ in A there is a nontrivial linear relation $\epsilon_0 x_0 + \cdots + \epsilon_m x_m = 0$. Using an averaging argument it is not hard to show that any m-arithmetically connected set A has at least $e^{-CM}|A|^3$ additive quadruples $a_1 + a_2 = a_3 + a_4$. This results can be combined with a previous result by Gowers [8, Proposition 12] and the weak Freiman theorem (proposition 12) to deduce:

Proposition 14. *If A is m-arithmetically connected nonempty subset of $G \setminus \{0\}$ for some $m \geq 1$, then there is a regular Bourgain system S satisfying:*

$$\dim(S) \leq e^{Cm}$$
$$|S| \geq e^{e^{-Cm}}|A|$$
$$\|\psi_S 1_A\|_\infty \geq e^{-Cm}$$

Proof of proposition 10. Take m comparable to M^4, say $m = [50M^4] > 4$. First consider the simpler case when $f \geq 0$. Define:

$$A := \mathrm{Supp}(f_{\mathbb{Z}})$$

It is sufficient to prove that A is m-arithmetically connected. If $A = G$ this is trivial, otherwise by translating f if necessary we can assume $0 \notin A$. Suppose. towards a contradiction, that A is not m-arithmetically connected. Then we can find dissociated $a_1, \ldots, a_m \in A$ such that $\langle a_1, \ldots, a_m \rangle \cap A = \{a_1, \ldots, a_m\}$. Consider the function $p(x)$ defined through a normalized Riesz product:

$$\widehat{p}(\gamma) := \frac{1}{|\widehat{G}|} \prod_{i=1}^{m} (1 + \frac{1}{2}(\gamma(a_i) + \overline{\gamma}(a_i)))$$

It is then not hard to see that $\|f_{\mathbb{Z}} p\|_A \leq 2M^2$, and the function p is real, nonnegative and supported on $\langle a_1, \ldots, a_m \rangle$, and,

$$p(a_i) = \sum_{\overrightarrow{\epsilon}: \epsilon_1 a_1 + \cdots + \epsilon_m a_m = a_i} 2^{-\sum_j |\epsilon_j|} \geq \frac{1}{2}$$

Now as $\langle a_1, \ldots, a_m \rangle \cap A = \{a_1, \ldots, a_m\}$ we have $f_{\mathbb{Z}} p(x) = \sum_{i=1}^{m} f_{\mathbb{Z}}(a_i) p(a_i) 1_{a_i}(x)$. This along with the dissociated property of $a_1, \ldots, a_m$ gives

$$\|\widehat{f_{\mathbb{Z}} p}\|_2^2 \geq \frac{m}{4|G|}$$

$$\|\widehat{f_{\mathbb{Z}} p}\|_4^4 \leq \frac{3}{|G|} \|\widehat{f_{\mathbb{Z}} p}\|_2^4$$

An application of Holder inequality then easily gives $\|f_{\mathbb{Z}} p\|_A \geq \sqrt{\frac{m}{12}}$, contradicts previous upper bound $\|f_{\mathbb{Z}} p\|_A \leq 2M^2$.

For general f, we consider $g = f^2$ and get the system S' for g. An argument similar to the proof of theorem 11 can be used to refine this system to obtain the desired system S for f. □

References

[1] B. J. Green and T. Sanders, *A quantitative version of the idempotent theorem in harmonic analysis.* to appear in Annals of Math.

[2] P.J. Cohen, *On a conjecture of Littlewood and Idempotent measures.* Amer. J. Math. **82**, no. 2, (1960), 191–212.

[3] H. Helson, *Note on harmonic functions.* Proc. Amer. Math. Soc. **4** (1953), 686-691.

[4] W. Rudin, *Idempotent measures on abelian groups.* Pacific J. Math **9** (1959) 195–209.

[5] B. J. Green and I. Z. Ruzsa, *Freiman's theorem in an arbitrary abelian group.* Jour. London Math. Soc. **75** (2007), no. 1, 163–175.

[6] M.-C. Chang, *A polynomial bound in Freiman's theorem.* Duke Math. J. **113** (2002), no. 3, 399–419.

[7] Bogolyubov, *Sur quelques propriétés des presque-périodes.* Ann. Chaire Math. Phys. Kiev **4** (1939), 185–194.

[8] W. T. Gowers, *A new proof of Szemerédi's theorem for arithemtic progressions of length four.* Geom. Funct. Anal. **8** (1998), 529–551.

Yen Do, UCLA
email: **qdo@math.ucla.edu**

3 A Sum-Product Estimate in Finite Fields, and Applications

after J. Bourgain, N. Katz, and T. Tao [1]
A summary written by Jacob Fox

Abstract

We prove that if $A \subset F =: \mathbb{Z}/q\mathbb{Z}$ with q prime and $|A| \leq |F|^{1-\delta}$, then $|A+A| + |A \cdot A| \geq c(\delta)|A|^{1+\epsilon}$. This sum-product estimate has applications to several problems for finite fields, including a Szemeredi-Trotter type theorem, estimates for the Erdős distance problem, and to the three-dimensional Kakeya problem.

3.1 Introduction

Let A be a non-empty subset of a finite field F. Let

$$A + A = \{a + b : a, b, \in A\}$$

and

$$A \cdot A = \{a \cdot b : a, b \in A\}.$$

It is clear that $|A+A|, |A \cdot A| \geq |A|$, and these bounds are sharp if A is a subfield of F. If $F = \mathbb{Z}/q\mathbb{Z}$ with q prime, so F has no proper subfields, and $|A| \ll |F|$, then one might expect to gain on this inequality. If A is an arithmetic progression, then $|A+A| = 2|A| - 1$ or $A + A = \mathbb{F}$. If A is a geometric progression, then $|A \cdot A| = 2|A| - 1$ or $A \cdot A = \mathbb{F}^*$, the set of invertible elements of F. It is natural to suspect that a subset A can not act both like an arithmetic progression and a geometric progression. This idea is captured in the main theorem.

Theorem 1. *Let $F = \mathbb{Z}/q\mathbb{Z}$ with q prime, and let A be a subset of F such that $|A| < |F|^{1-\delta}$ for some $\delta > 0$. Then*

$$\max\left(|A+A|, |A \cdot A|\right) \geq c(\delta)|A|^{1+\epsilon}$$

for some $\epsilon = \epsilon(\delta) > 0$.

Bourgain, Katz, and Tao prove this theorem with the extra assumption that $|A| > |F|^{\delta}$. This assumption was removed by Bourgain and Konyagin

[2]. We will follow most closely the presentation in Green's lecture notes [6]. The best bound on ϵ is given by Katz and Shen [7], improving on an earlier estimate of Garaev [5].

An integer version of Theorem 1 was proved by Erdős and Szemerédi [4] much earlier. They showed $\max(|A+A|, |A \cdot A|) \geq |A|^{1+\epsilon}$ for any set A of integers. The bound on ϵ has been improved several times, the most recent by Solymosi [8]. An old conjecture of Erdős, which is still open and considered quite difficult, says that for each $\epsilon > 0$, $\max(|A+A|, |A \cdot A|) \geq c(\epsilon)|A|^{2-\epsilon}$. An analogous conjecture for finite fields, which is at least as difficult, says $\max(|A+A|, |A \cdot A|) \geq c(\epsilon) \min(|A|^{2-\epsilon}, |F|^{1-\epsilon})$.

We next discuss applications of Theorem 1 to three combinatorial geometry problems over finite fields. The following theorem is the first of these results.

Theorem 2. *If one has N lines and N points in the finite plane $(\mathbb{Z}/q\mathbb{Z})^2$ with $N \ll q^2$, then there are at most $O(N^{3/2-\epsilon})$ incidences.*

Using the fact that no two lines are incident to the same two points, which means the point-line incidence graph is has no cycle of length four, it is easy to show the number of incidences between N points and N lines is at most $O(N^{3/2})$. Note that the above theorem does better. Theorem 2 is an analogue of a result of Szemerédi and Trotter [9] that N points and N lines in the plane $\mathbb{R}^2$ have $O(N^{4/3})$ incidences, which is tight.

The second application, which uses Theorem 2, is to the Erdős distinct distance problem. The problem asks: how many distinct distances are among any N points in the finite plane $(\mathbb{Z}/q\mathbb{Z})^2$? A standard argument using Theorem 2 demonstrates that any N points with $N \ll q^2$ determine $N^{1/2+\epsilon}$ distinct distances. The original problem of Erdős was in the plane $\mathbb{R}^2$.

The last application is to the finite field Kakeya problem. A Besicovitch set in d-dimensions is a subset of $(\mathbb{Z}/q\mathbb{Z})^d$ that contains a line in each direction. The finite field Kakeya problem asks how small can a Besicovitch set be? The finite field Kakeya conjecture states that every Besicovitch set has at least cq^d elements for some constant $c = c(d) > 0$. Wolff showed when $d = 3$ that every Besicovitch set has cardinality at least $cq^{5/2}$. Using Theorem 1, this bound can be improved to $cq^{5/2+\epsilon}$. The bound for the finite field Kakeya problem was improved by Dvir [3] using completely different methods (algebraic). Shortly after Dvir first posted his paper, Alon and Tao observed how Dvir's proof can be easily modified to resolve the finite field Kakeya conjecture.

3.2 Preliminaries

We list here two lemmas, the Plünnecke-Ruzsa inequality and the Balog-Szemerédi-Gowers lemma, that we use for the proof of Theorem 1. These two lemmas are important tools for many problems in additive combinatorics.

For a positive integer k, let kA denote the k-fold sumset $A + A + \cdots + A$ where the number of A's is k. We will also use the difference set

$$A - B := \{a - b : a \in A, b \in B\}$$

and the quotient set

$$A/B = \{a/b : a \in A, b \in B, b \neq 0\}.$$

The first lemma we use is the Plünnecke-Ruzsa inequality.

Lemma 3. *Let A and B be finite non-empty subsets of a finite field F with $|A+B| \leq K|A|$. Then for all positive integers k and ℓ, $|kA - \ell A| \leq K^{k+\ell}|A|$.*

The next lemma we recall is the versatile Balog-Szemerédi-Gowers lemma.

Lemma 4. *Let A, B be finite subsets of an additive group with $|A| = |B|$, and let G be a subset of $A \times B$ with $|G| \geq |A||B|/K$ and $|\{a + b : (a, b) \in G\}| \leq K|A|$. Then there exist subsets A', B' of A and B respectively with $|A'| \geq cK^{-C}|A|$ and $|B'| \geq cK^{-C}|B|$ such that $|A' - B'| \leq CK^C|A|$.*

3.3 Proof outline of Theorem 1

The next lemma is an important step toward the proof of the sum-product estimate. It says roughly that if neither $A + A$ nor $A \cdot A$ grows, then there is a large subset $A' \subset A$ such that $A' \cdot A' - A' \cdot A'$ does not grow. The proof of this lemma uses the Balog-Szemerédi-Gowers lemma twice, once for the multiplicative group F^* and once with F as an additive group.

Lemma 5. *Let A be a non-empty subset of F such that $|A+A|, |A \cdot A| \leq K|A|$. Then there is a subset A' of A with $|A'| \geq cK^{-C}|A|$ such that*

$$|A' \cdot A' - A' \cdot A'| \leq CK^C|A|.$$

The next lemma demonstrates that if $A \cdot A - A \cdot A$ does not grow much, then neither does any polynomial expression of A.

Lemma 6. *Let A be a non-empty subset of F such that $|A \cdot A - A \cdot A| \leq K|A|$ for some $K \geq 1$. Then for any polynomial P of several variables and integer coefficients, we have*

$$|P(A, A, \ldots, A)| \leq CK^C|A|$$

where the constant C depends on P.

The previous lemma together with the multiplicative version of the Plünnecke-Ruzsa inequality demonstrates the following extension. Indeed, if P_1 and P_2 are polynomials, and $Q = P_1/P_2$ is a rational function, then Lemma 6 implies that, if if $A \cdot A - A \cdot A$ does not grow, then $P_1 \cdot P_2$, also a polynomial, does not grow and hence $Q = P_1/P_2$ does not grow as well.

Lemma 7. *Let A be a non-empty subset of F such that $|A \cdot A - A \cdot A| \leq K|A|$ for some $K \geq 1$. Then for any rational function Q of several variables and integer coefficients, we have*

$$|Q(A, A, \ldots, A)| \leq CK^C|A|,$$

where the constant C depends on Q.

We now sketch the idea of the proof of Theorem 1. We may suppose for contradiction that $\max(|A + A|, |A \cdot A|) \leq K|A|$ with $K = |A|^\epsilon$, i.e, neither $A + A$ nor $A \cdot A$ grows. By Lemma 5, there is a subset $A' \subset A$ with $|A'| \geq cK^{-C}|A|$ such that $|A' \cdot A' - A' \cdot A'| \leq CK^C|A|$, i.e., A' is a large subset of A such that $A' \cdot A' - A' \cdot A'$ does not grow. By Lemma 7, any rational expression $Q(A', A', \ldots, A')$ does not grow, i.e., there is a constant C_1 depending only on Q such that $Q(A', A', \ldots, A') \leq C_1 K^{C_1}|A'|$. But the following proposition gives a rational function J such that $J(A')$ or $A' - A'$ always grows, a contradiction.

Lemma 8. *Let $J(A) = \{a_5 \left(\frac{a_1a_2 - a_3a_4}{a_1 - a_3} + a_6 : a_1, \ldots, a_6 \in A, a_1 \neq a_3\right)\}$. Then*

1. *If $|A| \geq \sqrt{q}$, then $|Q(A)| \geq q/2$.*
2. *If $|A| < \sqrt{q}$, then $|Q(A)| \geq \frac{|A|^3}{2|A-A|}$.*

References

[1] Bourgain J., Katz N., and Tao T., *A sum-product estimate in finite fields, and applications.* Geom. Funct. Anal. **14** (2004), no. 1, 27–57.

[2] Bourgain J., and Konyagin S., *Estimates for the number of sums and products and for exponential sums over subgroups in fields of prime order.* C.R. Acad. Sci. Paris, Ser. I **337** (2003) 75–80.

[3] Dvir Z., *On the size of Kakeya sets in finite fields.* arXiv:0803.2336v3 [math.CO]

[4] Erdős P., and Szemerédi E., *On sums and products of integers.* in: Studies in Pure Mathematics (Birkhauser, Basel, 1983) 213–218.

[5] Garaev M. Z., *An explicit sum-product estimate in* $\mathbb{F}_p$. arXiv:math/0702780v1 [math.NT]

[6] Green B., *Sum-product estimates*, lecture notes.

[7] Katz N., and Shen C.-Y., *A Slight Improvement to Garaev's Sum Product Estimate.* arXiv:math/0703614v1 [math.NT]

[8] Solymosi J., *An upper bound on the multiplicative energy.* arXiv:0806.1040v3 [math.CO]

[9] Szemerédi E., and Trotter W. T., *Extremal problems in discrete geometry.* Combinatorica **3** (1983), 381–392.

Jacob Fox, Princeton
email: **jacobfox@math.princeton.edu**

4 Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$

after H. A. Helfgott [He1]
A summary written by S. Zubin Gautam

Abstract

We summarize the proof in [He1] that, for any generating set A in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, the diameter of the associated Cayley graph is $O\big((\log p)^{O(1)}\big)$; we also discuss auxiliary results related to expander graphs.

4.1 Introduction

Given a finite group G and a generating set $A \subset G$, the *Cayley graph* of G with respect to A is the graph $\mathcal{G}(G, A)$ with vertex set G and edge set $\{(g, ag) \mid g \in G,\ a \in A \cup A^{-1}\}$. Thus the diameter of $\mathcal{G}(G, A)$ is the maximum length of a word with letters in $A \cup A^{-1}$ needed to produce an arbitary element of G.

An easy counting argument shows that $\mathrm{diam}(\mathcal{G}(G, A)) \geq \log |G| / \log |A|$, and in fact the diameter can be much larger in general for G abelian. However, in the setting of non-abelian simple groups, Babai conjectured that the diameter should not in fact be much larger than the trivial lower bound:

Conjecture 1 (Babai)**.** *For any finite, simple, non-abelian group G and any generating set $A \subset G$,*

$$\mathrm{diam}\big(\mathcal{G}(G, A)\big) \lesssim (\log |G|)^c$$

for some absolute constant c.[2]

For further background and related discussions, see *e.g.* [BHKLS]. At present, Babai's conjecture in full generality remains open; our goal is to verify it for the special case $G = \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ for p a prime.[3]

[2]Here and in the sequel, we write $X \lesssim Y$ or $X = O(Y)$ to mean $X \leq cY$ for some constant c; subscripts on the symbols "$\lesssim$" or "O" denote dependence of the constant c, and in the absence of such subscripts c is understood to be absolute.

[3]Of course, $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ is not simple, but verifying the conjecture for this choice of G is trivially equivalent to treating the case of $G = \mathrm{PSL}_2(\mathbb{Z}/p\mathbb{Z})$, which is simple for $p > 3$.

4.2 Outline of the proof

To prove Babai's conjecture for $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, we will show that a generic subset A of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ must grow rapidly when acting on itself repeatedly. More precisely, the central result of our discussion is the following:

Key Proposition. *Let p be a prime and A a subset of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ not contained in any proper subgroup.*

(a) *Suppose $|A| < p^{3-\delta}$ for some fixed $\delta > 0$. Then*

$$|A \cdot A \cdot A| \gtrsim_\delta |A|^{1+\varepsilon},$$

where ε depends only on δ.

(b) *Suppose $|A| > p^\delta$ for some fixed $\delta > 0$. Then*

$$\operatorname{diam}\big(\mathcal{G}(\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z}), A)\big) \lesssim_\delta 1.$$

In the setting of this proposition, fix $\delta = 1$, for instance. Then for any p and any generating set $A \subset \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, part (a) gives that one need only multiply A by itself at most $O((\log p)^c)$ times to obtain a set $\tilde{A}$ with $|\tilde{A}| > p$, where c is a universal constant. But then part (b) of the proposition yields

$$\operatorname{diam}\big(\mathcal{G}(\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z}), A)\big) \lesssim (\log p)^c \big(\mathcal{G}(\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z}), \tilde{A})\big) \lesssim (\log p)^c,$$

which resolves the conjecture for $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$.

Our main task is thus to prove the Key Proposition. Part (b) of the proposition actually relies on part (a); part (a) allows us to assume that A is sufficiently large, from which point part (b) can be proven by exploiting structural properties of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ in combination with Fourier-analytic sum-product estimates for "large" sets in the field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.[4]

Part (a) itself is also proved via a reduction to additive combinatorics in finite fields; this time, the transfer is to $\mathbb{F}_{p^2}$ and is achieved via the trace on SL_2. The starting point is the observation that if $A \subset \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ does not grow rapidly under repeated action on itself, then A is "highly commutative"

[4] A quicker and slightly stronger proof of part (b) from part (a) has been given by Nikolov and Pyber ([NP]); their approach works directly with the representation theory of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ in lieu of reducing matters to Fourier analysis on $\mathbb{Z}/p\mathbb{Z}$. ¿From the perspective of pushing results beyond SL_2, this is perhaps a more apt approach; see [He2].

in a suitable sense and thus contains a large simultaneously diagonalizable subset. This fact is used to show that, broadly speaking, one may consider the size of *traces* of subsets rather than subsets of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ themselves. From this point, we are reduced to solving a problem in $\mathbb{F}_{p^2}$; the salient tools used to finish the proof are the Balog–Szemerédi–Gowers regularity theorem, standard sum-product estimates for "small" sets in finite fields, and Ruzsa distances.

4.3 Part (b) from part (a)

In the sequel, for $A \subset \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, A_k denotes the set of products of at most k elements of $A \cup A^{-1}$:

$$A_k = \{a_1 \dots a_\ell \,|\, a_j \in A \cup A^{-1},\, \ell \leq k\}.$$

So we want to show that if $|A| > p^\delta$, there exists some k depending only on δ such that $A_k = \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. We can decompose SL_2 as $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z}) = LULU$, where L is the subgroup of lower-triangular unipotent elements and U is the subgroup of upper-triangular unipotents; thus, it suffices to show that $L \cup U \subset A_k$ for some k depending only on δ. Additionally, part (a) of the Key Proposition affords us the luxury of strengthening our assumption to $|A| > c_1 p^{c_2}$ for any suitable choice of universal constants c_j. Via an application of the pigeonhole principle, one can show that under such an assumption AA^{-1} contains many upper-triangular matrices and many lower-triangular matrices. We have thus reduced part (b) of the Key Proposition to a slightly weaker analogous statement in the Borel subgroups[5] of SL_2:

Lemma 2. *Let p be a prime, H a Borel subgroup of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, and $A \subset H$ with $|A| > 2p^{5/3} + 1$. Then A_8 contains all trace-2 elements of H (*i.e.*, all unipotent elements of H).*

Proof sketch. Without loss of generality, we take H to be the subgroup of upper-triangular matrices in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. For $C \subset H$ and $r \in (\mathbb{Z}/p\mathbb{Z})^*$, define $P_r(C) := \left\{ x \in \mathbb{Z}/p\mathbb{Z} \;\middle|\; \begin{pmatrix} r & x \\ 0 & r^{-1} \end{pmatrix} \in C \right\}$; we want to show $|P_1(A_8)| = p$.

[5] In the present setting, the reader may take this to mean the subgroups of lower- and upper-triangular matrices; these are not all of the Borel subgroups, but they are all we will need.

By the pigeonhole principle, there exists $r \in (\mathbb{Z}/p\mathbb{Z})^*$ for which $|P_r(A)| > 2p^{2/3}$; moreover, some computation and counting allows us to conclude that $|P_1(A_4)| \geq |P_r(A) - sP_r(A)|$ for all s lying in some large set $S \subset (\mathbb{Z}/p\mathbb{Z})^*$ with $|S| > \frac{1}{2}p^{2/3}$.

The following sum-product-type estimate then gives $|P_1(A_4)| \geq \frac{2}{3}p$:

Lemma 3. *Let p be a prime, $A \subset \mathbb{F}_p$, and $S \subset \mathbb{F}_p^*$. Then there exists $s \in S$ such that*

$$|A + sA| \geq \left(\frac{1}{p} + \frac{p}{|S||A|^2}\right)^{-1}.$$

Thus, by the pigeonhole principle, $|P_1(A_8)| = |P_1(A_4) + P_1(A_4)| = p$, as desired. □

Lemma 3 is proved via basic Fourier analysis, driven by the observation that, by Hölder,

$$\|\chi_A * \chi_{sA}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} \leq |A + sA| \cdot \|\chi_A * \chi_{sA}\|_{L^2(\mathbb{Z}/p\mathbb{Z})}.$$

4.4 Proof of part (a)

4.4.1 A reduction via additive combinatorics

As a preliminary observation, we note that there is nothing special about the threefold product in part (a) of the Key Proposition:

Lemma 4. *For an integer $n > 2$ and A a finite subset of a group G, suppose there are $c, \varepsilon > 0$ such that $|A_n| > c|A|^{1+\varepsilon}$. Then there exist constants $c', \varepsilon' > 0$ depending only on c, ε, and n such that $|A \cdot A \cdot A| > c'|A|^{1+\varepsilon'}$.*

The key idea of the proof (which we omit) is to use the triangle inequality for the Ruzsa distance. For a group G, the *Ruzsa distance* between two finite subsets $A, B \subset G$ is

$$d(A, B) = \log\left(\frac{|AB^{-1}|}{\sqrt{|A||B|}}\right).$$

The Ruzsa distance, though not actually a metric, satisfies the triangle inequality, and thus we have $|AC^{-1}||B| \leq |AB^{-1}||BC^{-1}|$ for any finite subsets $A, B, C \subset G$; this estimate is used repeatedly to prove Lemma 4.

4.4.2 Traces and growth

In this section, we seek to control $|A_k|$ by the size of traces of sets in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. As mentioned above, we begin by observing that, heuristically, a subset $A \subset \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ either grows rapidly under its action on itself or contains a "large" simultaneously diagonalizable subset:

Proposition 5. *Let K be a field and $\emptyset \neq A \subset \mathrm{SL}_2(K)$ a finite subset not contained in any proper subgroup of $\mathrm{SL}_2(K)$. Assume $|\mathrm{Tr}(A)| \geq 2$ and $|A| \geq 4$. Then A_4 contains at least $\frac{(|\mathrm{Tr}(A)|-2)(\frac{1}{4}|A|-1)}{|A_6|}$ simultaneously diagonalizable matrices.*

To prove this, one first shows that A_2 contains at least $\frac{1}{4}|A|-1$ elements with trace different from ± 2. A counting argument and a little work then give at least $\frac{(|\mathrm{Tr}(A)|-2)(\frac{1}{4}|A|-1)}{|A_6|}$ elements of $A_2^{-1}A_2$ that commute with some fixed $g \in A_2$, $\mathrm{Tr}(g) \neq \pm 2$. (The trace term here shows up as a lower bound for the number of conjugacy classes intersecting the non-unipotent elements of $A_2^{-1}A_2$, which in turn arises in the aforementioned counting argument.) Finally, one uses the fact that in any linear algebraic group, the centralizer of an element with distinct eigenvalues (such as our g) is abelian.

Next, one establishes some general results on SL_2 to the following effect: Let K be a field with algebraic closure $\overline{K}$, $|K| > 3$, and let $A \subset \mathrm{SL}_2(K)$ be a finite subset not contained in any proper subgroup. Then, loosely speaking, there is a universal constant k such that A_k acts "nontrivially" on generic pairs of vectors in $\overline{K}^2$ (or on pairs of points in the projective space $\overline{K}P^1$), in a suitable sense. These results are combined with Proposition 5 and still more counting arguments to provide both lower and upper bounds on the size of A (or boundedly many products thereof) in terms of traces:

Proposition 6. *For K and A as in Proposition 5 with $|K| > 3$, there exists $k \in \mathbb{N}$ independent of K, A such that*

$$|A_k| \geq \frac{1}{2}\left(\frac{1}{4}\frac{(|\mathrm{Tr}(A)|-2)(\frac{1}{4}|A|-1)}{|A_6|} - 5\right)\left(\frac{(|\mathrm{Tr}(A)|-2)(\frac{1}{4}|A|-1)}{|A_6|}\right)^2.$$

Proposition 7. *For any field K and A a finite subset of $\mathrm{SL}_2(K)$ not contained in any proper subgroup of $\mathrm{SL}_2(K)$, there exists $k \in \mathbb{N}$ independent of K, A such that*

$$|\mathrm{Tr}(A_k)| \gtrsim |A|^{1/3}.$$

4.4.3 A reduction to additive combinatorics

At last, we set out to prove part (a) of the Key Proposition. At any stage of our reasonings, we may of course assume p and $|A|$ to be larger than any fixed constants we wish. Moreover, given any reasonable fixed constants $c, \alpha > 0$ and $k \in \mathbb{N}$, we may assume $|A_k| < c|A|^\alpha$; if not, we are done by Lemma 4. Throughout the following, k_j and c_j will denote suitably chosen universal constants.

Combining these observations with Propositions 5 and 7, we see that A_{k_0} contains a large set B of simultaneously diagonalizable matrices; the set $V \subset \mathbb{F}_{p^2}$ of their eigenvalues must be at least as large. Adjusting constants if necessary as discussed in the previous paragraph, it turns out we may assume

$$|V| \gtrsim |A|^\alpha$$

for any $\alpha < 1/3$ and

$$C_\delta < |V| < p^{1-\delta/3}$$

for any given C_δ which depends only on δ. We know how to relate the size of A_k to the size of $\mathrm{Tr}(A_j) \subset \mathbb{F}_p$ for suitable k, j, so we might be tempted to use lower bounds for the size of the set $\{x + x^{-1} \mid x \in V\}$. This is not actually sufficient, but we might hope to do better by considering $|\{f(x, y) \mid x, y \in V\}|$ for some more general function f; this approach will work, up to replacing V by a suitable V_k. Based on the actual estimate for the size of $|V|$, an appeal to Proposition 6 and Lemma 4 shows that we need an estimate of the form

$$|\mathrm{Tr}(A_k)| \gtrsim |V|^{1+\varepsilon} \tag{1}$$

for suitably large k.

Now from the general results on $\mathrm{SL}_2(K)$ mentioned in the previous section, it so happens that there is some $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in A_{k_1}$ with a, b, c, d all nonzero. By computing $\mathrm{Tr}(h_1 g h_2 g^{-1})$ for $h_1, h_2 \in B$, we can show

$$\mathrm{Tr}(A_{k'}) \supset \{ad(xy + x^{-1}y^{-1}) - bc(x^{-1}y + xy^{-1}) \mid x, y \in V_k\}$$

for any $k \in \mathbb{N}$ (here $k' = k \cdot k_0 + k_1 + k \cdot k_0 + k_1 \lesssim k$, actually). Returning to the desired estimate (1), we see that the following sum-product-type result will finish the proof:

Proposition 8. *For $\mathbb{F}_q$ the finite field of q elements, let $a_1, a_2 \in \mathbb{F}_q^*$ and $\delta > 0$. Then there exist $C, \varepsilon > 0$ depending only on δ such that for any $A \subset \mathbb{F}_q^*$ with $C < |A| < p^{1-\delta}$,*

$$|\{f(x,y) := a_1(xy + x^{-1}y^{-1}) + a_2(x^{-1}y + xy^{-1}) \,|\, x, y \in A_{20}\}| > |A|^{1+\varepsilon}.$$

This result should be compared with the sum-product theorem of Bourgain–Katz–Tao ([BKT]), which essentially replaces the function $f(x,y)$ with the simpler function $x + y$; cf. also [Ko].

Proof sketch. A little algebra shows that it's enough to prove

$$|\{a_1(r + r^{-1}) + a_2(s + s^{-1}) \,|\, r, s \in bA_k^2\}| > |A|^{1+\varepsilon}$$

for any $b \in A_j$, where j and k are suitably chosen. (Here A_k^2 denotes the squares of elements of A_k, rather than $A_k A_k$.) Further, we can eliminate the coefficients a_1 and a_2 via an application of the Ruzsa triangle inequality. Now it so happens that we can cover A_4 by a well-controlled number of cosets of A_2^2, which eventually reduces us to showing

$$|\{(x + x^{-1}) \cdot (y + y^{-1}) \,|\, x, y \in A_2\}| > |A|^{1+\varepsilon}.$$

To finish the proof, we set $w(x) := x + x^{-1}$ and note that $w(x)w(y) = w(xy) + w(xy^{-1})$. If we suppose toward a contradiction that $|w(A_2)w(A_2)| \leq |A|^{1+\varepsilon}$, then a variant of the Balog–Szemerédi–Gowers theorem gives the bound $|w(A') + w(A')| \lesssim |A|^{1+O(\varepsilon)}$ for some large subset $A' \subset A_2$. This, combined with our original assumption on $w(A_2)w(A_2)$, contradicts established sum-product estimates (cf. *e.g.* [TV], §2.8).[6] □

4.5 Expander graphs

A good reference for the material in this section is the book [Lu].

Definition 9. *Let $X(V,E)$ be a connected k-regular graph with $|V| = n$. X is an (n,k,c)-*expander graph *if for all $A \subset V$ with $|A| \leq n/2$ we have*

$$|\partial A| \geq c|A|,$$

where $\partial A := \{x \in V \,|\, d(x,A) = 1\}$.

[6] Here one combines the aforementioned Bourgain–Katz–Tao theorem with a result of Heath-Brown–Konyagin (Lemma 5 of [H-BK]); the proof in [TV] unifies these and bypasses the number-theoretic techniques of [H-BK].

A family of k-regular graphs $X_i(V_i, E_i)$ with $|V_i| = n_i \to \infty$ is an *expander family* if each X_i is an (n_i, k, c)-expander for some fixed $c > 0$. Equivalently, given a k-regular graph $X(V, E)$, consider the normalized random walk operator

$$\mathrm{T}_X f(x) = \frac{1}{k} \sum_{(x,y)\in E} f(y)$$

on $L^2(V)$. The largest eigenvalue of T_X is 1 for any graph, and $\{X_i\}$ is an expander family if and only if the second-largest eigenvalues of T_{X_i} are uniformly bounded away from 1 (the family has "(uniform) spectral gap").

Cayley graphs have proven to be a a good source of expander graphs, and those of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ with respect to certain generating sets are known to be expander families in p. It is not hard to see that if $\{\mathcal{G}(G, A)\}_{G,A}$ yields a family of k-regular expanders, then $\operatorname{diam} \mathcal{G}(G, A) \lesssim_k \log |G|$, which gives an improvement over Babai's conjecture within the family. Consider the family

$$\mathcal{F}_k = \{\mathcal{G}(\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z}), A) | p \text{ prime}, \langle A\rangle = \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z}), |A \cup A^{-1}| = k\}.$$

Of course, we have not shown $\mathcal{F}_k$ to be an expander family. However, the proof we have given does imply that $\mathcal{G}(\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z}), A)$ has a spectral gap of at least $Ck(\log p)^{-O(1)}$.

4.6 Recent further progress

To conclude, we mention some recent results related to those presented thus far. Concerning expanders, building off of the results we have discussed, Bourgain and Gamburd ([BG1]) were able to characterize the sets in $\mathrm{SL}_2(\mathbb{Z})$ whose projections mod p provide expander Cayley graphs of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, as well as to prove that Cayley graphs of $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ with respect to *random* generating sets of fixed size form expander families in p, "asymptotically almost surely." Furthermore, in [BG2], they obtained analogous spectral gap results for appropriate subgroups of $\mathrm{SU}_2(\mathbb{C})$; the methods of both [BG1] and [BG2] are themselves highly additive-combinatorial.

Finally, Helfgott ([He2]) has proven the analogue of the Key Proposition for $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$, thus resolving Babai's conjecture for the family $\{\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})\}_p$. The proof follows the same approach to that which we have described, and it relies on the result for SL_2; however, it involves a bit more use of the group structure than we needed.

References

[BHKLS] L. Babai, G. Hetyei, W. M. Kantor, A. Lubotzky, and Á. Seress. On the diameter of finite groups. *Proc. 31st IEEE FOCS* (1990), 857–865.

[BG1] J. Bourgain and A. Gamburd. Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Ann. of Math.* **167** (2008), 625–642.

[BG2] ——. On the spectral gap for finitely-generated subgroups of SU(2). *Invent. Math.* **171** (2008), 83–121.

[BKT] J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *Geom. Funct. Anal.* **14** (2004), 27–57.

[H-BK] D. R. Heath-Brown and S. V. Konyagin. New bounds for Gauss sums derived from kth powers, and for Heilbronn's exponential sums. *Quart. J. Math.* **51** (2000), 221–235.

[He1] H. A. Helfgott. Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. *Ann. of Math.* **167** (2008), 601–623.

[He2] ——. Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$. Preprint, arXiv:math.GR/0807.2027v1, 2008.

[Ko] S. V. Konyagin. A sum-product estimate in fields of prime order. Preprint, arXiv:math.NT/0304217v1, 2003.

[Lu] A. Lubotzky. *Discrete Groups, Expanding Graphs and Invariant Measures, Progress in Math.* **195**, Birkhäuser, Basel, 1994.

[NP] N. Nikolov and L. Pyber. Product decompositions of quasirandom groups and a Jordan-type theorem. Preprint, arXiv:math.GR/0703343v3, 2007.

[TV] T. Tao and V. Vu, *Additive Combinatorics, Cambridge Studies in Adv. Math.* **105**, Cambridge UP, Cambridge, 2006.

S. Zubin Gautam, UCLA
email: **sgautam@math.ucla.edu**

5 The true complexity of a system of linear equations

after W.T. Gowers and J. Wolf [3]
A summary written by Derrick Hart

Abstract

Let $L_1, \dots L_m$ be a system of m linear forms with d variables from an abelian group G. We say that such a system has true complexity k if control over the Gowers U^{k+1} norm of subset A implies that $L_i(\mathbf{x}) \in A$ for $i = 1, \dots m$ for approximately the statistically correct number of $\mathbf{x}$'s. We give a necessary condition for certain systems to have true complexity 2 in the case $G = \mathbb{F}_p^n$.

5.1 Introduction

In many problems in additive combinatorics it is often necessary to understand when an appropriately uniform subset of an abelian group G will contain not only a system of linear forms but the expected number of them. Perhaps the simplest motivating example is the case of three term arithmetic progressions in $\mathbb{Z}_p$. It is a short exercise to show that a subset A of $\mathbb{Z}_p$ which uniform in the sense that $sup_{\xi \neq 0} |\widehat{A}(\xi)|$ is small then A will contain the statistically correct number of three term arithmetic progressions.

In order to deal with four term arithmetic progressions and in general systems of linear forms one needs a much more complicated analysis. Let $\mathcal{L} = (L_1, \dots, L_m)$ be a system of linear forms with d variables. In order to count the images of these forms in a set A one considers the expression

$$\mathbb{E}_{x_1, \dots, x_d \in G} \prod_{i=1}^{m} A(L_i(x_1, \dots, x_d)).$$

Let f be the balance function given by $f(x) = A(x) - |A|/|G|$. Substituting in the above expression,

$$\mathbb{E}_{\mathbf{x} \in G^d} \prod_{i=1}^{m} A(L_i(\mathbf{x})) = \left(\frac{|A|}{|G|} \right)^m + \text{Middle Terms} + \mathbb{E}_{\mathbf{x} \in G^d} \prod_{i=1}^{m} f(L_i(\mathbf{x})).$$

If we can show that the second and third term are small in absolute value then we will get the desired result. The Middle Terms will automatically be

small if one has control over $\mathbb{E}_{\mathbf{x}\in G^d}\prod_{i=1}^m f(L_i(\mathbf{x}))$ and so this term will be the main focus of our attention.

The key tools in controlling this term are the Gowers uniformity norms.

Definition 1. [2] *Let G be a finite Abelian group. For any $k \geq 1$ and $f : G \to \mathbb{C}$, define the Gowers U^k-norm by the formula*

$$\|f\|_{U^k}^{2^k} := \mathbb{E}_{x,h_1,\ldots,h_k\in G} \prod_{\boldsymbol{w}\in\{0,1\}^k} C^{|\boldsymbol{w}|} f(x + \boldsymbol{w}\cdot\boldsymbol{h}),$$

where $C^{|\boldsymbol{w}|}f = f$ is $\sum_i w_i$ is even and $\overline{f}$ otherwise.

In this article we are chiefly interested in the U^2 and U^3 norms. We say that a function is c-uniform if $\|f\|_{U^2} \leq c$ and c-quadratically uniform if $\|f\|_{U^3} \leq c$.

It turns out that in most cases controlling $\|f\|_{U^k}$ for k large enough will allows one to control $\mathbb{E}_{\mathbf{x}\in G^d}\prod_{i=1}^m f(L_i(\mathbf{x}))$ and guarantee the expected number of images. However, since in most applications in which one needs to say more the associated analysis becomes increasingly complicated (and is not completely understood) as k grows and therefore finding the minimal k is of some importance.

Definition 2 (True Complexity)**.** *The true complexity of $\mathcal{L}$ is the smallest k with the following property . For every $\epsilon > 0$ there exists $\delta > 0$ such that if G is any finite abelian group and $f : G \to \mathbb{C}$ is any function with $\|f\|_\infty \leq 1$ and $\|f\|_{U^{k+1}} \leq \delta$, then*

$$\left|\mathbb{E}_{x_1,\ldots,x_d\in G}\prod_{i=1}^m f(L_i(x_1,\ldots,x_d))\right| \leq \epsilon.$$

In Green and Tao [1] the another notion of complexity of a system of linear forms is given.

Definition 3 (Cauchy-Schwarz Complexity)**.** *Let $\mathcal{L} = (L_1,\ldots,L_m)$ be a system of m linear forms in d variables. For $1 \leq i \leq m$ and $s \geq 0$, we say that $\mathcal{L}$ is k-complex at i if one can partition the $m-1$ forms $\{L_j : j \neq i\}$ into $k+1$ classes such that L_i does not lie in the linear span of any of these classes. The Cauchy-Schwarz complexity of $\mathcal{L}$ is defined to be the least k for which the system is k-complex at i for all $1 \leq i \leq m$, or ∞ if no such k exists.*

In the same paper it is shown that with this concept of complexity in tow one can prove the following theorem.

Lemma 4. *Let $f_1, \dots f_m$ be functions from $G \to [-1,1]$. Let $\mathcal{L}$ be have Cauchy-Schwarz complexity k. Then*

$$\left| \mathbb{E}_{x \in G^d} \prod_{i=1}^{m} f_i(L_i(x)) \right| \leq \min_i \|f_i\|_{U^{k+1}}.$$

Therefore in order to get the expected number of images of systems of linear forms with Cauchy-Schwarz complexity k one only needs to control the U^{k+1} norm. The first question one asks then is this upper bourd sharp? The answer at least for Cauchy-Schwarz complexity 2 is yes. There exist uniform sets in $\mathbb{Z}_p$ which contain too many four term arithmetic progressions. This result relies on the fact that x^2, $(x+s)^2$, $(x+3s)^2$ and $(x+2s)^2$ are linearly dependent. Is then the Cauchy-Schwarz complexity actually equal to the true complexity of a system linear forms? At least in the case of Cauchy-Schwarz complexity 2 Gowers and Wolf ([3]) show this to be false. However, they make the following conjecture.

Conjecture 5. *The true complexity of a system $\mathcal{L}$ is equal to the smallest k such that the functions L_i^{k+1} are linearly independent.*

Our goal is to give Gowers' and Wolf's proof of this conjecture in the case that $G = \mathbb{F}_p^n$ and the system of linear forms is of Cauchy-Schwarz complexity 2 .

5.2 Quadratic fourier analysis and initial reductions

¿From this point on we will refer to the property of the L_i^2 being linearly independent that of as square-independence. In the context of $\mathbb{F}_p^n$ when we say that a system of linear forms is square-independent what we really mean is that the quadratic forms $L_i^T L_i$ are linearly independent, i.e. that the matrices made of the coefficients of $L_i^T L_i$ are linearly independent over $\mathbb{F}_p$.

Theorem 6 (Main Theorem)**.** *For every $\epsilon > 0$ there exists a constant $c > 0$ with the following property. Let $f : \mathbb{F}_p^n \to [-1,1]$ be a c-uniform function. Let $\mathcal{L}$ be a square-independent system of linear forms, with Cauchy-Schwarz*

complexity at most 2. Then

$$\mathbb{E}_{\boldsymbol{x}\in(\mathbb{F}_p^n)^d}\prod_{i=1}^{m} f(L_i(\boldsymbol{x})) \leq \epsilon.$$

Our main tool will be that of quadratic Fourier Analysis. Consider a surjective linear map $\Gamma_1 : \mathbb{F}_p^n \to \mathbb{F}_p^{d_1}$. Also define a quadratic map $\Gamma_2 : \mathbb{F}_p^n \to \mathbb{F}_p^{d_2}$ with $\Gamma_2(x) = (q_1(x), \ldots, q_{d_2}(x))$ where the q_i are quadratic forms on $\mathbb{F}_p^n$. By the rank(Γ_2) we shall mean the rank of the the bilinear form $\sum_i \lambda_i\beta_i$ where λ_i are not all zero and β_i is the bilinear form associated with q_i. For each $a \in \mathbb{F}_p^{d_1}$ and $b \in \mathbb{F}_p^{d_2}$ consider the corresponding sets $M_a = \{x : \Gamma_1(x) = a\}$ and $N_b = \{x : \Gamma_2(x) = b\}$. Define $\mathcal{B}_1$ to be the algebra generated by the set M_a and $\mathcal{B}_2$ to be the algebra generated by N_b. Then $\mathcal{B}_1$ is referred to as a linear factor of complexity d_1 while the pair $(\mathcal{B}_1, \mathcal{B}_2)$ is referred as a quadratic factor of complexity $(d_1, d_2)!$. By the rank$(\mathcal{B}_1, \mathcal{B}_2)$ we shall mean the rank(Γ_2).

The following theorem allows us to decompose f into the sum of three functions. A function f_1 which is "quadratically structured" in the sense that it is constant on the atoms of $\mathcal{B}_2$ of the quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ which is of bounded complexity. A function f_2 which is small in L^2 which allows us to guarantee that $(\mathcal{B}_1, \mathcal{B}_2)$ has high rank. And finally, a function f_3 which is quadratically uniform.

Theorem 7 (Structure Theorem)**.** *Let p be a fixed prime. Let $\delta > 0$, $r : \mathbb{N} \to \mathbb{N}$ be a function which may depend on δ. Suppose that $n > n_0(r, \delta)$ is sufficiently large. Then given any function $f : \mathbb{F}_p^n \to [-1, 1]$, there exists a $d_0 = d_0(r, \delta)$ and a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ with*

$$\text{rank}(\mathcal{B}_1, \mathcal{B}_2) \geq r(d_1 + d_2) \quad \textit{and} \quad \text{complexity}(\mathcal{B}_1, \mathcal{B}_2) \leq (d_1, d_2),$$

with $d_1, d_2 \leq d_0$ together with a decomposition

$$f = f_1 + f_2 + f_3,$$

where

$$f_1 := \mathbb{E}(f|\mathcal{B}_2), \quad \|f_2\|_2 \leq \delta \quad \textit{and} \quad \|f_3\|_{U^3} \leq \delta.$$

The Structure Theorem allows to essentialy replace f with f_1 in the Main Theorem. To see this let $\delta > 0$. Let $r : d \mapsto 2md + C$. From the Structure Theorem (Theorem 7) there exists a d_0 and a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ with

$$\text{rank}(\mathcal{B}_1, \mathcal{B}_2) \geq 2m(d_1 + d_2) + C \quad \text{and} \quad \text{complexity}(\mathcal{B}_1, \mathcal{B}_2) \leq (d_1, d_2),$$

with $d_1, d_2 \leq d_0$. Replacing the first f,

$$\mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^m f(L_i(\mathbf{x})) = \mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d}(f_1 + f_2 + f_3)(L_1(\mathbf{x}))(\prod_{i=2}^m f(L_i(\mathbf{x})).$$

Applying Cauchy-Schwarz to the second term,

$$\left|\mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} f_2(L_1(\mathbf{x})) \prod_{i=2}^m f(L_i(\mathbf{x}))\right| \leq \mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} |f_2(L_1(\mathbf{x}))| \leq \|f_2\|_2 \leq \delta.$$

followed in turn by Lemma 4 to the third term,

$$\left|\mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} f_3(L_1(\mathbf{x})) \prod_{i=2}^m f(L_i(\mathbf{x}))\right| \leq \|f_3\|_{U^3} \leq \delta.$$

This gives the bound,

$$\mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^m f(L_i(\mathbf{x})) \leq \mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} f_1(L_1(\mathbf{x})) \prod_{i=2}^m f(L_i(\mathbf{x})) + 2\delta.$$

Continuing this process replacing one f at a time,

$$\mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^m f(L_i(\mathbf{x})) \leq \mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^m f_1(L_i(\mathbf{x})) + 2m\delta.$$

5.3 Dealing with f_1

In order to deal with $\mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^m f_1(L_i(\mathbf{x}))$ we need some preliminary calculations.

Lemma 8. *Let $\mathcal{L}$ be a square-independent system of linear forms and let $\Gamma_2 = (q_1, \ldots q_{d_2})$ be a quadratic map from $\mathbb{F}_p^n \to \mathbb{F}_p^{d_2}$ with $\operatorname{rank}(\Gamma_2) \geq r$. Let $\phi_1, \ldots, \phi_m$ be linear maps from $(\mathbb{F}_p^n)^d$ to $\mathbb{F}_p^{d_2}$ and $b_1, \ldots, b_m$ be elements of $\mathbb{F}_p^{d_2}$. Then for a randomly chosen element $\boldsymbol{x}$,*

$$\left|\Pr\left[\Gamma_2(L_i(\boldsymbol{x})) = \phi_i(\boldsymbol{x}) + b_i, i = 1, \ldots, m\right] - p^{-md_2}\right| \leq p^{-r/2}.$$

Proof. Let

$$\Pr\left[\Gamma_2(L_i(\mathbf{x})) = \phi_i(\mathbf{x}) + b_i, i = 1, \ldots, m\right] - p^{-md_2} =$$

$$\mathbb{E}_{\mathbf{x}}\mathbb{E}_{\lambda\in\Lambda^*}\prod_{i=1}^{m}\prod_{i=1}^{d_2}\chi(\lambda_{ij}(q_j(L_i(\mathbf{x}))-\phi_{ij}(x)-b_{ij})$$

Let $L_i(\mathbf{x})=\sum_{u=1}^{d}c_{iu}x_u$. If β_j is the bilinear form associated with q_j then

$$q_j(L_i(\mathbf{x}))=\sum_{u,v}c_{iu}c_{iv}\beta_j(x_u,x_v).$$

Consider a j such that λ_{ij} is non-zero for at least one i. Then from the square-independence of $\mathcal{L}$ there exists a u,v such that $\sum_i\lambda_{ij}c_{iu}c_{iv}\neq 0$. Setting $\beta_{tw}=\sum_{i,j}\lambda_{ij}c_{it}c_{iw}$ we note that since $\mathrm{rank}(\Gamma_2)\geq r$ and square-independence imply $\mathrm{rank}(\beta_{uv})\geq r$. It is now possible after some restatement to consider the

$$\mathbb{E}_{x_u,x_v}\chi(\beta_{x_u,x_v}+\psi_u(x_u)+\psi_v(x_v)-b),$$

where ψ_u and ψ_v are linear functionals. Now for $u=v$ we may apply the following lemma.

Lemma 9. *Let M be matrix of rank r and $b\in\mathbb{F}_p^n$. Then*

$$\left|\mathbb{E}_{x\in\mathbb{F}_p^n}\chi(x^TMx+b^Tx)\right|\leq p^{-r/2}.$$

When $u\neq v$ then for a fixed x_v then $\beta_{x_u,x_v}+\psi_u(x_u)+\psi_v(x_v)-b$ is linear in u. However, then the expection is zero unless $\beta_{x_u,x_v}+\psi_u(x_u)$ is constant. This occurs on a subspace of dimension $n-r$ and so the expectation is bounded by p^{-r}. □

Using this one can then take into account Γ_1 as well. For a full proof see Gowers and Wolf ([3]).

Lemma 10. *Let the $\mathcal{L}$ be a square-independent and with the dimension of the linear span of $\mathcal{L}$ be d'. Let $\boldsymbol{a}=(a_1,\ldots,a_m)\in(\mathbb{F}_q^{d_1})^m$ and $\boldsymbol{b}=(b_1,\ldots,b_m)\in(\mathbb{F}_q^{d_2})^m$. Then for a randomly chosen $\boldsymbol{x}$ we have that*

$$\Pr(a,b)=\Pr\left[\Gamma_1(L_i(\boldsymbol{x}))=a_i \text{ and } \Gamma_2(L_i(\boldsymbol{x}))=b_i \text{ for } i=1,\ldots,m\right]=$$

$$\begin{cases}0 & \text{for } \boldsymbol{a}\notin Z\\ p^{-d_1d'-d_2m}+R(a,b) & \text{for } \boldsymbol{a}\in Z,\end{cases}$$

where $|R(a,b)|\leq p^{d_1-d_1d'-r/2}$ and Z is a subspace of $(\mathbb{F}_q^{d_1})^m$ with dimension $d'd_1$.

We can now deal with f_1.

Theorem 11. *Let $\Gamma_1 : \mathbb{F}_p^n \to \mathbb{F}_p^{d_1}$ be a linear map and $\Gamma_2 : \mathbb{F}_p^n \to \mathbb{F}_p^{d_1}$ be a quadratic map with corresponding quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ with $rank(\mathcal{B}_1, \mathcal{B}_2) \geq r$. Let $f : \mathbb{F}_p^n \to [-1, 1]$ and let $f_1 := \mathbb{E}(f|\mathcal{B}_1)$. Let $\mathcal{L}$ be a square-independent system of linear forms. Then*

$$\mathbb{E}_{\boldsymbol{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^{m} f_1(L_i(\boldsymbol{x})) \leq \|f\|_{U^2} 4^m p^{d_1/4} + 2^{m+1} p^{m(d_1+d_2)-r/2}.$$

In order to prove this we first appeal to a result which is really a combination of a few basic results on projections onto linear factors and the U^2 norm.

Lemma 12. *Let $g = \mathbb{E}(f_1|\mathcal{B}_1)$. Then $\|g\|_2^4 \leq p^n \|f\|_{U^2}^4$.*

Setting $h = f_1 - g$ then we have

$$\mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^{m} f_1(L_i(\mathbf{x})) = \mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^{m} (h+g)(L_i(\mathbf{x})),$$

which we may then expand into 2^m terms containing products of g's and h's. For the $2^m - 1$ terms which contain at least one g we may use the fact that $\|g\|_\infty \leq 1$ and $\|h\|_\infty \leq 2$ until there is one remaining g followed by Lemma 12 to $\|g\|_1 \leq \|g\|_2 \leq \|f\|_{U^2} p^{d_1/4}$. The worst case occurs when there is only one g giving the bound $2^m c p^{d_1/4}$. Therefore,

$$\mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^{m} f_1(L_i(\mathbf{x})) \leq c 4^m p^{d_1/4} + \mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^{m} h(L_i(\mathbf{x})).$$

Let H be defined by the formula $H(\Gamma_1 x, \Gamma_2 x) = h(x)$. From Lemma 10

$$\begin{aligned}
\mathbb{E}_{\mathbf{x}} \prod_i h(L_i(\mathbf{x})) &= \mathbb{E}_{\mathbf{x}} \prod_i H(\Gamma_1(L_i(\mathbf{x})), \Gamma_2(L_i(\mathbf{x}))) = \sum_{\mathbf{a}\in Z, \mathbf{b}} \Pr(\mathbf{a}, \mathbf{b}) \prod_i H(a_i, b_i) \\
&= p^{-d_1 d' - d_2 m} \sum_{\mathbf{a}\in Z, \mathbf{b}} \prod_i H(a_i, b_i) + \sum_{\mathbf{a}\in Z, \mathbf{b}} R(\mathbf{a}, \mathbf{b}) \prod_i H(a_i, b_i) \\
&= \mathbb{E}_{\mathbf{a}\in Z} \prod_i \mathbb{E}_{b_i} H(a_i, b_i) + 2^m p^{d_1 d' + d_2 m + d_1 - d_1 d' - r/2}.
\end{aligned}$$

Similarly, one can derive from Lemma 10 that $|\mathbb{E}_{b_i} H(a_i, b_i)| \leq 2p^{d_1+d_2-r/2}$, which shows that

$$|\mathbb{E}_{\mathbf{a}\in Z} \prod_i \mathbb{E}_{b_i} H(a_i, b_i)| \leq (2p^{d_1+d_2-r/2})^m.$$

Combining these terms gives us the desired result.

5.4 Finishing the proof of the main theorem

¿From Lemma 11,

$$\mathbb{E}_{\mathbf{x}\in(\mathbb{F}_p^n)^d} \prod_{i=1}^m f(L_i(\mathbf{x})) \leq c4^m p^{d_1/4} + 2^{m+1} p^{m(d_1+d_2)-r/2} + 2m\delta.$$

$$\leq c4^m p^{d_1/4} + 2^{m+1} p^{-C/2} + 2m\delta.$$

Retroactively setting C such that $2^{m+1}p^{-C/2} \leq \epsilon/3$,

$$\leq c4^m p^{d_0/4} + \frac{\epsilon}{3} + 2m\delta.$$

Setting $\delta = \epsilon/6m$ gives

$$c4^m p^{d_0/4} + \frac{2\epsilon}{3}.$$

This now lets us set the uniformity $c = 4^{-m}p^{-d_0/4}\epsilon/3$, to give the desired result.

References

[1] Green, B. and Tao, T., *Linear equations in the primes.* Preprint;

[2] Gowers, W. T., *A new proof of Szemerédi's theorem.* GAFA, **11** (2001), 465–588;

[3] Gowers, W. T. and Wolf, J., *The true complexity of a system of linear equations.* Preprint;

Derrick Hart, Rutgers University
email: **dnhart@math.rutgers.edu**

6 On an Argument of Shkredov on Two-Dimensional Corners

after Michael T. Lacey and William McClain [1]
A summary written by Vjekoslav Kovač

Abstract

We consider the cardinality of the largest subset of $\mathbb{F}_2^n \times \mathbb{F}_2^n$ containing no corner, which is a triple of the form (x, y), $(x + d, y)$, $(x, y + d)$ with $d \neq 0$. We prove that this quantity is bounded by a constant multiple of $2^{2n \frac{\log_2 \log_2 n}{\log_2 n}}$.

6.1 Some history and the main result

In the spirit of the famous Roth-Szemerédi theorem on arithmetic progressions, it is natural to investigate density of sets that do not contain some fixed two-dimensional sub-structure. The simplest such sub-structure is a *corner*, i.e. a triple of points of the form

$$(x, y),\ (x + d, y),\ (x, y + d),$$

for some integers x, y and a positive integer d.

In order to state some quantitative results we define

$$r_\angle(N) := \max\{|A| \,:\, A \subseteq \{1, 2, \ldots, N\}^2,\ A \text{ contains no corners}\}.$$

The first result about asymptotics of $r_\angle(N)$ was found by Ajtai and Szemerédi [2] who proved $r_\angle(N) = o(N^2)$. Their proof actually gives an explicit bound

$$r_\angle(N) \lesssim \frac{N^2}{(\log^* N)^c},$$

where $c > 0$ is an absolute constant and $\log^*$ is the *iterated logarithm* function, i.e. the number if times one must take the logarithm in order to produce a number less than or equal to 1. Several other authors obtained similar explicit but very weak bounds. Shkredov [5, 6] was the first to produce a "reasonable" bound:

$$r_\angle(N) \lesssim \frac{N^2}{(\log \log N)^c}.$$

Green observed in [3, 4] that Shkredov's argument can be simplified in the context of finite fields. More generally, for a finite abelian group G we define a *corner* to be a triple (x, y), $(x + d, y)$, $(x, y + d)$, where $x, y, d \in G$ and $d \neq 0$. Again we denote

$$r_{\angle}(G) := \max\{|A| \,:\, A \subseteq G \times G,\ A \text{ contains no corners}\}.$$

Although the original problem is essentially the particular case $G = \mathbb{Z}/N\mathbb{Z}$, the greatest simplification of Shkredov's arguments is present in the case $G = \mathbb{F}_2^n$. If we denote $N = |\mathbb{F}_2^n| = 2^n$, then the result can be stated as

$$r_{\angle}(\mathbb{F}_2^n) \lesssim \frac{N^2}{(\log_2 \log_2 N)^{1/25}} = \frac{2^{2n}}{(\log_2 n)^{1/25}}.$$

Later Shkredov proved in [7] an analogous estimate for an arbitrary finite abelian group.

The result of the paper [1] is a further improvement of the above bound on $r_{\angle}(\mathbb{F}_2^n)$. Lacey and McClain show that

$$r_{\angle}(\mathbb{F}_2^n) \lesssim N^2 \frac{\log_2 \log_2 \log_2 N}{\log_2 \log_2 N} = 2^{2n} \frac{\log_2 \log_2 n}{\log_2 n}.$$

Here we elaborate on their proof.

6.2 Outline of the proof

Basic structure of the proof is the same as in the original Shkredov's (or adapted Green's) proof. It is iterative and its main ingredients are:

- Denifition of appropriate "box norms"
- Generalized von Neumann estimate
- Density increment on a sublattice
- Uniformizing a sublattice

The key new ingredient in [1] is considering three different "box norms" associated to three coordinate systems in $\mathbb{F}_2^n \times \mathbb{F}_2^n$. Accordingly, product lattices are replaced by intersections of two product lattices in different coordinate systems. Details are to follow.

6.3 Some notation

Let H denote a subspace of $\mathbb{F}_2^n$. Its dimension will decrease at every iterative step of the proof. $X, D, Y \subset H$ are subsets. We view X as a subset of the first coordinate associated to basis element e_1, Y as a subset of the second coordinate, associated to basis element e_2, and D as a subset of the "diagonal" coordinate associated to $e_1 + e_2$. We will be working with subsets of

$$S := X \times Y \cap X \overset{\text{diag}}{\times} D.$$

The density of X in H is

$$\delta_X := \mathbb{P}(X \mid H) = \tfrac{|X|}{|H|}\,,$$

and analogously for δ_Y and δ_D. In the iterative procedure these densities will decrease.

The following quantity measures "uniformity" of the distribution of X in H:

$$\|X\|_{\text{uni}} := \sup_{\xi \neq 0} \tfrac{|\widehat{X}(\xi)|}{|H|}\,.$$

If $\|X\|_{\text{uni}} \leq \eta$ then we say that X is *η–uniform.* Here $\widehat{X}$ represents the the Fourier transform of X:

$$\widehat{g}(\xi) := \sum_{x \in H} g(x)(-1)^{x \cdot \xi}.$$

After the deletion of a small subset, a uniform set is again uniform.

The density of A is

$$\delta := \mathbb{P}(A \mid S)\,.$$

This quantity will increase in the iterative procedure of the proof. We define the *balanced function* of A to be the function supported on S as

$$f(x, y) := A(x, y) - \delta S\,.$$

Throughout this proof we assume:

$$\|X\|_{\text{uni}}\,,\ \|Y\|_{\text{uni}}\,,\ \|D\|_{\text{uni}} \leq \upsilon\,, \qquad \upsilon := (\delta\delta_X\delta_Y\delta_D)^C$$

where C is a large constant which we need not specify exactly, as its precise value only influences implied constants in our main Theorem. We will use the notation υ' for a fixed function of υ, that tends to zero as υ does.

For a function $f\colon S \longrightarrow \mathbb{C}$, define the "box norm"

$$\|f\|_\square := \delta_D^{-4}\mathbb{E}_{\substack{x,x'\in X\\ y,y'\in Y}} f(x,y)f(x',y)f(x,y')f(x',y')$$

where we use the standard basis (e_1, e_2). When f is the balanced function of A, the norm being 'large' is an obstacle to A having the expected number of corners. We use two additional "box norms". In the $(e_1, e_1 + e_2)$ coordinate system we define

$$\|f\|_{\square,X}^4 := \delta_Y^{-4}\mathbb{E}_{\substack{x,x'\in X\\ d,d'\in D}} f(x,d)f(x',d)f(x,d')f(x',d')\,.$$

Also with respect to the $(e_2, e_1 + e_2)$ coordinate system we define

$$\|f\|_{\square,Y}^4 := \delta_X^{-4}\mathbb{E}_{\substack{y,y'\in Y\\ d,d'\in D}} f(y,d)f(y',d)f(y,d')f(y',d')\,.$$

6.4 Main ingredients of the proof

6.4.1 Generalized von Neumann lemma

The following lemma provides sufficient conditions for A to have a corner.

Lemma 1. *Suppose that $A \subset S$ with $\mathbb{P}(A \mid S) = \delta$ and we have the inequalities*

$$\delta_X\delta_Y\delta_D\delta^2 N > C\,,$$
$$\max\{\|f|_\square\,,\ \|f|_{\square,X}\,,\ \|f|_{\square,Y}\} \le \kappa\delta^{5/4}\,.$$

Then A has a corner.

Here C represents a large absolute constant and $0 < c, \kappa, \kappa' < 1$ are small fixed constants.

6.4.2 Density increment lemma

If the conditions of the previous lemma are not satisfied, then we can find a sublattice on which A has increased density. This is a result of the following lemma.

Lemma 2. *For $0 < \kappa$ there is a constant $0 < \kappa' < 1$ for which the following holds. Suppose that $A \subset S = X \times Y \cap X \overset{\text{diag}}{\times} D$ with $\mathbb{P}(A \mid S) = \delta$, that f is the balanced function of A on S, and that*

$$\max\{\|f|_{\square}\,,\ \|f|_{\square,X}\,,\ \|f|_{\square,Y}\} > \kappa\delta^{5/4}\,.$$

Then there exists $X' \subset X$, $Y' \subset Y$, $D' \subset D$ such that three conditions hold.

$$\text{either } X' = X, \text{ or } Y' = Y, \text{ or } D' = D;$$

$$\mathbb{P}(A \mid S') \geq \delta + \kappa'\delta^2\,, \qquad S' = X' \times Y' \cap X' \overset{\text{diag}}{\times} D'\,;$$
$$\mathbb{P}(X' \mid X)\,,\ \mathbb{P}(Y' \mid Y)\,,\ \mathbb{P}(D' \mid D) \geq \kappa'\delta^2.$$

We need only refine two of the three sets X, Y and D above. With uniformity in coordinate that is not refined, we then have that the set $S' = X' \times Y' \cap X' \overset{\text{diag}}{\times} D'$ has about the expected number of points in it.

6.4.3 Uniformizing a sublattice

The last auxiliary result tells us that we can find a *uniform* sublattice on which A has increased density. Uniformity is important since it is required in applying the Generalized von Neumann Lemma.

Lemma 3. *Suppose that X, Y, D are as above and*

- *$X' \subset X$, $Y' \subset Y$, $D' \subset D$, with $\mathbb{P}(X' \mid X) \geq c\delta^2$ and similarly for Y and D;*
- *Either $X' = X$, $Y' = Y$ or $D' = D$;*
- *$S' = X' \times Y' \cap X' \overset{\text{diag}}{\times} D'$;*
- *$\mathbb{P}(A \mid S') = \delta + c\delta^2$;*
- *$\dim(H) > C[\delta^4(\upsilon'')^2]^{-1}$, where $0 < \upsilon'' < 1$ is fixed.*

Then there exists $X'' \subset X'$, $Y'' \subset Y$, $D'' \subset D'$ and H', H'', translates of the same *subspace $H_0 \leq H$, so that*

$$\|X''\|_{\text{uni}}\,,\ \|Y''\|_{\text{uni}}\,,\ \|D''\|_{\text{uni}} \leq \upsilon'',$$
$$\mathbb{P}(A \mid S'') \geq \delta + \tfrac{c}{2}\,\delta^2\,, \qquad S'' = X'' \times Y'' \cap X'' \overset{\text{diag}}{\times} D''\,,$$
$$\dim(H_0) \geq \dim(H) - C[\delta^4(\upsilon'')^2]^{-1}\,,$$
$$\mathbb{P}(X'' \mid H') \geq \kappa\delta^2\mathbb{P}(X' \mid H)\,.$$

Analogously for Y'' and D''. In particular $\mathbb{P}(D'' \mid H' + H'') \geq \kappa\delta^2\mathbb{P}(D' \mid H)$

We emphasize that H' and H'' are translates of the same subspace of $H_0 < H$, After a joint translation of A, X, Y and D, we can assume that H' and H'' are in fact the same subspace H.

6.5 Proof of the main theorem

The proof is recursive and we describe its conditional loop.

Initialize $X \leftarrow \mathbb{F}_2^n$, $Y \leftarrow \mathbb{F}_2^n$, $D \leftarrow \mathbb{F}_2^n$, $S \leftarrow \mathbb{F}_2^n \times \mathbb{F}_2^n$, $H \leftarrow \mathbb{F}_2^n$. Likewise $\delta_X, \delta_Y, \delta_D \leftarrow 1$. Fix a set A_0 with density δ_0 in $\mathbb{F}_2^n \times \mathbb{F}_2^n$. Initialize $A \leftarrow A_0$ and $\delta \leftarrow \mathbb{P}(A \mid S)$.

We iteratively apply the following steps:

- If $\max\{\|f\|_\square\,,\ \|f\|_{\square,X}\,,\ \|f\|_{\square,Y}\} > \kappa\delta^{5/4}$, apply the density increment lemma.

- If X', Y' or D' is not $\upsilon = (\delta\delta_{X'}\delta_{Y'}\delta_D)^C$ uniform, apply the uniformity lemma. Suppose these sets are as in the Lemma: subsets $X'' \subset X'$, $Y'' \subset Y'$, $D'' \subset D'$ and affine subspaces $H', H'' \subset H$ containing X'', Y'', D''. After joint translation of X'', Y'', D'', A and H', H'', we can assume that $H' = H''$ and are subspaces of H.

- Update variables:

$$X \leftarrow X''\,, \quad Y \leftarrow Y''\,, \quad D \leftarrow D''\,, \quad H \leftarrow H'\,,$$
$$\delta_X \leftarrow \mathbb{P}(X'' \mid H')\,, \quad \delta_Y \leftarrow \mathbb{P}(Y'' \mid H')\,, \quad \delta_D \leftarrow \mathbb{P}(D'' \mid H')\,,$$
$$S \leftarrow X \times Y \cap X \overset{\text{diag}}{\times} D\,, \quad \delta \leftarrow \mathbb{P}(A \mid S).$$

- The density of the incremented A on the set S has increased by at least $\kappa\delta_0^2$. The incremented densities δ_X, δ_Y, and δ_D have decreased by at most $(\kappa\delta_0)^C$.

Once this loop stops we conclude that A has a corner, provided that the initial dimension is large enough. This loop must stop in $\lesssim \delta_0^{-1}$ iterates, since itherwise the density of A on the sublattice would exceed one. Thus we need to be able to apply our lemmas $\lesssim \delta_0^{-1}$ times. In order to do that, both X and H must be sufficiently large at each stage of the loop.

This requirement places lower bounds on $N = 2^n$. The most stringent of these comes from the loss of dimensions. Note that before the loop terminates, we can have δ_X as small as

$$\delta_X \geq (\kappa\delta_0)^{(\kappa\delta_0)^{-1}}.$$

In order to apply the uniformity lemma at that stage, we need

$$N > 2^{(C\delta_0)^{-C\delta_0^{-1}}}.$$

From this we get the bound stated in the main theorem. ■

References

[1] Lacey M. T., McClain W., *On an Argument of Shkredov on Two-Dimensional Corners*, Online Journal of Analytic Combinatorics, No. 2, 2007,
`arXiv:math/0510491v3 [math.CO]`

[2] Ajtai M., Szemerédi E., *Sets of lattice points that form no squares*, Stud. Sci. Math. Hungar., 9, 1974

[3] Green B., *Finite field models in additive combinatorics*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., Vol. 327,
`arXiv:math/0409420v1 [math.NT]`

[4] Green B., *An Argument of Shkredov in the Finite Field Setting*,
`http://www.dpmms.cam.ac.uk/~bjg23/`

[5] Shkredov I. D., *On one problem of Gowers*, Izv. ross. Akad. Nauk Ser. Mat., 70, 2006, No. 2, 179–221,
`arXiv:math/0405406v1 [math.NT]`

[6] Shkredov I. D., *On a Generalization of Szemeredi's Theorem*, `arXiv:math/0503639v1 [math.NT]`

[7] Shkredov I. D., *On a two-dimensional analog of Szemeredi's Theorem in Abelian groups*, `arXiv:0705.0451v1 [math.NT]`

Vjekoslav Kovač, UCLA
email: **vjekovac@math.ucla.edu**

7 An inverse theorem for the Gowers $U^3(G)$ norm over the finite field $\mathbb{F}_5^n$

after B. Green and T. Tao [1]
A summary written by Choongbum Lee

Abstract

We give an inverse theorem for functions with large Gowers $U^3(G)$ norm over $G = \mathbb{F}_5^n$ which can be stated as following. A bounded function $f : G \to \mathbb{C}$ has large $U^3(G)$ norm if and only if it has a large inner product with a function $e^{2\pi i\phi}$, where $\phi : \mathbb{F}_5^n \to \mathbb{R}/\mathbb{Z}$ is a quadratic phase function.

7.1 Introduction

Let's start by quoting a paragraph from Green and Tao [1]

> There has been much recent progress in the study of arithmetic progressions in various sets, such as dense subsets of the integers or of the primes. One key tool in these developments has been the sequence of *Gowers uniformity norms* $U^d(G), d = 1, 2, 3, \ldots$ on a finite additive group G; in particular, to detect arithmetic progressions of length k in G it is importatnt to know under what circumstances the $U^{k-1}(G)$ norm can be large.

The goal of this summary is to provide preliminaries and show the outline of the proof of an inverse theorem for functions with large $U^3(\mathbb{F}_5^n)$ norm and thereby answering the question above about the necessary condition for large $U^{k-1}(G)$ norm in a special case. Using this theorem we can deduce $r_4(\mathbb{F}_5^n) \ll N(\log\log N)^c$ for some constant c where $r_4(G)$ is the largest cardinality of a set $A \subset G$ which does not contain an arithmetic progression of length 4.

We also note that in the same paper the authors also proves an inverse theorem for general abelian group G but this is not the scope of this summary. The statement of definitions and theorems will be quoted from [1].

7.2 Preliminaries

In this section we will define the Gowers uniformity norm $U^d(G)$ and a seminorm $u^d(B)$ which will be defined over a subset $B \subset G$.

7.2.1 Notation

Here we collection the notations we will use in this summary.

G will denote a finite additive group where an additive group is an abelian group with operation $+$. If $f : G \to H$ is a function from one additive group to another, and $h \in G$, we define the shift operator T^h applied to f by $T^h f(x) := f(x+h)$, and the difference operator $h \cdot \nabla := T^h - 1$ applied to f by the formula $(h \cdot \nabla) f(x) = f(x+h) - f(x)$. We extend these definitions to functions of several variables by using subscript as following. $T^h_x(x,y) = f(x+h,y)$ and $(h \cdot \nabla_x) f(x,y) = f(x+h,y) - f(x,y)$

The expectation $\mathbb{E}_{x \in B} f(x) := \frac{1}{|B|} \sum_{x \in B} f(x)$ will be used to denote the average of f over B.

We will use $e : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ is the exponential map $e(x) := e^{2\pi i x}$ and $\mathcal{D} := \{z \in \mathbb{C} : |z| \leq 1\}$ for the unit disk.

7.2.2 Gowers uniformity norm $U^d(G)$, $\|\cdot\|_{U^d(G)}$

The following norm was first introduced by Gowers [4] to prove Szemerédi's theorem.

Definition 1. *(Gowers uniformity norm). Let $d \geq 0$, and let $f : G \to \mathbb{C}$ be a function. We define the Gowers uniformity norm $\|f\|_{U^d(G)} \geq 0$ of f to be the quantity*

$$\|f\|_{U^d(G)} := \big(\mathbb{E}_{x \in G, h \in G^d} \prod_{w \in \{0,1\}^d} \mathcal{C}^{|w|} T^{w \cdot h} f(x)\big)^{1/2^d}$$

where $w = (w_1, \ldots, w_d), h \in (h_1, \ldots, h_d), w \cdot h := w_1 h_1 + \ldots + w_d h_d, |w| := w_1 + \ldots + w_d$ and $\mathcal{C}$ is the conjugation operator $\mathcal{C} f(x) := \overline{f(x)}$

To feel comfortable with the $U^d(G)$ norm, here we explicitly write down the $U^2(G)$ norm and $U^3(G)$ norm.

$$\begin{aligned}
\|f\|_{U^2(G)} &:= \big(\mathbb{E}_{x,a,b \in G} f(x) \overline{f(x+a) f(x+b)} f(x+a+b)\big)^{1/4} \\
\|f\|_{U^3(G)} &:= \big(\mathbb{E}_{x,a,b,c \in G} f(x) \overline{f(x+a) f(x+b) f(x+c)} \\
&\qquad f(x+a+b) f(x+b+c) f(x+c+a) \overline{f(x+a+b+c)}\big)^{1/8}
\end{aligned}$$

We can also define $U^d(G)$ norms recursively as following.

$$\|f\|_{U^0(G)} := \mathbb{E}(f); \qquad \|f\|_{U^1(G)} := |\mathbb{E}(f)|$$

$$\|f\|_{U^d(G)} := (\mathbb{E}_{h \in G} \left\|T^h f \overline{f}\right\|^{2^{d-1}}_{U^{d-1}(G)})^{1/2^d}$$

Here are some basic properties of $\|\cdot\|_{U^d(G)}$.

1. (Norm) $\|\cdot\|_{U^d(G)}$ is a norm for $d > 1$
2. (Monotonicity) $\|f\|_{U^d(G)} \leq \|f\|_{U^{d+1}(G)}$
3. (Conjugation Symmetry) $\left\|\overline{f}\right\|_{U^d(G)} = \|f\|_{U^d(G)}$
4. (Phase invariance) $\|fe(\phi)\|_{U^d(G)} = \|f\|_{U^d(G)}$ whenever ϕ is a global polynomial phase function of order at most $d-1$ [7] .
5. (Shift Invariance) $\left\|T^h f\right\|_{U^d(G)} = \|f\|_{U^d(G)}$ for all $h \in G$

7.2.3 Local polynomial bias of order d, $\|\cdot\|_{u^d(B)}$

Here we define a seminorm which is called the local polynomial bias of order d. This seminorm will be used for the main theorem. i.e. we will later explore the relation between Gowers norm and local polynomial bias.

Definition 2. *(Locally polynomial phase functions) If B is any non-empty subset of a finite additive group G and $d \geq 1$, we say that a function $\phi : B \to \mathbb{R}/\mathbb{Z}$ is a polynomial phase function of order at most $d-1$ locally on B if we have*

$$(h_1 \cdot \nabla_x) \dots (h_d \cdot \nabla_x)\phi(x) = 0$$

whenever the cube $(x + w_1h_1 + \dots + w_dh_d)_{w_1,\dots w_d \in \{0,1\}}$ is contained in B. If $f : B \to \mathbb{C}$ is a function, we define the local polynomial bias of order d on B to be the quantity

$$\|f\|_{u^d(B)} := sup|\mathbb{E}_{x\in B}(f(x)e(-\phi(x)))|$$

where ϕ ranges over all local polynomial phase functions of order at most $d-1$ on B.

From now on we will refer to polynomial phase functions of degree at most 1 as linear phase functions, and of degree at most 2 as quadratic phase functions. Note that to measure the local polynomial bias, we are taking the inner product of f with functions of the form $e(\phi)$ where ϕ is a polynomial

[7] Global polynomial phase function will be defined in the next section.(See Definition 2) But we include this property here to see the similarity between the properties of the Gowers Uniformity norm and that of the local polynomial bias defined in next section

phase function. Therefore in some sense we are measuring the correlation of f with polynomial oscillations.

We can deduce the folllowing properties from the definition which are similar to the Gowers uniformity norm.

1. (Seminorm) $\|\cdot\|_{u^d(B)}$ is a seminorm for any $B \subset G$
2. (Monotonicity) $\|f\|_{u^d(B)} \leq \|f\|_{u^{d+1}(B)}$
3. (Conjugation Symmetry) $\left\|\overline{f}\right\|_{u^d(B)} = \|f\|_{u^d(B)}$
4. (Phase invariance) $\|fe(\phi)\|_{u^d(B)} = \|f\|_{u^d(B)}$ whenever ϕ is a locally polynomial phase function of order at most $d-1$ on B.
5. (Shift Invariance for $B = G$) $\left\|T^h f\right\|_{u^d(G)} = \|f\|_{u^d(G)}$ for all $h \in G$

By using these properties together with the properties of the Gowers $U^d(G)$ norm we can prove the following proposition.

Proposition 3. *Let G be an additive group and $f : G \to \mathbb{C}$ be a function. Then for all $d > 1$,*

$$\|f\|_{U^d(G)} \geq \|f\|_{u^d(G)}$$

Proof. Given any global polynomial phase function of degree at most $d-1$ we can use the monotonicity and phase invariance properties to deduce

$$\|f\|_{U^d(G)} = \|fe(-\phi)\|_{U^d(G)} \geq \|fe(-\phi)\|_{U^1(G)} = |\mathbb{E}_{x\in G}(f(x)e(-\phi(x)))|$$

and taking the supremum over all ϕ we have

$$\|f\|_{U^d(G)} \geq \|f\|_{u^d(G)}$$

□

This proposition asserts that large polynomial bias of order d forces large $U^d(G)$ norm. Now we are interested in the reverse direction. So the question becomes whether large $U^d(G)$ norm forces large polynomial bias of order d or not. For $d = 2$ case we indeed have an inverse theorem which isn't that difficult to deduce.

Proposition 4. *(Inverse theorem for $U^2(G)$ norm). Let $f : G \to \mathcal{D}$ be a bounded function. Then*

$$\|f\|_{u^2(G)} \leq \|f\|_{U^2(G)} \leq \|f\|_{u^2(G)}^{1/2}$$

Unfortunately for $d > 2$ in general additive group G, the converse isn't true. This was discovered by Furstenberg and Weiss [3] and the example can be found in [1], Example 2.4. This is why we only worry about $\mathbb{F}_5^n$ where an inverse theorem for $U^3(\mathbb{F}_5^n)$ norm does exist.

7.3 Main Theorem

The main theorem will positively answer our question for $G = \mathbb{F}_5^n$ and $d = 3$. That is, large $U^3(\mathbb{F}_5^n)$ norm does indeed force large polynomial bias of order 3. Actually we can replace $\mathbb{F}_5$ by $\mathbb{F}_p$ for any odd prime p where an almost identical proof will be applied. But we will work in $\mathbb{F}_5$ for concreteness and simplicity.

Theorem 5. *(Inverse theorem for $U^3(\mathbb{F}_5^n)$). Let $f : \mathbb{F}_5^n \to \mathcal{D}$ be a bounded function and let $0 < \eta \leq 1$*

(i) If $\|f\|_{U^3(\mathbb{F}_5^n)} \geq \eta$, then there exists a subspace $W \leq \mathbb{F}_5^n$ of codimension at most $(2/\eta)^C$ such that

$$\mathbb{E}_{y \in \mathbb{F}_5^n} \|f\|_{u^3(y+W)} \geq (\eta/2)^C$$

where we can take $C = 2^{16}$. In particular, there exists $y \in G$ such that $\|f\|_{u^3(y+W)} \geq (\eta/2)^C$.

(ii) Conversely, given any subspace $W \leq \mathbb{F}_5^n$ and any function $f : \mathbb{F}_5^n \to \mathbb{C}$ we have $\|f\|_{U^3(\mathbb{F}_5^n)} \geq \|f\|_{u^3(\mathbb{F}_5^n)} \geq 5^{-n}|W| \, \|f\|_{u^3(y+W)}$ for any $y \in \mathbb{F}_5^n$

Combining the two parts of Theorem 5 we see that

$$\|f\|_{u^3(\mathbb{F}_5^n)} \leq \|f\|_{U^3(\mathbb{F}_5^n)} \leq \frac{C}{\log^c(1 + 1/\|f\|_{u^3(\mathbb{F}_5^n)})}$$

for some absolute constants $c, C > 0$. Thus we obtained an inverse theorem as promised. But note that this bound is much more weaker than the bound of Proposition 4 for $U^2(G)$. Actually the first part of Theorem 5 asserts that the control is much better when we look at a quadratic bias over a well chosen affine subspace.

7.4 Outline of the proof

Throughout this section G will be fixed as $\mathbb{F}_5^n$

We will describe the easier proof of the second part first. Proposition 3 already gives us one side of the inequality so we only need to look at the second inequality $\|f\|_{u^3(\mathbb{F}_5^n)} \geq 5^{-n}|W| \, \|f\|_{u^3(y+W)}$ for any subspace W and any $y \in \mathbb{F}_5^n$. Each side of the norm is determined by the supremum of the inner product of f with functions of the form $e(p(x))$ where $p(x)$ is a quadratic phase function. But by using the following proposition we can see that every possible choice of $e(p(x))$ for the RHS can be also considered as a choice of $e(p(x))$ for the LHS and therefore prove the second part of the theorem.

Proposition 6. *(Quadratic Extension Theorem). Let G be an additive group, let $H \leq G$ be a subgroup, and suppose that $y \in G$. Then any quadratic phase function $\phi : y + H \to \mathbb{R}/\mathbb{Z}$ can be extended (non-uniquely in general) to a globally quadratic phase function on G.*

Now let's prove the first part of theorem 5 which asserts that if we have a function f with large $U^3(\mathbb{F}_5^n)$ norm, then it must have quadratic polynomial bias. Recall that quadratic polynomial bias is defined as the supremum of $|\mathbb{E}_{x\in\mathbb{F}_5^n} f(x)e(-p(x))|$ where $p(x)$ varies over polynomials of degree at most 2. Introducing $\phi(x)$ as the phase of $f(x)$ (That is $f(x) = e(\phi(x))$) will help us understand the heuristic of the argument. With this phase $\phi(x)$, the polynomial bias can be rewritten as,

$$\|f\|_{u^3(G)} = \sup|e(\phi(x) - p(x))|$$

Thus our goal of proving that f has large quadratic polynomial bias can be changed into proving that the phase $\phi(x)$ is 'almost' a quadratic polynomial. [8]

STEP 1 : Linearization of phase derivative.

A nice way of showing that a function $\phi(x)$ behaves like a quadratic polynomial is by proving that $(h \cdot \nabla)\phi(x) = \phi(x + h) - \phi(x)$ is approximately linear. The proof will go in this direction. Since we are talking about polynomial bias, we will first examine $\phi(x + h) - \phi(x)$ and prove it's linear bias which certainly exists if we allow to restrict ourself to a subspace V of

[8] Here the formal meaning of 'almost a quadratic polynomial' guides us back to the original formulation. Thus formally this will mean that f has large quadratic polynomial bias

$\mathbb{F}_5^n$.(Fortunately we have a bound on codimension). The formal statement we derive is

Proposition 7. *(Large $U^3(\mathbb{F}_5^n)$ gives linear phase derivative). Let $G = \mathbb{F}_5^n$, and let $f : G \to \mathcal{D}$ be a bounded function such that $\|f\|_{U^3(G)} \geq \eta$ for some $\eta > 0$. Then there exists a linear subspace V of G with the codimension bound*

$$n - dim(V) \leq 2^{C_1}\eta^{-C_1'}$$

and a translate $x_0 + V$ of this subspace, together with a linear transformation $M : V \to \widehat{G}$ and an element $\xi_0 \in \widehat{G}$ such that

$$\mathbb{E}_{h\in V}|\mathbb{E}_{x\in G}T^{x_0+h}f(x)\overline{f(x)}e(-(2Mh + \xi_0)\cdot x)| \geq 2^{-C_2}\eta^{C_2'}$$

It is permissible to take $C_i, C_i' = 2^{16}$ for $i = 1, 2$.

Note that the conclusion

$$\mathbb{E}_{h\in V}|\mathbb{E}_{x\in G}T^{x_0+h}f(x)\overline{f(x)}e(-(2Mh + \xi_0)\cdot x)| \geq 2^{-C_2}\eta^{C_2'}$$

can be rewritten as following by using the phase $\phi(x)$ instead of $f(x)$.

$$\mathbb{E}_{h\in V}|\mathbb{E}_{x\in G}e(\phi(x + x_0 + h) - \phi(x) - (2Mh + \xi_0)\cdot x)| \geq 2^{-C_2}\eta^{C_2'}$$

which exhibits the linearity of $\phi(x + h) - \phi(x)$[9].

STEP2: The symmetry argument

Even though we have established the linearity of $\phi(x + h) - \phi(x)$ by approximating it as a linear operator M on a large subspace V of $\mathbb{F}_5^n$, this doesn't immediately give us the 'almost' quadratic property of $\phi(x)$. For our argument to work, we need the linear operator M to be self-adjoint. Again this is possible if we restrict ourself to a subspace W of V with a bound on codimension.

STEP3: Eliminating the quadratic phase component

By using the results of STEP1 and STEP2 we can indeed show that f has large quadratic bias but this is not an immediate consequence. If we had an approximation such as $\phi(x + h) - \phi(x) - 2Mh \cdot x \approx 0$ then it is not that difficult to deduce $\phi(x) \approx Mx \cdot x + q(x)$ where $q(x)$ is a linear

[9] in the sense of linear bias

polynomial. Unfortunately the formal meaning of $\phi(x+h)-\phi(x)$ being linear in STEP1 is somewhat different from above. But still we can follow the idea (with almost identical steps) of the above argument and deduce the following, which implies that translations of $f|_W$ have large correlation with the quadratic phase function $Mx \cdot x$.

$$\mathbb{E}_{y\in G}\left\|\overline{f(x+y)}e(Mx\cdot x)\right\|_{u^3(W)} \geq 2^{-C_2}\eta^{C_2'}$$

and by the properties from section 7.2.3 we can conclude that

$$\mathbb{E}_{y\in G}\|f\|_{u^3(y+W)} \geq (\eta/2)^C$$

which is the conclusion of Theorem 5.

7.5 Application

As an application of Theorem 5 we can give an upper bound on $r_4(\mathbb{F}_5^n)$ where $r_4(G)$ is the largest cardinality of a set $A \subset G$ which doesn't contain any arithmetic progression of length 4.

Theorem 8. *(Szemerédi theorem for $r_4(\mathbb{F}_5^n)$). write $N = 5^n$. Then we have the bound*

$$r_4(\mathbb{F}_5^n) \ll N(\log\log N)^{-2^{-21}}$$

This theorem can be easily deduced from the next proposition (which can be proved using Theorem 5).

Proposition 9. *Let $\delta > 0$, suppose that*

$$n > 6(2/\delta)^{2^{20}}$$

and let $A \subset \mathbb{F}_5^n$ be a set with size at least δN. Suppose that A contains no four-term arithmetic progressions. Then we can find an affine subspace $x_0 + V$ of $\mathbb{F}_5^n$ with dimension $dim(V) \geq n/3$ such that we have the density increment

$$\mathbb{E}_{x\in x_0+V}1_A(x) \geq \mathbb{E}_{x\in\mathbb{F}_5^n}1_A(x) + (\delta/2)^{2^{20}}$$

References

[1] Green, B. and Tao, T., *An inverse theorem for the Gowers $U^3(G)$ norm*, To appear, Proc. Edin. Math. Soc.

[2] Tao, T. and Vu, V., *Additive Combinatorics* Cambridge studies in advanced mathematics 105, Cambridge University Press, 2006.

[3] H. Furstenberg and B. Weiss, *A mean ergodic theorem for* $N^{-1}\sum_{n=1}^{N} f(T^n x)g(T^{n^2}x)$, , Convergence in ergodic theory and probability (Columbus OH 1993), 193-227, Ohio State Univ. Math. Res. Inst. Publ., 5. de Truyter, Berlin, 1996.

[4] W.T.Gowers, *A new proof of Szemerédi's theorem for progressions of length four*, GAFA 8 (1998), no. 3, 529-551

[5] W.T.Gowers, *A new proof of Szemerédi's theorem*, GAFA 11 (2001), 465-588

Choongbum Lee, UCLA
email: **abdesire@ucla.edu**

8 New bounds for Szemerédi's Theorem, II: A new bound for $r_4(N)$

after B. Green and T. Tao [1]
A summary written by Kenneth Maples

Abstract

For N a large positive integer, let $r_k(N)$ denote the maximum cardinality of a subset $A \subset [N] = \{1, ..., N\}$ that does not contain an arithmetic progression of length k. Szemerédi showed that $r_3(N) \ll Ne^{-c\sqrt{\log\log N}}$, and later that $r_k(N) = o(N)$ for $k \geq 4$. We improve the known effective bounds for $r_4(N)$ to $r_4(N) \ll Ne^{-c\sqrt{\log\log N}}$.

8.1 The Problem

Let A be a subset of an abelian group Z, which will typically be the integers of $\mathbb{Z}/p\mathbb{Z}$. An *arithmetic progression of length* k is a subset of A of the form $\{x, x+r, x+2r, ..., x+(k-1)r\}$ for some initial value x and step r. The sequence $\{3, 8, 13, 18\}$ is an arithmetic progression of length 4 and step 5 in the positive integers.

As mentioned in the abstract, let $r_k(N)$ denote the maximum cardinality of a subset $A \subseteq [N] = \{1, ..., N\}$ *without* an arithmetic progression of length k. For example, $r_3(5) = 4$, which is achieved by the subset $A = \{1, 2, 4, 5\}$.

Roth proved that $r_3(N) \ll N(\log\log N)^{-1}$ in 1953 [3], which was subsequently improved by Szemerédi to $r_3(N) \ll Ne^{-c\sqrt{\log\log N}}$ (unpublished). Szemerédi later showed $r_k(N) = o(N)$ in [4] [5], which was improved by Gowers in 1998 to the effective bound $r_k(N) \ll N(\log\log N)^{-c_k}$.

In a series of papers, Green and Tao refined the analysis of Gowers to improve the effective bounds on $r_4(N)$ to the level of the estimates for $r_3(N)$. This note summarizes the second paper in the series, which establishes the following improvement on Gowers' bound in the $r_4(N)$ case; the inequality is equivalent to Szemerédi's bound on $r_3(N)$.

Theorem 1. *For N sufficiently large,*

$$r_4(N) \ll Ne^{-c\sqrt{\log\log N}}.$$

Stronger bounds for $r_3(N)$ and $r_4(N)$ have also been found but we will not discuss them here.

8.2 Notation

We will use standard notations for asymptotic estimates. In particular, the Vinogradov notation $f \ll g$ means that there is a positive constant C such that $f \leq Cg$ for all (sufficiently large) arguments of f and g. If the implied constant C depends on another value, say δ, then we will denote the dependence with a subscript: $f \ll_\delta g$. We will also employ Landau notation, where O, o, Ω, Θ have their usual significance. If we write c or C then they are placeholders for suitable small or large constants, respectively, which may denote different values even if used in the same expression.

If Z is any finite set and $E \subseteq Z$ is a subset, we write the density of E in Z as $\mathbb{P}(E) = \mathbb{P}_Z(E) = \mathbb{P}_{\zeta \in Z}(E) = \frac{|E|}{|Z|}$. Similarly, for any function $f : E \to \mathbb{C}$ we can define its expectation by $\mathbb{E}_Z f = \mathbb{E}_{\zeta \in Z} f(\zeta) = \frac{1}{|Z|} \sum_{\zeta \in Z} f(\zeta)$. L^p spaces take their usual meaning against this expectation.

This summary uses the language of factors which are σ-algebras on finite sets. In particular, a factor $\mathcal{B} \subseteq 2^X$ is a collection of subsets of a finite set X that is closed under unions, intersections, and complements. The minimal non-empty elements of the factor are called atoms. In other words, a factor partitions $\mathcal{P}$ of X into atoms, where the factor itself consists of all possible unions of the atoms. We let $\mathcal{B} \vee \mathcal{C}$ denote the smallest factor containing both $\mathcal{B}$ and $\mathcal{C}$. The conditional expectation of f relative to the factor $\mathcal{B}$ is denoted by $\mathbb{E}(f|\mathcal{B})$ and may be defined as the usual orthogonal projection.

8.3 Strategy and Initial Reductions

Theorem 1 is proven using the *density increment strategy* invented by Roth. The key step is to establish the following density increment proposition.

Proposition 2. *Let $\delta > 0$ and assume that $N \geq e^{C\delta^{-C}}$. Let $A \subseteq [N]$ with $|A| \geq \delta N$ be a subset with density at least δ. Suppose that A does not contain an arithmetic progressions of length 4. Then there is an arithmetic progression $P \subseteq [N]$ of length $|P| \gg N^{c\delta^C}$ on which A has a larger density; explicitly, we have the density increment*

$$\frac{|A \cap P|}{|P|} \geq (1 + c)\delta.$$

Let us suppose that this proposition is proven and then derive a proof of the main theorem. Suppose that we have constructed a subset $A \subseteq [N]$

with cardinality $|A| = \delta N$ that does not contain any (non-trivial) arithmetic progressions of length 4. We can apply the above proposition iteratively in the following way. We initialize the values $A_0 = A$, $N_0 = N$, and $\delta_0 = \delta$, and we will construct each A_k, N_k, and δ_k in turn. For each k, we apply the proposition to N_k, A_k, and δ_k and note that one of the following alternatives holds:

1. $N_k \leq e^{C\delta^{-C}}$ and our iteration must terminate, as the conditions of the proposition are not satisfied.

2. $N_k < e^{C\delta^{-C}}$ and the conditions of the proposition are satisfied. In particular, we can find an arithmetic progression P on which A_k has a larger density as described above.

We let the next set A_{k+1} be the image of $A_k \cap P$ under the affine transformation that takes P to $N_{k+1} = \{1, ..., |P|\}$. The new set has density $\delta_{k+1} \geq (1+c)\delta_k$ in the smaller interval, which we expand to $\delta_{k+1} \geq (1+c)^{k+1}\delta$. We also have the lower bound $N_{k+1} \gg N_k^{c\delta_k^C}$ from the proposition.

Clearly the algorithm cannot continue past $\delta_k \geq 1$. Because δ_k grows exponentially, the algorithm will terminate before $C\log(1/\delta)$ steps. Once this has happened, we can combine our two estimates for N_k at the termination time to see that

$$N^{(c\delta^C)^{C\log(1/\delta)}} \ll e^{C\delta^{-C}}.$$

Taking two logarithms and rearranging shows that

$$\delta \ll e^{-c\sqrt{\log\log N}}.$$

As this was true for any construction of A, it must hold for any subset which achieves the cardinality $r_4(N)$; in this case $\delta = r_4(N)/N$ and we are done. $\square$

We now embed the interval $[N]$ with the cyclic group $\mathbb{Z}/p\mathbb{Z}$ for some large prime p so we can use Fourier analytic techniques. The cautious reader will note that this may introduce spurious arithmetic progressions that "wrap around" the cyclic group; these are prevented by taking p is sufficiently large, say $p \in [4N, 8N]$.

We will analyze the quadrilinear form Λ, which is defined for functions $f_j : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ by

$$\Lambda(f_0, f_1, f_2, f_3) = \mathbb{E}_{x,h\in\mathbb{Z}/p\mathbb{Z}} f_0(x)f_1(x+h)f_2(x+2h)f_3(x+3h).$$

Notice that if 1_A is the characteristic function of some subset $A \subseteq \mathbb{Z}/p\mathbb{Z}$, then $\Lambda(1_A, 1_A, 1_A, 1_A) \leq 1/p$ if A does not contain an arithmetic progression of length 4; in fact, the product is positive only when $h = 0$ and $x \in A$.

The purpose of this quadrilinear form is the following reduction of the previous proposition.

Theorem 3. *Let p be a large prime, $N \in [p/8, p/4]$ an integer, and $f : \mathbb{Z}/p\mathbb{Z} \to [0,1]$ be a 1-bounded nonnegative function which vanishes outside $[N]$. Let $\delta = \mathbb{E}_{[N]}(f)$. Suppose $p \gg \exp(C\delta^{-C})$ for some suitably large constant C, and*

$$|\Lambda(f, f, f, f) - \Lambda(\delta 1_{[N]}, \delta 1_{[N]}, \delta 1_{[N]}, \delta 1_{[N]})| \gg \delta^4.$$

Then we can find an arithmetic progression P in $[N]$ obeying the length bound $|P| \gg p^{c\delta^C}$ and the density increment

$$\mathbb{E}_P(f) \geq (1 + c)\delta.$$

To see how this theorem implies Proposition 2, take $f = 1_A$ and explicitly count the number of length 4 arithmetic progressions in $[N]$. It therefore remains to prove this theorem.

8.4 Relevant Definitions

Before we can discuss the proof of Theorem 3, we must make some relevant definitions. The analysis depends on finding subsets of $[N]$ that roughly obey linear or quadratic equations. To this end, we will look at the level sets of "phase functions", which will be generalizations of linear and quadratic polynomials from $\mathbb{Z}/p\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$. We say that a function $\phi : \mathbb{Z}/p\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is a globally linear phase function if it satisfies the equation,

$$\phi(x + h_1 + h_2) - \phi(x + h_1) - \phi(x + h_2) + \phi(x) = 0.$$

Note that $\phi(x) = \xi x/p + \alpha$ is a globally linear phase function, where $\xi \in \mathbb{Z}/p\mathbb{Z}$ and $\alpha \in \mathbb{R}/\mathbb{Z}$.

We say that a function $\phi : \mathbb{Z}/p\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is a locally quadratic phase function on $B \subseteq \mathbb{Z}/p\mathbb{Z}$ if it satisfies the equation,

$$\begin{aligned}\phi(x + h_1 + h_2 + h_3) - \phi(x + h_1 + h_2) - \phi(x + h_2 + h_3) - \phi(x + h_1 + h_3)\\ + \phi(x + h_1) + \phi(x + h_2) + \phi(x + h_3) - \phi(x) = 0.\end{aligned}$$

where each sum $x + \cdots \in B$. Note that this definition is local, i.e. restricted to the subset B of the cyclic group.

Phase functions generate factors as their level sets in the following precise way. A linear factor of complexity at most d and resolution K is a factor of the form $\mathcal{B} = \mathcal{B}_{\phi_1} \vee \cdots \vee \mathcal{B}_{\phi_{d'}}$ where $d' \leq d$, each ϕ_k is a globally linear phase function, and $\mathcal{B}_\phi$ is the factor whose atoms are the sets $\{x : \|\phi(x) - j/K\|_{\mathbb{R}/\mathbb{Z}} < 1/2K\}$ for $j = 0, ..., K-1$. Here $\|\cdot\|_{\mathbb{R}/\mathbb{Z}}$ denotes the distance from the argument to the nearest integer. The complexity refers to the number of phase functions in its definition, and its resolution refers to the width of the level sets. A *pure* quadratic factor is defined in the same way but with locally quadratic phase factors relative to B, which is the atom of a linear factor We can therefore define a (general) quadratic factor to be a pair $(\mathcal{B}_1, \mathcal{B}_2)$ of factors where $\mathcal{B}_1$ is a linear factor and $\mathcal{B}_2$ is a factor whose restriction to each atom of $\mathcal{B}_1$ is a pure quadratic factor. These will turn out to be the factors relevant in our analysis.

8.5 Proof Outline

We will now prove Theorem 3. It turns out that if f satisfies the given hypotheses then it has a high relative expectation on a quadratic factor. The following proposition finds this factor; its proof is discussed in Section 8.6. Here $\mathcal{B}_{\text{triv}}$ denotes the factor generated by $[N]$, namely $\mathcal{B}_{\text{triv}} = \{\emptyset, [N], \mathbb{Z}/p\mathbb{Z} \setminus [N], \mathbb{Z}/p\mathbb{Z}\}$.

Proposition 4. *Let the assumptions be as in Theorem 3. Then there exists a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity and resolution bounded by $O(\delta^{-C})$ and an atom B_2 of $\mathcal{B}_2 \vee \mathcal{B}_{triv}$ of density $\mathbb{P}_{\mathbb{Z}/p\mathbb{Z}}(B_2) \gg \exp(-O(\delta^{-C}))$ and contained in $[N]$ such that*

$$\mathbb{E}_{B_2}(f) \geq (1 + c)\delta$$

for some absolute constant $c > 0$.

Next, we need to partition the quadratic atom into arithmetic progressions, as in the following proposition.

Proposition 5. *Let $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most (d_1, d_2) and some resolution K, and let B_2 be an atom of $\mathcal{B}_2$. Then one can partition $B_2 \cap [N]$ as the union of $\ll d_2^{O(d_2)} N^{1-c/(d_1+1)(d_2+1)^3}$ disjoint arithmetic progressions in $\mathbb{Z}/p\mathbb{Z}$.*

An appropriate form of the pigeonhole principle gives us an arithmetic progression of length at least $\exp(-O(\delta^{-C}))N^{c\delta^C}$ on which f has average value at least $(1+\frac{1}{2}c_0)\delta$. Careful control of the constants shows that the length of the arithmetic progression is indeed large enough. $\square$

8.6 The Gowers U^3-norm and the Quadratic Bohr Sets

The proof of Proposition 4 relies on the Gowers U^3-norm, which is defined for functions $\phi : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ by

$$\|f\|^8_{U^3(\mathbb{Z}/p\mathbb{Z})} = \mathbb{E}_{x,h_k\in\mathbb{Z}/p\mathbb{Z}} f(x)\overline{f(x+h_1)f(x+h_2)f(x+h_3)}f(x+h_1+h_2)\times \\ \times f(x+h_2+h_3)f(x+h_1+h_3)\overline{f(x+h_1+h_2+h_3)}.$$

The reason for using the Gowers U^3-norm is the following form of the Gowers Inverse U^3 Theorem.

Theorem 6 (Inverse $U^3(\mathbb{Z}/p\mathbb{Z})$ theorem for quadratic factors)**.** *Let $f : \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ be a 1-bounded function such that $\|f\|_{U^3(\mathbb{Z}/p\mathbb{Z})} \geq \eta$ for some $\eta \in (0,1)$. Suppose also that K is an integer such that $K \geq C\eta^{-C}$ for some sufficiently large constant $C > 0$. Then there exists a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\eta^{-C}), 1)$ and resolution K such that*

$$\|\mathbb{E}(f|\mathcal{B}_2)\|_{L^1(\mathbb{Z}/p\mathbb{Z})} \gg \eta^C.$$

The original theorem is proven in [2]. This version can be derived by an averaging argument from the usual form for locally quadratic phase functions.

Iterating the U^3 theorem gives the following bound.

Theorem 7 (Quadratic Koopman-von Neumann)**.** *Let $f : \mathbb{Z}/p\mathbb{Z} \to [-1,1]$ be a 1-bounded function, and let $\eta > 0$. Suppose also that K is an integer such that $K \geq C\eta^{-C}$ for some sufficiently large constant $C > 0$. Then there exists a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\eta^{-C}), O(\eta^{-C}))$ and resolution K such that*

$$\|f - \mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{triv})\|_{U^3(\mathbb{Z}/p\mathbb{Z})} \leq \eta.$$

Explicitly, this theorem is proven using an energy increment method. We construct the quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ iteratively. Initialize $\mathcal{B}_1 = \mathcal{B}_2 = \{\emptyset, \mathbb{Z}/p\mathbb{Z}\}$. At each step we apply Theorem 6 to $f - \mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}})$. Elementary estimates show that the "energy" $\|\mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}})\|^2_{L^2}$ increases by

at least η^C at each step. Because the energy is bounded above by 1, there are at most η^{-C} steps. This algorithm therefore constructs the desired quadratic factor with the given complexity.

We are now in a position to prove Proposition 4. If $g = \mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}})$, then applying the quadratic Koopman-von Neumann theorem above to Λ shows that we can replace f with g in the following sense.

Corollary 8 (Anomalous AP4 count on a quadratic factor)**.** *Let the assumptions be as in Theorem 3. Then there exists a quadratic factor* $(\mathcal{B}_1, \mathcal{B}_2)$ *in* $\mathbb{Z}/p\mathbb{Z}$ *of complexity at most* $(O(\delta^{-C}), O(\delta^{-C}))$ *and resolution* $O(\delta^{-C})$ *such that the function* $g := \mathbb{E}(f|\mathcal{B}_2 \vee \mathcal{B}_{triv})$ *obeys*

$$|\Lambda(g, g, g, g) - \Lambda(\delta 1_{[N]}, \delta 1_{[N]}, \delta 1_{[N]}, \delta 1_{[N]})| \gg \delta^4.$$

With this corollary proven, let Ω be the set of numbers in $\mathbb{Z}/p\mathbb{Z}$ where $g \geq (1+c)\delta$. Norm control of Λ, along with the previous reduction, show that $\mathbb{P}_{\mathbb{Z}/p\mathbb{Z}}(\Omega) \gg \delta^4$. Our bounds on the complexity and resolution of $\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}}$ further bound the number of atoms, which by the pigeonhole principle gives us our desired high-density atom. $\square$

References

[1] Green, B. and Tao, T., *New bounds for Szemerédi's theorem, II: a new bound for* $r_4(N)$. arXiv:math/0610604v1 [math.NT], pp. 1–26.

[2] Green, B. and Tao, T., *An inverse theorem for the Gowers* $U^3(G)$ *norm.* Proc. Edinb. Math. Soc. (2) 51 (2008), no. 1, pp. 73–153.

[3] Roth, K., *On certain sets of integers.* J. London Math. Soc. 28 (1953), pp. 245–252.

[4] Szemerédi, E., *On sets of integers containing no four elements in arithmetic progression.* Acta Math. Acad. Sci. Hungary. 20 (1949), pp. 89–104.

[5] Szemerédi, E., *On sets of integers containing no* k *elements in arithmetic progression.* Acta Arith. 27 (1975), 299–345.

Kenneth Maples, UCLA
email: **maples@math.ucla.edu**

9 An inverse theorem for the Gowers $U^3(G)$ norm

after Ben Green and Terence Tao [1]
A summary written by Eyvindur Ari Palsson

Abstract

The sequence of Gowers uniformity norms $U^d(G)$, $d = 1, 2, 3, \ldots$ on a finite additive group G, are a great tool in the study of arithmetic progressions in various sets. To detect arithmetic progressions of length k in G it is important to know when the $U^{k-1}(G)$ norm can be large. We state an inverse theorem for the $U^3(G)$ norm on a finite, additive group G of odd order and give an outline of the proof.

9.1 Introduction

Let G be a finite additive group, that is a finite group with a commutative group operation $+$. Throughout the talk, we will use $N := |G|$ for the cardinality of G.

Let $f : G \to H$ be a function from one additive group to another and $h \in H$. We define the shift operator T^h applied to f by the formula $T^h f(x) := f(x + h)$. We define the difference operator $h \cdot \nabla := T^h - 1$ applied to f by the formula $(h \cdot \nabla)f(x) := f(x+h) - f(x)$. We then extend these definitions to functions of several variables by subscripting the variable to which the operator is applied. For example $T^h_x f(x, y) = f(x + h, y)$.

If $f : G \to \mathbb{C}$ is a complex valued function, and $B \subseteq G$ is a non-empty subset of G, we will write $\mathbb{E}_{x\in B} f(x) := \frac{1}{|B|} \sum_{x\in B} f(x)$, which is the average of f over B. We will further write $\mathbb{E}_{x\in G} f(x)$ as $\mathbb{E} f(x)$ when the domain G of f is clear from the context.

If $f_0, \ldots, f_{k-1} : G \to \mathbb{C}$, we define the k-linear form $\Lambda_k(f_0, \ldots, f_{k-1}) \in \mathbb{C}$ by

$$\Lambda_k(f_0, \ldots, f_{k-1}) := \mathbb{E}_{x,r\in G} f_0(x) T^r f_1(x) \ldots T^{(k-1)r} f_{k-1}(x).$$

Note that if $A \subseteq G$ and $f_0 = \ldots = f_{k-1} = 1_A$ then $\Lambda_k(1_A, \ldots, 1_A)$ is the number of progressions of length k, including those with common difference 0, divided by the normalizing factor of N^2.

If $(N, (k-1)!) = 1$ and A contains no proper progressions of length k then we see that $\Lambda_k(1_A, \ldots, 1_A) = |A|/N^2$, which will be quite small when

N is large. We are thus interested in determining whether $\Lambda_k(1_A, \ldots, 1_A)$ is small or large.

Definition 1. *(Gowers uniformity norm). Let $d \geq 1$, and let $f : G \to \mathbb{C}$ be a function. We define the Gowers uniformity norm $\|f\|_{U^d(G)} \geq 0$ of f to be the quantity*

$$\|f\|_{U^d(G)} := \left(\mathbb{E}_{x \in G, h \in G^d} \prod_{\omega \in \{0,1\}^d} \mathcal{C}^{|\omega|} T^{\omega \cdot h} f(x) \right)^{1/2^d},$$

where $\omega = (\omega_1, \ldots, \omega_d)$, $h = (h_1, \ldots, h_d)$, $\omega \cdot h := \omega_1 h_1 + \ldots + \omega_d h_d$, $|\omega| := \omega_1 + \ldots + \omega_d$, and $\mathcal{C}$ is the conjugation operator $\mathcal{C} f(x) := \overline{f(x)}$.

It turns out that $\|\cdot\|_{U^1(G)}$ is not a norm. However for $d > 1$, one can show that $\|\cdot\|_{U^d(G)}$ is indeed a norm.

An equivalent definition of the $\|f\|_{U^d(G)}$ norms is given by the recursive formulae

$$\|f\|_{U^1(G)} = |\mathbb{E}(f)|$$

$$\|f\|_{U^d(G)} := \left(\mathbb{E}_{h \in G} \|T^h f \bar{f}\|_{U^{d-1}(G)}^{2^{d-1}} \right)^{1/2^d}$$

for all $d \geq 2$.

This latter definition and induction give us a monotonicity property

$$\|f\|_{U^d(G)} \leq \|f\|_{U^{d+1}(G)} \text{ for } d = 0, 1, 2, \ldots$$

Since the k-linear form Λ_k can be used to count arithmetic progressions in a set then the following theorem shows how the Gowers uniformity norms can be used to bound the number of those arithmetic progressions.

Proposition 2. *(Generalized von Neumann Theorem). Let G be a finite abelian group with $(N, (k-1)!) = 1$. Let $f_0, \ldots, f_{k-1} : G \to \mathcal{D} := \{z \in \mathbb{C} : |z| \leq 1\}$ be functions. Then we have*

$$|\Lambda_k(f_0, \ldots, f_{k-1})| \leq \min_{1 \leq j \leq k} \|f_j\|_{U^{k-1}(G)}$$

Let $e : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ be the exponential map $e(x) := e^{2\pi i x}$. If f has the form $f(x) := e(\phi(x))$, for some phase function $\phi : G \to \mathbb{R}/\mathbb{Z}$, then

$$\|f\|_{U^d(G)}^{2^d} = \mathbb{E}_{x, h_1, \ldots, h_d \in G} e((h_1 \cdot \nabla_x) \ldots (h_d \cdots \nabla_x) \phi(x)).$$

This shows that the U^d norm in some sense measures the oscillation in the dth "derivative" of the phase, since $h \cdot \nabla$ is the discrete analogue of a directional derivative operator.

Definition 3. *(Locally polynomial phase functions). If B is any non-empty subset of a finite additive group G and $d \geq 1$, we say that a function $\phi : B \to \mathbb{R}/\mathbb{Z}$ is a polynomial phase function of order at most $d-1$ locally on B if we have*

$$(h_1 \cdot \nabla_x) \ldots (h_d \cdot \nabla_x)\phi(x) = 0$$

whenever the cube $(x + \omega_1 h_1 + \ldots + \omega_d h_d)_{\omega_1,\ldots,\omega_d \in \{0,1\}}$ is contained in B. If $f : B \to \mathbb{C}$ is a function, we define the local polynomial bias of order d on B $\|f\|_{u^d(B)}$ to be the quantity

$$\|f\|_{u^d(B)} := \sup |\mathbb{E}_{x \in B}(f(x)e(-\phi(x)))|$$

where ϕ ranges over all local polynomial phase functions of order at most $d-1$ on B.

$\|f\|_{u^d(B)}$ is a seminorm, $\|f\|_{u^d(B)} \leq \|f\|_{u^{d+1}(B)}$, $\|\bar{f}\|_{u^d(B)} = \|f\|_{u^d(B)}$ and $\|T^h f\|_{u^d(G)} = \|f\|_{u^d(G)}$. When ϕ is a locally polynomial phase of degree at most $d-1$ on B then $\|fe(\phi)\|_{u^d(B)} = \|f\|_{u^d(B)}$. Similarly when ϕ is a global polynomial phase function of degree at most $d-1$ then it turns out that $\|fe(\phi)\|_{U^d(G)} = \|f\|_{U^d(G)}$.

Using this last invariance we get

$$\|f\|_{U^d(G)} = \|fe(-\phi)\|_{U^d(G)} \geq \|fe(-\phi)\|_{U^1(G)} = |\mathbb{E}_{x \in G}(f(x)e(-\phi(x)))|$$

whenever ϕ is a global polynomial phase of degree at most $d-1$ and taking suprema over all such ϕ we end up with

$$\|f\|_{U^d(G)} \geq \|f\|_{u^d(G)}$$

for all $d \geq 1$.

Let $\hat{G}$ be the Pontryagin dual of G, in other words the space of homomorphisms $\xi : x \longmapsto \xi \cdot x$ from G to $\mathbb{R}/\mathbb{Z}$. $\hat{G}$ is an additive group which is isomorphic to G. We define the Fourier coefficient $\hat{f}(\xi)$ of f at the frequency $\xi \in \hat{G}$ by the formula

$$\hat{f}(\xi) = \mathbb{E}_{x \in G} f(x)e(-\xi \cdot x).$$

We have some classical Fourier results for those new Fourier coefficients, for example the Fourier inversion formula and the Plancherel identity.

Using the fact that polynomials of degree at most 0 are constant we can easily see that

$$\|f\|_{U^1(G)} = \|f\|_{u^1(G)}.$$

Next using some straightforward Fourier techniques one can proof

Proposition 4. *(Inverse theorem for $U^2(G)$ norm). Let $f : G \to \mathcal{D}$ be a bounded function. Then*

$$\|f\|_{u^2(G)} \leq \|f\|_{U^2(G)} \leq \|f\|_{u^2(G)}^{1/2}$$

It is tempting to conjecture that the two norms $U^d(G)$ and $u^d(G)$ are also related for higher d.

The paper establishes an inverse theorem in the case where $d = 3$ and $G = \mathbb{F}_5^n$ is a finite field. However due to applications then we are interested in more general groups, for example $\mathbb{Z}/N\mathbb{Z}$. Unfortunately then in general it does not hold that any bounded function with a small $u^d(G)$ norm must have a small $U^d(G)$ norm. Let's look at an example from Furstenberg and Weiss that illustrates this point.

Example 5. *(Furstenberg and Weiss). Let N be a large prime number, and let M be the largest integer less than or equal to $\sqrt{N}$. Let $G := \mathbb{Z}/N\mathbb{Z}$, and let $f : G \to \mathbb{C}$ be the bounded function defined by setting*

$$f(yM + z) := e(yz/M)\psi(y/M)\psi(z/M)$$

whenever $-M/10 \leq y, z \leq M/10$, and $f = 0$ otherwise. Here $\psi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a non-negative smooth cutoff function which equals one on the interval $[-1/20, 1/20]$ and vanishes outside of $[-1/10, 1/10]$. Then a direct calculation shows that $\|f\|_{U^3(G)} \geq c_0$ for some absolute constant $c_0 > 0$ wheras a Weyl sum computation reveals that $\mathbb{E}(fe(-\phi)) = O(N^{-c})$ for any quadratic phase function ϕ and some explicit constant $c > 0$.

The key element in the example was that the function $yM + z \longmapsto yz/M$ is locally quadratic on the set $B := \{yM + z : -M/10 \leq y, z \leq M/10\}$ but does not extend to a globally quadratic phase function on all of G. We need to account for those local quadratic phase functions in order to produce a genuine inverse theorem for the $U^3(G)$ norm. We also must understand the generalization of sets like B. We'll accomplish that by looking at Bohr sets.

9.2 The inverse theorem

Definition 6. *(Bohr sets). Let G be a finite additive group, and let $S \subseteq \hat{G}$, $|S| = d$ be a subset of the dual group. We define a sub-additive quantity $\|\cdot\|_S$ on G by setting*

$$\|x\|_S := \sup_{\xi \in S} \|\xi \cdot x\|_{\mathbb{R}/\mathbb{Z}},$$

where $\|\cdot\|_{\mathbb{R}/\mathbb{Z}}$ denotes the distance to the nearest integer, and define the Bohr set $B(S, \rho) \subseteq G$ for any $\rho > 0$ to be the set

$$B(S, \rho) := \{x \in G : \|x\|_S < \rho \text{ for all } \xi \in S\}$$

The dependence of the Bohr set $B(S, \rho)$ on ρ can be rather discontinuous but fortunately we can get past that in applications by restricting our attention to regular Bohr sets. We can show that there is an abundance of those regular Bohr sets.

Definition 7. *(Regular Bohr sets). Let $S \subseteq \hat{G}$, $|S| = d$, be a set of characters, and suppose that $\rho \in (0, 1)$. A Bohr set $B(S, \rho)$ is said to be regular if one has*

$$(1 - 100d|\kappa|)|B(S, \rho)| \leq |B(S, (1 + \kappa)\rho| \leq (1 + 100d|\kappa|)|B(S, \rho)|$$

whenever $|\kappa| \leq 1/100d$.

Theorem 8. *(Inverse theorem for $U^3(G)$). Let G be a finite, additive group of odd order, let $f : G \to \mathcal{D}$ be a bounded function and let $0 < \eta \leq 1$.*

(i) *If $\|f\|_{U^3(G)} \geq \eta$, then there exists a regular Bohr set $B := B(S, \rho)$ in G with $|S| \leq (2/\eta)^C$ and $\rho \geq (\eta/2)^C$ such that*

$$\mathbb{E}_{y \in G} \|f\|_{u^3(y+B)} \geq (\eta/2)^C,$$

where it is permissible to take $C = 2^{24}$. In particular, there exists $y \in G$ such that $\|f\|_{u^3(y+B)} \geq (\eta/2)^C$.

(ii) *Conversely, if $B = B(S, \rho)$ is a regular Bohr set, $f : G \to \mathcal{D}$ is a bounded function and $\|f\|_{u^3(y+B)} \geq \eta$, then we have*

$$\|f\|_{U^3(G)} \geq (\eta^3 \rho^2 / C' d^3)^d$$

for some absolute constant C'.

9.3 Outline of proof for the inverse theorem

The proof of part (ii) is relatively straightforward so we'll focus on how to prove part (i) of the inverse theorem. The first step is to establish a "phase derivative" $h \longmapsto \xi_h$ for the function f and to establish some additivity properties on this phase derivative. We use the Balog-Szemerdi-Gowers theorem and Plünnecke inequalities, both from additive combinatorics. The following proposition is the main result in this step.

Proposition 9. *Let G be an arbitrary finite additive group, and let $f : G \to \mathcal{D}$ be a bounded function such that $\|f\|_{U^3(G)} \geq \eta$ for some $\eta > 0$. Then there exists a set $H' \subseteq G$, a function $\xi : H' \to \hat{G}$ whose graph $\Gamma' := \{(h, \xi_h) : h \in H'\} \subseteq G \times \hat{G}$ obeys the estimates*

$$|\Gamma'| \geq 2^{-C_1}\eta^{C_1'}N$$

and

$$|k\Gamma' - l\Gamma'| \leq (2^{C_2}\eta^{-C_2'})^{k+l}|\Gamma'| \text{ for all } k, l \geq 1.$$

Furthermore for each $(h, \xi_h) \in \Gamma'$ we have

$$|\mathbb{E}_{x\in G}T^h f(x)\bar{f}(x)e(-\xi_h \cdot x)| \geq \eta^4/2.$$

The second step is the linearization of the phase derivative. That is, the function $h \longmapsto \xi_h$, which roughly speaking captures the derivative of the phase of f, matches up with a locally linear function. Using the previous proposition we get the following key result in this step.

Proposition 10. *(Large $U^3(G)$-norm implies locally linear phase derivative). Let G be an arbitrary finite additive group, and let $f : G \to \mathcal{D}$ be a bounded function such that $\|f\|_{U^3(G)} \geq \eta$ for some $\eta > 0$. Then there exists a set $S \subseteq \hat{G}$ with*

$$d_1 := |S| \leq 2^{C_3}\eta^{-C_3'},$$

a regular Bohr set $B_1 := B(S, \rho) \subseteq B(S, \frac{1}{4}) = B_0$ with $\rho \in [\frac{1}{16}, \frac{1}{8}]$, elements $x_0 \in G$ and $\xi_0 \in \hat{G}$, and a function $M : B_0 \to \hat{G}$ obeying the local linearity property

$$M(h \pm h') = Mh \pm Mh' \text{ whenever } \|h\|_S, \|h'\|_S \leq \frac{1}{8},$$

and such that

$$\mathbb{E}_{h\in B_1}|\mathbb{E}_{x\in G}T^{x_0+h}f(x)\overline{f(x)}e(-(\xi_0 + 2Mh) \cdot x)| \geq 2^{-C_4}\eta^{C_4'}.$$

The third step is a symmetry argument.

Lemma 11. *(Symmetry of derivative). Let the notation be as in the previous proposition. For any $x, y \in B_0$, let $\{x, y\}$ denote the anti-symmetric form*

$$\{x, y\} := M(x) \cdot y - M(y) \cdot x.$$

Then there exists a set S_3 of frequencies with $S_3 \supseteq \frac{1}{2}S$ and $d_3 := |S_3| \leq 2^{C_5}\eta^{-C_5'}$, and a Bohr set $B_3 = B(S_3, 2^{-C_6}\eta^{C_6'}) \subseteq B_1$, such that

$$\|\{x, z\}\|_{\mathbb{R}/\mathbb{Z}} \leq 2^{C_7}\eta^{-C_7'}\|x\|_{S_3}$$

The fourth and final step is then to use the symmetry of the derivative to get rid of the quadratic phase component in the last proposition.

References

[1] Green, B. and Tao, T., *An inverse theorem for the Gowers $U^3(G)$ norm*, arXiv:math/0503014v3 [math.NT];

Eyvindur Ari Palsson, Cornell
email: **eap48@cornell.edu**

10 On the Erdös-Volkmann and Katz-Tao ring conjectures

after J. Bourgain [2]
A summary written by Chun-Yen Shen

Abstract

We summarize the results of [2], where it is shown that a Borel subring of $\mathbf{R}$ cannot have Hausdorff dimension strictly intermediate between 0 and 1.

10.1 Introduction

The problem is raised in [1] on the existence of subrings R of the reals $\mathbf{R}$, R a Borel set, with Hausdorff dimension strictly between 0 and 1. It was already shown by Falconer that if the Hausdorff dimension $H-dimR > 1/2$, then $H - dimR = 1$. But his argument, based on dimension considerations of the distance set $\{|a-b|; a, b \in R \times R\} \subset \sqrt{R}$, does not seem to cover the range $H-dimR \leq 1/2$. In fact, the particular issue whether there is a Borel ring satisfying $H - dimR = 1/2$ and more specifically the discretized ring conjecture due to Katz and Tao [3] mentioned below, became recently a focus of attention because of its relevance to another problem of combinatorial measure theory, known as the Kakeya conjecture which says if A is a Borel subset of $\mathbf{R}^3$ containing a unit line segment in every direction, then $H - dimA = 3$.

Now we state the version of the ring problem as considered in [3]. First a definition, a bounded subset A of $\mathbf{R}$ is called a $(\delta, \sigma)_1$-set provided A is a union of δ-intervals and satisfies

$$|A \cap I| < (\frac{\gamma}{\delta})^{1-\sigma} \delta^{1-\epsilon} \tag{1}$$

whenever $I \subset \mathbf{R}$ is an arbitrary interval of sizes $\delta \leq \gamma \leq 1$ and ϵ is a small parameter. Roughly speaking, a $(\delta, \sigma)_1$-set behaves like the δ-neighbourhood of a σ-dimensional set.

It is conjectured in [3] that if A is a $(\delta, 1/2)_1$ set satisfying $|A| > \delta^{1/2+\epsilon}$, then one must have $|A+A| + |AA| > \delta^{1/2-c}$, with $c > 0$ an absolute constant. The following is the main result in [2]

Theorem 1. *If A is a $(\delta,\sigma)_1$-set, $0<\sigma<1$, such that $|A|>\delta^{\sigma+\epsilon}$. Then*

$$|A+A|+|AA|>\delta^{\sigma-c}$$

for an absolute value $c=c(\sigma)>0$.

The case $\sigma = 1/2$ is given special attention because of its many connections and equivalent formulations described in [3]. In particular, for the discrete version of this problem, when measure is replaced by cardinality, there is a result of Elekes that when A has finite cardinality $\sharp A$, at least one of $A+A$ and AA has cardinality $\geq \sharp A^{5/4}$. The proof of this result exploits the Szemerédi-Trotter theorem. This is heuristic evidence for the ring conjecture if one accepts the (somewhat questionable) analogy between discrete models and δ-discretized models.

Here we outline the proof of theorem 1. We proceed by contradiction, assuming

$$|A+A|+|AA|<\delta^{\sigma-\epsilon}.$$

The initial stages of the argument use only the additive information, thus $|A+A|<\delta^{-\epsilon}|A|$. It is processed through a multi-scale construction, based on Ruzsa's sumset estimates and, most importantly, quantitative versions of Freimann's famous theorem on finite sets of reals with small doubling set. The key difficulty comes from the fact that the hypothesis $|A+A|<\delta^{-\epsilon}|A|$ is by far too weak for a direct application to Freimann's theorem and the doubling constant $\delta^{-\epsilon}$ needs to be reduced. This is achieved using certain sub-multiplicativity properties of the doubling constants at various scales. The final product is a subset C of A with a tree-structure that exhibits a "multiscale porosity property". At this point, we start using the multiplicative structure and prove the existence of elements $x_1, x_2 \in A-A$ such that

$$|x_1C+x_2C|>\delta^{\sigma-\kappa},$$

where $\kappa>0$ is an absolute constant. In fact, the elements x_1, x_2 are obtained randomly according to a probability measure on $A-A$. Thus the conclusion is that

$$|AA-AA+AA-AA|\geq|A(A-A)+A(A-A)|>\delta^{\sigma-\kappa}.$$

The contradiction comes from the fact that under our assumptions

$$|A|>\delta^{\sigma+\epsilon}, |A+A|+|AA|<\delta^{\sigma-\epsilon},$$

there is a subset A' of A satisfying $|A'| > \delta^{\sigma+C\epsilon}$ and $|A'A' - A'A' + A'A' - A'A'| < \delta^{\sigma-C\epsilon}$ (for some constant C). This non-trivial fact was proven by Katz and Tao [3] for $\delta = 1/2$ and the arguments can be extended to the general cases.

10.2 Preliminary results

We present a number of results to be used to prove theorem 1. The first one is the so-called Freimann theorem and Lemma 4 is so-called Katz-Tao lemma.

Theorem 2. *Let $A \subset \mathbf{R}$ be a finite set such that*

$$|A + A| < K|A|.$$

Then A is contained in a translation of a proper d-dimensional progression P

$$P = \{b_1\xi_1 + \cdots + b_d\xi_d; 0 < b_i \leq h_i, 1 \leq i \leq d\}$$

satisfying

$$d < K$$

and

$$|P| = \Pi_{i=1}^{d} h_i < exp(K^2(logK)^3)|A|.$$

Lemma 3. *Let $A \subset \mathbf{R}$ be a bounded set with $\delta N(A, \delta) < \kappa$ and assume*

$$\frac{N(2A, \delta)}{N(A, \delta)} < K$$

where $N(A, \delta)$ denotes the δ-covering number. Then there are $\sigma > 0$, $\sigma\delta' > \delta$ and a subset A' of A contained in a union of intervals of size $\sigma\delta'$ and separation at least δ' satisfying the following conditions.

$$\sigma < (logK)^C \kappa^{(logK)^{-C}}.$$

$$N(A', \delta) > \frac{N(A, \delta)}{Klog1/\delta}.$$

Lemma 4. *Let A be a Borel set obtained as union of δ-intervals and $0 < \sigma < 1$, satisfying $|A| > \delta^{\sigma+\epsilon}, |A + A| + |AA| < \delta^{\sigma-\epsilon}$. Then there is $A' \subset A$ with $|A'| > \delta^{\sigma+\epsilon'}$ such that*

$$|A'A' - A'A'| < \delta^{\sigma-\epsilon'}.$$

where $\epsilon' \to 0$ when $\epsilon \to 0$.

Proof. We outline the main steps and refer the reader to [3] for further details. As in [3], we use $X \lessapprox Y$ to mean $X \leq C_\epsilon \delta^{-C\epsilon} Y$ and $X \approx Y$ to mean $X \lessapprox Y$ and $Y \lessapprox X$. A subset B of A is called a refinement of A provided B is a union of δ-intervals and $|B| \gtrapprox |A|$.

First using the assumptions, we may obtain refinements C, D of A such that for all $(c, d) \in C \times D$

$$|\{(a_1, a_2, a_3, a_4, a_5, a_6) \in A^6; c-d = (a_1-a_4)-(a_2-a_5)+(a_3-a_6)\}| \approx \delta^{5\sigma} \quad (2)$$

The measure $||$ refers here to the corresponding hyperplane in $\mathbf{R}^6$. We may also assume moreover

$$|c - d| \approx 1, c \in C, d \in D.$$

Next, fix $a_1, a_2, a_3, a_4, a_5 \in A$, $c \in C$, $d \in D$. Multiplying in (2) with $\frac{a_1 a_2 a_3}{a_4 a_5}$, we get

$$|\{(e_1, e_2, e_3, e_4, e_5, e_6) \in (\frac{AAAA}{AA})^6; \frac{a_1 a_2 a_3}{a_4 a_5}(c-d) = (e_1-e_4)-(e_2-e_5)+(e_3-e_6)\}| \approx \delta^{5\sigma}.$$

Since by assumption on the set A (together with the multiplicative version of Ruzsa sumset estimates), we have

$$|\frac{AAAA}{AA}| \sim \delta^\sigma,$$

and by Fubini theorem we have

$$|\frac{AAA(C-D)}{AA}| \lesssim \delta^\sigma.$$

Since

$$|C|, |D|, |CD| \approx \delta^\sigma,$$

from the multiplicative version of (2), there are further refinements $C' \subset C$, $D' \subset D$ such that for all $c \in C', d \in D'$

$$|\{(c_1, c_2, c_3, d_1, d_2, d_3) \in C^3 \times D^3; cd = c_1 d_1 (c_2 d_2)^{-1} c_3 d_3\}| \approx \delta^{5\sigma} \quad (3)$$

Now we fix $c, c' \in C'$ and $d, d' \in D'$. Denote X the set obtained in (3). If $(c_1, c_2, c_3, d_1, d_2, d_3) \in X$, we have

$$cd - c'd' = x_1 - x_2 + x_3 - x_4,$$

where

$$x_1 = \frac{(c_1 - d')d_1c_3d_3}{c_2d_2}$$
$$x_2 = \frac{(c' - d_1)d'c_3d_3}{c_2d_2}$$
$$x_3 = \frac{(c_3 - d_2)d'c'd_3}{c_2d_2}$$
$$x_4 = \frac{(c_2 - d_3)d'c'd_2}{c_2d_2}$$

The map $(c_1, c_2, c_3, d_1, d_2, d_3) \mapsto (x_1, x_2, x_3, x_1, c_2, d_2)$ is clearly a diffeomorphism. Therefoe we have

$$|\{(x_1, x_2, x_3, x_1, c_2, d_2) \in (\frac{AAA(C-D)}{AA})^4 \times C \times D, cd - c'd' = x_1 - x_2 + x_3 - x_4\}| \gtrsim \delta^{5\sigma}.$$

Finally, by Fubini theorem again, we may conclude that

$$\delta^{4\sigma} \times \delta^{\sigma} \times \delta^{\sigma} \gtrsim |\frac{AAA(C-D)}{AA}|^4 |C||D| \gtrsim |C'D' - C'D'|$$

and hence by letting $A' = C''$ we get the desired result.

□

10.3 Outline of Proof of Theorem 1

Using notation from [3], a bounded subset A of $\mathbf{R}$ is called a $(\delta, \sigma)_1$-set provided A is a union of δ-intervals and

$$|A \cap I| < (\frac{\gamma}{\delta})^{1-\sigma} \delta^{1-\epsilon}.$$

for any interval $I(x, \gamma)$, $\delta \leq \gamma \leq 1$ and ϵ is a very small number. We proceed by contradiction, thus we assume

$$|A| > \delta^{\sigma+\epsilon}, |A + A| + |AA| < \delta^{\sigma-\epsilon}.$$

¿From lemma 4, we may assume moreover

$$|AA - AA| < \delta^{\sigma-\epsilon}$$

and hence

$$|AA - AA + AA - AA| < \delta^{\sigma-\epsilon}.$$

Next at the first stage, we use the bound on the sumset, i.e $|A + A| < \delta^{\sigma-\epsilon}$. ¿From lemma 3, we construct a subset C of A, $|C| > \delta^{\epsilon_1}|A|$, which has a tree structure. The key property is multi-scale porosity. This means that at each level of the tree, C is a union of subsets contained in disjoint intervals of a certain size which is small w.r.t their mutual separation. One then considers a set of the form

$$x_0C + xC$$

where $x_0, x \in A-A$, x_0 fixed and x considered as a random variable governed by a measure ν on $A - A$. Using the porosity property, we have that

$$\int |x_0C + xC|\nu(dx) > \delta^{-\epsilon_2}|C|.$$

Here $\epsilon_2 \gg \epsilon_1 \gg \epsilon$. Thus

$$|AA - AA + AA - AA| \geq |(A - A)C + (A - A)C| > \delta^{-(\epsilon_2-\epsilon_1-\epsilon)+\sigma}$$

and in conclusion proves theorem 1.

10.4 Applications

In [3], Katz and Tao investigate three unsolved conjectures in geometric combinatorics, namely Falconer's distance problem, the dimension of sets of Furstenburg's type, and Erdós ring problem. They reduce these geometric problems to δ-discretized variants and show that these variants are all equivalent.

10.4.1 The Falconer distance problem

For any compact subset K of the plane $\mathbf{R}^2$, define the distance set $dist(K)$ of K by

$$dist(K) = \{|x - y|; x, y \in K\}.$$

Falconer conjectured that if $dimK \geq 1$, then $dim(dist(K)) = 1$, where $dimK$ denotes the Hausdorff dimension of K. As progress towards this conjecture, the best result to date is by Wolff who showed that $dim(dist(K)) = 1$ provided $dimK \geq 4/3$.

Now suppose that one only assumes that $dimK \geq 1$. An argument of Mattila shows that $dim(dist(K)) \geq 1/2$. Hence by theorem 1 we have

Theorem 5. *There exists an absolute constant $c > 0$ such that*

$$dim(dist(K)) \geq 1/2 + c,$$

whenever K is compact and satisfies $dimK \geq 1$.

10.4.2 Dimension of sets of Furstenburg type

Let $0 < \beta \leq 1$. We define a β-set to be a compact set $K \subset \mathbf{R}^2$ such that for every direction $\omega \in S^1$ there exists a line segment ℓ_ω with direction ω which intersects K in a set with Hausdorff dimension at least β. We let $\gamma(\beta)$ be the infimum of the Hausdorff dimensions of β-sets.
At present the best bounds known are

$$\max(\beta + \frac{1}{2}, 2\beta) \leq \gamma(\beta) \leq \frac{3}{2}\beta + \frac{1}{2}.$$

The most interesting value of β appears to be $\beta = 1/2$. In this case the two lower bounds on $\gamma(\beta)$ coincide to become $\gamma(1/2) \geq 1$. Again, by theorem 1 we have

Theorem 6. *The 1/2-sets must have Hausdorff dimension at least $1 + c$ for some absolute constants $c > 0$.*

Remark 7.

The fact that $\mathbf{R}$ is a totally ordered field is relevant, since the analogue of Erdös's ring problem is false for non-ordered fields such as the complex numbers or the finite field F_{p^2}. The analogues of Falconer's distance problem and the conjectures for Furstenburg sets also fail for these fields.

These problems are also related to the Kakeya problem in three dimensions, although the connection is more tenuous. The proof of ring conjecture also leads to an alternate proof of a result in [4], namely that Besicovitch sets in $\mathbf{R}^3$ have Minkowski dimension strictly greater than 5/2, and would not rely as heavily on the assumption that the line segments all point in different directions. Very informally, the point is that the arguments in [4] can be pushed a bit further to conclude that a Besicovitch set of dimension exactly 5/2 must essentially be a Heisenberg group over a ring of dimension 1/2.

References

[1] P. Erdös and B. Volkmann, *Additive Gruppen mit vorgegebener Hausdorffscher Dimension* . J. Reine Angew. Math **221** (1966),pp. 203–208;

[2] J. Bourgain, *On the Erdös-Volkmann and Katz-Tao ring conjectures.* GAFA, 13 (2003), pp. 334-365.

[3] N. Katz and T. Tao, *Some connections between Falconer's distance set conjectures and sets of Furstenberg type.* New York J.Math, 7 (2001), pp. 149-187.

[4] N. Katz, I. Łaba and T. Tao, *An improved bound for the Minkowski dimension of Besicovitch sets in R^3.* Annals of Math, 152 (2002), pp. 346-383.

Chun-Yen Shen, Indiana University
email: **shenc@indiana.edu**

11 Quantitative bounds for Freiman's Theorem

Abstract

A summary of the Chang/Ruzsa proof of Freiman's theorem.

11.1 Introduction

Let Z be an additive group (usually $\mathbb{Z}$ or the cyclic group $\mathbb{Z}_N$). If $d \geq 1$ is an integer, a *d-dimensional generalized arithmetic progression* (g.a.p.) in Z is a set P which can be written

$$\begin{aligned} P &= \{a + x_1 v_1 + \cdots + x_d v_d : 0 \leq x_i < l_i,\, 1 \leq i \leq d\} \\ &=: P(v_1, \ldots, v_d; l_1, \ldots, l_d; a), \end{aligned} \tag{4}$$

for $a, v_1, \ldots, v_d \in Z$ and non-negative integers $l_1, \ldots, l_d$. The length of P is

$$\ell(P) := \prod_{j=1}^{d} l_d.$$

P is *proper* if $|P| = \ell(P)$, or equivalently if each sum in (4) is distinct.

Freiman's theorem states that if $|2A| = |A + A|$ is small relative to $|A|$, then A is comparable to a g.a.p.:

Theorem 1. *Let $A \subset \mathbb{Z}$ be an additive set satisfying $|2A| \leq K|A|$ for some $K > 0$. Then A is contained in a g.a.p. of dimension $d \leq d(K)$ and length $\ell(P) \leq C(K)|A|$.*

According to Chang [2], Freiman's original proof (1973) was a difficult read, and it wasn't until the 90s that Bilu and Ruzsa offered more accessible arguments. In all of these results, the bounds on the function C were doubly exponential in K. In 2002, Chang [2] improved on Ruzsa's argument, addressing two inefficiencies and thereby proving the following theorems:

Theorem 2. *Given an additive set $A \subset \mathbb{Z}$ with $|2A| \leq K|A|$, the conclusions of Freiman's theorem hold with $d(K)$ and $\log C(K)$ bounded by $CK^2(\log K)^2$.*

It is conjectured that one can find $P \supset A$ with $d(K) = O(K)$ and $C(K) = e^{O(K)}$. See next section for other conjectures and an explanation why this is optimal.

Theorem 3. *Given an additive set $A \subset \mathbb{Z}$ with $|2A| \leq K|A|$, A is contained in a g.a.p. P which is proper, has dimension $d \leq [K-1]$, and whose cardinality satisfies*

$$\log \frac{|P|}{|A|} \leq CK^2(\log K)^3.$$

Along the way, Chang made an improvement to a structure theorem of Ruzsa, which has found applications independent of Freiman's theorem.

Theorem 4. *If $A \subset Z$ is an additive set and $|2A| \leq K|A|$, then $2A - 2A$ contains a proper symmetric g.a.p. P of dimension $d \leq O(K + K \log K)$ and cardinality $|P| \geq \exp(-O(K(1 + \log^2 K)))|A|$.*

11.2 Applications, remaining conjectures

Much of the material for this section was taken from [5], particularly chapter 5.

Freiman's theorem, particularly the quantitative bounds established by Chang have been useful in much of the recent literature on arithmetic combinatorics, including several of the papers discussed at this summer school, in particular, the articles of: Bourgain, Katz, and Tao on the sum-product estimate; Tao and Green on Gowers' $U_3(G)$ norm; Bourgain on the ring conjectures; and Green and Sanders on the idempotent theorem.

The following conjectures are still open:

Conjecture 5. *(Polynomial Freiman-Ruzsa) Suppose $A \subset \mathbb{Z}$ is an additive set satisfying $|2A| \leq K|A|$. Then there exists a g.a.p. P with*

$$\dim(P) \leq CK^{O(1)} \qquad |P| \leq CK^{O(1)}|A|$$

such that one has $|A \cap P| \geq cK^{-O(1)}|A|$.

Note: the conclusion is that $A \cap P$ is large, not that $A \subset P$. One can see that with the conclusion $A \subset P$, the best one could hope for is

$$\dim(P) \leq CK \qquad |P| \leq C\exp(CK)|A|$$

by considering (for instance) the set $2^{[0,K-1]} = \{1, 2, \ldots, 2^{K-1}\}$.

Conjecture 6. *(Gowers, see [3]) If $A \subset \mathbb{Z}$ is an additive set satisfying $|2A| \leq K|A|$ and $\varepsilon > 0$, then there exists a g.a.p. P with $\dim(P) \leq CK^{\varepsilon}$ such that $|P|$ is reasonably small and $|A \cap P|$ is reasonably large.*

In many of the above mentioned applications, a proof of either of these conjectures would yield substantially improved numerical results.

11.3 The proof of Theorem 2

Roughly the Chang/Ruzsa's argument breaks down into three major portions. First, reduce to the case when A is contained in $\mathbb{Z}_N$ for some prime N. Next, find a large proper g.a.p. $P_0 \subset 2A - 2A$ of small degree. Finally, use P_0 to produce the g.a.p. P (having small dimension and length) which contains A.

11.3.1 Reduction to $A \subset \mathbb{Z}_N$.

For details, particularly the proof of the theorems of Ruzsa and Plünecke, see [4].

Let A and B be subsets of additive groups Y and Z, respectively. Let $h \geq 1$ be an integer. A function $\phi : A \to B$ is a *Freiman homomorphism of order* h if whenever $a_1, \ldots, a_h, a_1{}', \ldots, a_h{}' \in A$,

$$a_1 + \cdots + a_h = a_1{}' + \cdots + a_h{}' \implies \phi(a_1) + \cdots + \phi(a_h) = \phi(a_1{}') + \cdots + \phi(a_h{}').$$

The function ϕ is a *Freiman isomorphism of order* h if ϕ is a bijection from A onto B and both ϕ and ϕ^{-1} are Freiman homomorphisms of order h

What's the point? We cannot identify non-trivial subsets of $\mathbb{Z}$ with subsets of $\mathbb{Z}_N$ while preserving arithmetic relations of all orders, but fortunately, for the purposes of Freiman's theorem, we only need to preserve arithmetic relations of low order and g.a.p.'s. For this one can show that Freiman isomorphisms suffice.

Ruzsa proved that one can indeed identify subsets of A with subsets of $\mathbb{Z}_N$ via Freiman isomorphisms:

Theorem 7. *Let $A \subset \mathbb{Z}$ be finite, non-empty. Let $h \geq 2$ and let $m \geq 4h|hA - hA|$. Then there exists $A' \subset A$ with*

$$|A'| \geq \frac{|A|}{h}$$

which is Freiman isomorphic of order h to $\mathbb{Z}_m$.

The remainder of the proof proceeds as follows: Identify $A' \subset A$ with a subset $R \subset \mathbb{Z}_N$ via a Freiman isomorphism ϕ of order 8. Find a g.a.p. $P_0 \subset 2R - 2R$. Then $Q_0 := \phi^{-1}(P_0)$ is a g.a.p. contained in $2A - 2A$. Use Q_0 to construct a g.a.p. containing A.

To obtain quantitative bounds on P_0 in the middle step, we will need for N to be small relative to $|A|$ and prime. Once we have achieved smallness of N, primality is easy since there is always a prime between N and $2N$. For smallness, looking back at the previous theorem, it suffices to show that $|2A|$ small implies $|8A - 8A|$ is small. For this, one uses Plünnecke's theorem, which states that if $|2A|$ is small relative to $|A|$, then the same holds for all difference sets $kA - lA$ with k, l non-negative integers.

11.3.2 Finding a progression in $2A - 2A$.

Suppose $A \subset \mathbb{Z}_N$, where N is prime, $|2A| \leq K|A|$ and $\delta := |A|/N > cK^{-16}$.

By a method due to Bogolyubov, given any subset $A \subset \mathbb{Z}_N$ (N prime), there is a proper g.a.p. $P_0 \subset 2A - 2A$ such that

$$\dim(P_0) \leq p_1(\delta^{-1}) \qquad \log(\frac{|P_0|}{N}) \geq p_2(\delta^{-1}),$$

for polynomials p_1 and p_2. In our case, this implies bounds which are polynomial in K. This was the method used by Ruzsa.

Chang modified Bogolyubov's argument to take advantage of the extra information $|2A| \leq K|A|$. This gave improved bounds of the same type as those above. We will briefly sketch the argument behind this improvement.

Two crucial concepts for this portion of the proof are *Bohr sets* and *dissociated sets*.

Let $S \subset \mathbb{Z}_N$ and $\varepsilon > 0$. Then the *Bohr neighborhood of S with radius ε* is defined as

$$B(S; \varepsilon) := \{x \in \mathbb{Z}_N : \|\frac{nx}{N}\|_{\mathbb{R}/\mathbb{Z}_N} < \varepsilon, \forall n \in S\}.$$

A subset $\{\lambda_1, \ldots, \lambda_d\}$ of an additive group Z is *dissociated* if whenever $\varepsilon_j \in \{-1, 0, 1\}$, $1 \leq j \leq d$ are not all zero,

$$\sum_j \varepsilon_j \lambda_j \neq 0.$$

For some choice of $c, \varepsilon \sim 1$, if one lets

$$\Gamma := \{x \in \mathbb{Z}_N : |\hat{1_A}(x)| \geq cK^{-1/2}\delta\}$$

and $B := B(\Gamma; \varepsilon)$, then $B \subset 2R - 2R$ and $d := |\Gamma|$ is small.

Then, we refine Γ to a maximal dissociated subset of $\Lambda \subset \Gamma$. An application of Rudin's theorem shows that such Γ has small size, and by the maximality of Γ,

$$B(\Lambda, \frac{\varepsilon}{|\Lambda|}) \subset B(\Gamma, \varepsilon).$$

At this stage, one may apply a theorem of Bogolyubov which guarantees the existence of proper g.a.p.'s of small dimension, large cardinality in Bohr sets. This allows one to produce a g.a.p. $P_0 \subset B \subset 2A - 2A$ which obeys the bounds in the structure theorem, Theorem 4.

11.3.3 From $P_0 \subset 2A - 2A$ to $P \supset A$

It is in the passage from the proper g.a.p. $P_0 \subset 2A - 2A$ to the (not necessarily proper) g.a.p. $P \supset A$ that Chang's argument offers the greatest numerical gains (from doubly to singly exponential) over Ruzsa's argument, though the two are similar in spirit. We describe the simpler argument of Ruzsa here because of limited space.

Ruzsa's argument: Let $\{a_1, \ldots, a_s\} \subset A$ to be a maximal set with the property that the sets $a_i + P_0$ are pairwise disjoint. One can show that $A \subset \{a_1, \ldots, a_s\} + P_0 - P_0 \subset P_1$ for some (not necessarily proper) g.a.p. P_1. If $d := \dim(P_0)$, then one can show that the g.a.p. P_1 satisfies

$$\dim(P_1) \leq s + d \qquad \ell(P_1) \leq 2^{s+d}|P_0|.$$

One can also show (using Plünnecke's inequality) that $s \leq C^d K^5 d^d$, which together with the bounds for d and $|P_1|$ described above, proves Freiman's theorem, but with non-optimal values of $d(K)$ and $|C(K)|$.

11.4 Producing a proper progression of small rank.

Starting from a g.a.p. P as described in Theorem 2, Chang can proves the existence of an g.a.p. P' satisfying the conclusions of Theorem 3 via a modified (to maintain the small cardinality bound) argument of Freiman, as described by Bilu in [1].

Roughly, this proceeds as follows. We have a g.a.p.

$$P = P(v_1, \ldots, v_d; l_1, \ldots, l_d; a).$$

We can assume a is zero. Associated to P are the homomorphism

$$\phi : \mathbb{Z}^d \to \mathbb{Z} \qquad \phi(e_j) = v_j$$

and the parallelogram $B := \prod_{j=1}^{d}[-l_i + 1, l_i - 1]$. Note that $\phi(B) \supset A$. If $\phi|_{2B\cap\mathbb{Z}^d}$ is not one-to-one, one can reduce the dimension by 1, eventually obtaining a homomorphism which *is* one-to-one on $2B$.

This homomorphism ϕ is then a 2-Freiman isomorphism from $2B \cap \mathbb{Z}^d$ to its image in $\mathbb{Z}$. One uses ϕ to pull A back to a subset $A' \subset \mathbb{Z}^d$, maintaining the inequality $|2A'| \leq K|A|$. The inequality implies that A' must be a $[K-1]$-dimensional set. Let Γ be the affine space spanned by A'. The g.a.p. P' is the image of $\phi|_{\Gamma\cap B\cap\mathbb{Z}^d}$. We still have $A \subset P'$, but now $\dim(P) \leq [K-1]$.

References

[1] Bilu, Y., "Structure of sets with small sumset" in *Structure Theory of Set Addition,* Astrisque 258, Soc. Math. France, Montrouge, 1999, 77–108.

[2] Chang, M.-C., *A polynomial bound in Freiman's theorem.* Duke Math. J. 113 (2002), no. 3, 399–419.

[3] Green, B. J., *Structure Theory of Set Addition*, Lecture notes, http://www-math.mit.edu/ green/notes.html.

[4] Nathanson, M. B., *Additive Number Theory: Inverse Problems and the Geometry of Sumsets,* Springer, New York, 1996.

[5] Tao, T. and Vu, V. H., *Additive Combinatorics,* Cambridge Univ. Press, Cambridge, 2006.

Betsy Stovall, UC Berkeley
email: **betsy@math.berkeley.edu**

12 Norm Convergence of Multiple Ergodic Averages of Commuting Transformations

after Terence Tao [2]
A summary written by Zhiren Wang

Abstract

In [2] it is shown that the averages $\frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^{l} f_i(T_i^n x)$ converge in $L^2(X)$ for commuting measure-preserving transformations $\{T_i\}_{i=1}^{l}$ and L^∞ functions $\{f_i\}_{i=1}^{l}$ on a probability $(X, \mathcal{X}, \mu)$. The proof is purely combinatorial by converting the original problem to a finitary one, where both the underlying space and the number of terms in the converging sequence which need to be studied are finite.

12.1 Introduction

Let $(X, \mathcal{X}, \mu)$ be a probability space and $T : X \mapsto X$ be a measure preserving transformation, i.e. T is $\mathcal{X}$-measurable and $T_*\mu = \mu$. The von Neumann mean ergodic theorem claims for any $f \in L^2(X, \mathcal{X}, \mu)$ the averages $\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)$ are convergent in $L^2(X)$ as $N \to \infty$.

The main theorem of the present work [2] deals with a generalization where there are more than one transformations and functions.

Theorem 1. *For $l \geq 1$, if $T_1, \cdots, T_l$ are commuting measure-preserving transformations in a probability space $(X, \mathcal{X}, \mu)$ and $f_1, \cdots, f_l \in L^\infty(X, \mathcal{X}, \mu)$ then $\{\frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^{l} f_i(T_i^n x)\}_{N=1}^{\infty}$ converge in $L^2(X, \mathcal{X}, \mu)$.*

When $l = 1$ this gives the conventional mean ergodic theorem. Many other special cases (with small l or additional assumptions on T_i) have been previously studied

12.2 Finitary versions of the main theorem

The first part of the proof reduces the main theorem in to a finitary statement. This is achieved in two steps.

Before we start, it should be remarked that $(X, \mathcal{X}, \mu)$ can be assumed to be ergodic under the $\mathbb{Z}^l$ action generated by $T_1, \cdots, T_l$, as it satisfies Theorem 1 as long as all members in its ergodic decomposition do, which can be easily demonstrated with dominated convergence theorem.

Notations: The average operator over a finite set Γ is denoted by $\mathbf{E}_{\gamma\in\Gamma}$, i.e. $\mathbf{E}_{\gamma\in\Gamma}f(\gamma) = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} f(\gamma)$. Note $[N] := \{0, \cdots, N-1\}$ for all $N \in \mathbb{N}$.

12.2.1 Finite type convergence statement

Theorem 1 is equivalent to the following fact, which allows us to calculate estimates for only a finite number of terms.

Theorem 2. *Given the same setting as in Theorem 1,* $\forall\epsilon > 0$, $\forall F : \mathbb{N} \mapsto \mathbb{N}$, $\exists M \in \mathbb{N}$ *such that*

$$\|\mathbf{E}_{n\in[N]}\prod_{i=1}^{l} f_i \circ T_i^n - \mathbf{E}_{n\in[N']}\prod_{i=1}^{l} f_i \circ T_i^n\|_{L^2(X)} < \epsilon, \forall N, N' \in [M, F(M)].$$

It's clear that Theorem 1 implies Theorem 2. To see the other direction, observe if Theorem 1 fails then $\exists\epsilon$, $\forall M \in \mathbb{N}$, $\exists F(M) > M$, $\|\mathbf{E}_{n\in[M]}\prod_{i=1}^{l} f_i \circ T_i^n - \mathbf{E}_{n\in[F(M)]}\prod_{i=1}^{l} f_i \circ T_i^n\|_{L^2(X)} > \epsilon$, which contradicts Theorem 2. This type of reduction works for general convergence statements.

12.2.2 Discretization of the space

In his ergodic proof ([1]) to Szemerédi theorem, Furstenberg introduced the Furstenberg correspondence principle which assigns a dynamical system to a combinatorial object. The current step is the reverse of that correspondence, similar to the implication from Szemerédi theorem to Furstenberg multirecurrence theorem.

Equip the finite additive group $\mathbb{Z}_P^l$ with uniform probability measure, denote by S_i the shift on the i-th coordinate: $S_i(a_1, \cdots, a_l) = (a_1, \cdots, a_{i-1}, a_i + 1, a_{i+1}, \cdots, a_l)$.

Theorem 2 is implied by the following

Theorem 3. *Fix* $l \in \mathbb{N}$, $\forall\epsilon > 0$, $\forall F : \mathbb{N} \mapsto \mathbb{N}$, $\exists M^* \in \mathbb{N}$ *such that the following property holds:*

if $P \in \mathbb{N}$, $\{f_i\}_{i=1}^{l}$ *are functions on the finite additive group* $\mathbb{Z}_P^l$ *whose values are bounded in* $[-1, 1]$, *then* $\exists M \leq M^*$ *such that*

$$\|\mathbf{E}_{n\in[N']}\prod_{i=1}^{l} f_i \circ S_i^n - \mathbf{E}_{n\in[N']}\prod_{i=1}^{l} f_i \circ S_i^n\|_{L^2(\mathbb{Z}_P^l)} < \epsilon, \forall N, N' \in [M, F(M)].$$

Proof of Theorem 3 $\Rightarrow$ Theorem 2: Fix $(X, \mathcal{X}, \mu)$, f_i's, T_i's, ϵ and F, apply Theorem 3 to $\frac{\epsilon}{2}$ and get a M^*.

For $v \in \mathbb{Z}^l$ and $x \in X$, let $T^v x = T_1^{v_1} T_2^{v_2} \cdots T_l^{v_l} x$. As we assumed the system is ergodic under the joint action of T_i's, for any $g \in L^\infty(X)$ there exists a null set Ω_g with $\mu(\Omega_g) = 0$ and $\mathbf{E}_{v\in[P]^l} g(T^v x) \to \int g \mathrm{d}\mu, \forall x \notin \Omega_g$ since $\{[P]^l\}_{P=1}^\infty$ is a Følner sequence in the amenable group $\mathbb{Z}^l$. Let $\mathcal{G}$ be the collection of functions which are rational polynomials of functions of the form $f_i \circ T_i^n$, then $\mathcal{G}$ is countable, thus we can pick a generic point $x_0 \notin \bigcup_{g\in\mathcal{G}} \Omega_g$.

$g_{N,N'}(x) := |\mathbf{E}_{n\in[N']} \prod_{i=1}^l f_i(T_i^n x) - \mathbf{E}_{n\in[N']} \prod_{i=1}^l f_i(T_i^n x)|^2 \in \mathcal{G}$, therefore for P large enough,

$$\mathbf{E}_{v\in[P]^l} g_{N,N'}(x_0) - \int g_{N,N'} \mathrm{d}\mu < \frac{\epsilon^2}{10}, \forall N, N' \leq F(M^*), \tag{1}$$

here we can suppose F is increasing.

Now we construct a new system on the space $\mathbb{Z}_P^l$. $\forall v \in \mathbb{Z}_P^l$, identify v with an element of $[P]^l \subset \mathbb{Z}^l$ in the obvious way, still noted by v. Let $f_i'(v) = f_i(T^v x_0)$. Without loss of generality assume $\|f_i\|_{L^\infty(X)} \leq 1, \forall i$. So Theorem 3 applies to $\{f_i'\}$: $\exists M \leq M^*$,

$$\|\mathbf{E}_{n\in[N']} \prod_{i=1}^l f_i' \circ S_i^n - \mathbf{E}_{n\in[N']} \prod_{i=1}^l f_i' \circ S_i^n\|_{L^2(\mathbb{Z}_P^l)} < \frac{\epsilon}{2}, \forall N, N' \in [M, F(M)]. \tag{2}$$

The square of the left hand side is approximately

$$\mathbf{E}_{v\in[P]^l} |\mathbf{E}_{n\in[N']} \prod_{i=1}^l f_i(T_i^n x_0) - \mathbf{E}_{n\in[N']} \prod_{i=1}^l f_i(T_i^n x_0)|^2 = E_{v\in[P]^l} g_{N,N'}(x_0). \tag{3}$$

The error arises from those terms with v near the boundary of the region $[P]^l$ and thus "wraparound" in $\mathbb{Z}_P^l$ under some shift S_i^n. As $n < \max(N, N') \leq F(M) \leq F(M^*)$, the contribution of these problematic terms is of order $O(\frac{F(M^*)}{P})$ and is at most $\frac{\epsilon^2}{10}$ when $P \geq O(\frac{F(M^*)}{\epsilon^2})$.

Fix such a sufficiently large P and compare (1),(2),(3), we get $\forall N, N' \in [M, F(M)]$,

$$(\int g_{N,N'} \mathrm{d}\mu)^{\frac{1}{2}} \leq (\mathbf{E}_{v\in[P]^l} g_{N,N'}(x_0) + \frac{\epsilon^2}{10})^{\frac{1}{2}} \leq \sqrt{(\frac{\epsilon}{2})^2 + \frac{\epsilon^2}{10} + \frac{\epsilon^2}{10}} < \epsilon,$$

which is exactly the claim in Theorem 2. $\square$

Notice

$$\prod_{i=1}^{l} f_i \circ S_i^n(v_1, \cdots, v_l) = \prod_{i=1}^{l} f_i^*(v_1, \cdots, v_l, -\sum_{i=1}^{l} v_i - n)$$

where

$$f_i^*(v_1, \cdots, v_l, v_{l+1}) := f_i(v_1, \cdots, v_{i-1}, -\sum_{1 \le j \le l+1, j \ne i} v_j, v_{i+1}, \cdots, v_l)$$

is a function on $\mathbb{Z}_P^{l+1}$. f_i^* depends on only l coordinates of $\mathbb{Z}_P^{l+1}$.

Definition 4. *Let $(\Omega, \mathcal{O}, \eta)$ be a probability space. On $\mathbb{Z}_P^{l+1} \times \Omega$, a* **elementary function of complexity** *(d, J) is a function which can be expressed as*

$$\sum_{j=1}^{J} \prod_{e \subset \{1, \cdots, l+1\}, |e| = d} g_{e,j}$$

where $g_{e,j} : \mathbb{Z}_P^{l+1} \times \Omega \mapsto [-1, 1]$ depends only on the d coordinates inside e and on the Ω coordinate.

Definition 5. *$\forall g : \mathbb{Z}_P^{l+1} \times \Omega \mapsto \mathbb{R}, \forall N \in \mathbb{N}$, define $\triangle_N g : \mathbb{Z}_P^l \times \Omega \mapsto \mathbb{R}$ by $\triangle_N g(v, \omega) := \mathrm{E}_{n \in [N]} g((v, -\sum_{i=1}^{l} v_i - n), \omega)$.*

It's clear that $\prod_{i=1}^{l} f_i^*$ is an elementary function of complexity $(l, 1)$ on $\mathbb{Z}_P^{l+1}$ (with $\Omega = \{point\}$). Thus Theorem 3 is a special case of

Theorem 6. *Fix $l \in N$, $M_* \ge 1$, $0 \le d \le l$ and $J \ge 1$. Then $\forall \epsilon > 0$, $\forall F : \mathbb{N} \mapsto \mathbb{N}$, $\exists M^* \in \mathbb{N}$ such that the following property holds:*

$\forall P \in \mathbb{N}$, $\forall (\Omega, \mathcal{O}, \eta)$, if $g : \mathbb{Z}_P^{l+1} \times \Omega \mapsto \mathbb{R}$ is an elementary function of complexity (d, J) then $\exists M_* \le M \le M^*$ such that

$$\|\triangle_N g - \triangle_{N'} g\|_{L^2(\mathbb{Z}_P^l \times \Omega)} \le \epsilon, \forall N, N' \in [M, F(M)].$$

We remark that in Theorem 3 and Theorem 6, the key feature is that M^* is independent of: the functions f_i (or g), the extrinsic probability space Ω and most importantly, the scale P of the space.

Ω, J and M_* appear in the statement merely because of technical needs in a later inductive argument. They can be dropped without weakening Theorem 6.

Theorem 7. *Fix $l \in N$ and $0 \leq d \leq l$. Then $\forall \epsilon > 0$, $\forall F : \mathbb{N} \mapsto \mathbb{N}$, $\exists M^* \in \mathbb{N}$ such that the following property holds:*

$\forall P \in \mathbb{N}$, if $g : \mathbb{Z}_P^{l+1} \mapsto \mathbb{R}$ is an elementary function of complexity $(d, 1)$ then $\exists M \leq M^*$ such that

$$\|\triangle_N g - \triangle_{N'} g\|_{L^2(\mathbb{Z}_P^l)} \leq \epsilon, \forall N, N' \in [M, F(M)].$$

Theorem 7 is equivalent to Theorem 6. In fact, the space Ω can be ignored because of an elegant theorem ([2], Theorem A.2) which is basically Lebesgue dominated convergence theorem, but translated into the finitary language appeared in Theorem 2. By definition a function of complexity (d, J) can be written as the sum of J functions of complexity $(d, 1)$, as $\triangle_N$ and $\triangle_{N'}$ are linear operators the L^2 norm bound for the (d, J) case is at most $J^{\frac{1}{2}}$ times as large as that for $(d, 1)$. By adjusting ϵ, one can reduce to the case $J = 1$. Finally, M_* can also be supposed to be 1 as we can replace F by $F_{M_*}(M) = F(max(M, M_*))$, M_* by 1 and apply the theorem, then the M^* we get would also work for the original problem.

12.3 Sketch of proof

We shall prove Theorem 6 & 7 by induction on d. In the original paper [2] the induction starts at $d = 1$ but there would be no problem to begin with $d = 0$, in which case g is a constant and the theorem is trivial. For larger d, we are going to assume Theorem 6 for $d' = d - 1$ and deduce Theorem 7.

We want some control over the rate of convergence of $\triangle_N(g)$ which is uniform in P. g can be roughly decomposed into two parts, one is "locally flat"(called anti-uniform) and the other is oscillatory(called uniform). The uniform part would be nice for us because of its self-cancellation when averaged over long intervals. The anti-uniform part behaves less friendly but surprisingly it can be restricted in at most a constant finite number ($O_\epsilon(1)$, independent of P) of scale levels so that the left-over is oscillatory enough for an error smaller than ϵ. After that we shall approximate anti-uniform functions of these scale levels with elementary functions of complexity $(d - 1, J)$ and make use of the inductive hypothesis.

12.3.1 Koopman-von Neumann type decomposition

Let $g = \prod_{e \subset \{1, \cdots, l+1\}, |e| = d} g_e$ be an elementary function of complexity $(d, 1)$ where $g_e : \mathbb{Z}_P^{l+1} :\mapsto [-1, 1]$ is a function that depends only on the coordinates

inside e. Since if $l+1 \notin e$ then $\triangle_N g_e h = g_e \triangle_N h, \forall N, h$, in the proof of Theorem 7 we can suppose $g_e = 1$ when $l+1 \notin e$, i.e. $g = \prod_{e \in I} g_e$ where $|e| = d, l+1 \in e, \forall e \in I \subset \{1, \cdots, l+1\}$.

Definition 8. *Let $M \in \mathbb{N}$ and let $e \in I$, an* **basic e-anti-uniform function of scale** *M is a function $\phi_e : \mathbb{Z}_P^{l+1}$ of the form*

$$\phi_e(v) = \mathbf{E}_{m \in [M]} \prod_{i \in e} b_i\big((v_j)_{j \in e, j \neq i}, \sum_{k \in e} v_k + m\big).$$

where b_i is a function $\mathbb{Z}_P^{e \setminus \{i\}} \times \mathbb{Z}_P \mapsto [-1, 1]$.

A basic e-anti-uniform function depends only on coordinates in e.

Lemma 9. *For $M \geq 1$ and $\epsilon > 0$, if $\|\triangle_N g\|_{L^2(\mathbb{Z}_P^l)} > \epsilon$ for some $N \geq \frac{10M}{\epsilon^2}$, then $\forall e_0 \in I$, there is a basic e_0-anti-uniform function ϕ_{e_0} of scale M such that $|\langle g_{e_0}, \phi_{e_0} \rangle|_{L^2(\mathbb{Z}_P^l)} \geq \frac{\epsilon^2}{2}$.*

This lemma is crucial. It asserts that "uniform" functions, or functions approximately orthogonal to anti-uniform ones, become small in norm under average operators.

Now let $K = \lceil \frac{10^6 (2^{l+1})^5}{\epsilon^4} \rceil \geq \frac{10^6 |I|^5}{\epsilon^4}$ and $1 = M_1 \leq M_2 \leq \cdots \leq M_K$ be defined by $M_{k+1} = \tilde{F}(M_k)$ where $F : \mathbb{N} \mapsto \mathbb{N}$ will be determined later.

Theorem 10. *$\exists 2 \leq k \leq K+1$ and a decomposition $g_e = g_{e,U^\perp} + g_{e,U}$ for each $e \in I$, where $g_{e,U^\perp}, g_{e,U} : \mathbb{Z}_P^{l+1} \mapsto [-1, 1]$ depend only on coordinates inside e and satisfy:*
(i) (anti-uniform part) $\forall e \in I$, $\forall k \leq j \leq K$, there is a basic e-anti-uniform function $\phi_{e,j}$ and a polynomial Ψ_e whose degree and coefficients are both bounded by $O_\epsilon(1)$ (i.e., independent of P and g), such that

$$\begin{aligned} \|g_{e,U^\perp} - \Psi_e(\phi_{e,k}, \cdots, \phi_{e,K})\|_{L^1(\mathbb{Z}_P^{l+1})} &\leq \tfrac{\epsilon^2}{400|I|^2}, \\ \|g_{e,U^\perp} - \Psi_e(\phi_{e,k}, \cdots, \phi_{e,K})\|_{L^\infty(\mathbb{Z}_P^{l+1})} &\leq 1. \end{aligned} \tag{4}$$

(ii) (uniform part) $\forall e \in I$, $\forall N \geq \frac{1000|I|^2 M_{k-1}}{\epsilon^2}$, if a function $h_{e'} : \mathbb{Z}_P^{l+1} \mapsto [-1, 1]$ which only depends on the cooordinates inside e' is given for every $e' \neq e, |e'| = d$ then

$$\|\triangle_N (g_{e,U} \prod_{e' \neq e, |e'| = d} h_{e'})\| \leq \frac{\epsilon}{10|I|}. \tag{5}$$

The idea of the proof of Theorem 5 is simple. Initialize $k = K + 1$ and $g_{e,U} = g_e$, $g_{e,U^\perp} = 0$. During each step check if (5) holds for all $N \geq \frac{1000|I|^2 M_{k-1}}{\epsilon^2}$ and all e. If yes then stop the algorithm. Otherwise it means for some $e \in I$ there is a basic e-anti-uniform function $\phi_{e,k-1}$ which is highly correlated with $g_{e,U}$. In this case reallocate the part inside g_e, U which is correlated with $\phi_{e,k-1}$ to $g_{e,U^\perp}$, decrease k by 1 and repeat the step. This involves some complication as the reallocation is not by calculating an inner product but by take the conditional expectation (i.e. an orthogonal projection) of g_e with respect to a finitely generated σ-algebra which is generated by truncations of $\phi_{e,j}$'s with $j \geq k$. Lemma 3.6 in [2] constructs these truncations and the σ-algebra explicitly, guarantees that (4) holds.

The main problem here is if the algorithm can stop in a constant number of steps. In fact, each time when we decrease k by 1, $\|g_{e,U^\perp}\|_{L^2}$ increases at least by $(\frac{\epsilon^2}{200|I|^2})^2$ for one of the e's according to Lemma 9. However as $\|g_{e,U^\perp}\|$ is an orthogonal projection of g_e, $\|g_{e,U^\perp}\|_{L^2} \leq 1$. So $\sum_{e\in I} \|g_{e,U^\perp}\|_{L^2} \leq |I| < K \cdot (\frac{\epsilon^2}{200|I|^2})^2$. Thus the processus always stops before k reaches 1.

12.3.2 Inductive step

First of all we can get rid of the uniform part of g. Let $M_{**} = \lceil \frac{1000|I|^2 M_{k-1}}{\epsilon^2} \rceil$. Then for any $N \geq M_{**}$, $\triangle_N(\prod_{e\in I} g_e)$ can be expanded as the sum of $\triangle_N(\prod_{e\in I} g_{e,U^\perp})$ with $|I|$ terms, each of which is of the form $\triangle_N(g_{e,U} \prod_{e'\neq e, |e'|=d} h_{e'})$ for some $e \in I$. By (5), each of these $|I|$ terms is of L^2 norm at most $\frac{\epsilon}{10|I|}$, thus $\|\triangle_N(\prod_{e\in I} g_e) - \triangle_N(\prod_{e\in I} g_{e,U^\perp})\|_{L^2(\mathbb{Z}_P^l)} \leq \frac{\epsilon}{10}$. By triangle inequality it suffices to find a constant M^{**} such that $\exists M_{**} < M < M^{**}, \|\triangle_N(\prod_{e\in I} g_{e,U^\perp}) - \triangle_{N'}(\prod_{e\in I} g_{e,U^\perp})\|_{L^2(\mathbb{Z}_P^l)} \leq \frac{\epsilon}{2}, \forall N, N' \in [M, F(M)]$.

With a similar argument (4) implies

$$\|\triangle_N(\prod_{e\in I} g_{e,U^\perp}) - \triangle_N(\prod_{e\in I} \Psi_e(\phi_{e,k}, \cdots, \phi_{e,K}))\|_{L^1(\mathbb{Z}_P^l)} \leq \frac{\epsilon^2}{400|I|^2} \cdot |I|,$$

$$\|\triangle_N(\prod_{e\in I} g_{e,U^\perp}) - \triangle_N(\prod_{e\in I} \Psi_e(\phi_{e,k}, \cdots, \phi_{e,K}))\|_{L^\infty(\mathbb{Z}_P^l)} \leq |I|,$$

(here we used the fact that $\triangle_N$ contracts norms). Thus

$$\|\triangle_N(\prod_{e\in I} g_{e,U^\perp}) - \triangle_N(\prod_{e\in I} \Psi_e(\phi_{e,k}, \cdots, \phi_{e,K}))\|_{L^2(\mathbb{Z}_P^l)} \leq \frac{\epsilon}{20}.$$

Again by triangle inequality it suffices to find a constant M^{**} such that $\exists M_{**} < M < M^{**}, \forall N, N' \in [M, F(M)]$

$$\|\triangle_N(\prod_{e\in I}\Psi_e(\phi_{e,k},\cdots,\phi_{e,K}))-\triangle_{N'}(\prod_{e\in I}\Psi_e(\phi_{e,k},\cdots,\phi_{e,K}))\|_{L^2(\mathbb{Z}_P^l)} \leq \frac{\epsilon}{2}. \quad (6)$$

Denote $h = \Psi_e(\phi_{e,k},\cdots,\phi_{e,K}$, set $M^{**} = \lceil F^{-1}(M_k^{\frac{1}{4}})\rceil$ (we assume as before F is increasing). Then $M_{k-1} < M_{**} < M^{**} < M_k$ as long as $\tilde{F}$ grows sufficiently faster than F does.

Let $L = \lceil M_k^{\frac{1}{2}}\rceil$, we now approximate h with elementary functions of complexity $d-1$.

Lemma 11. *For $v \in \mathbb{Z}_P^l$, $w \in [L]^l$ and $n \in [N] \cup [N']$,*

$$h(v+w, -\sum_{i=1}^{l}(v_i+w_i)-n) = \mathbf{E}_{\omega\in\Omega}\phi_{v,\omega}(w, -\sum_{i=1}^{l}w_i - n) + O_\epsilon(M_k^{-\frac{1}{2}}),$$

where Ω is a finite probability space and $\forall\omega\in\Omega, v\in\mathbb{Z}_P^l$, $\phi_{v,\omega}:\mathbb{Z}^l\mapsto\mathbb{R}$ is an elementary function of complexity $(d-1, O_\epsilon(1))$.

Sketch of proof. h can be expanded, each term is a product of many factors $\phi_{e,j}$. Again expand

$$\begin{aligned} & \phi_{e,j}(v+w, -\textstyle\sum_{i=1}^{l}(v_i+w_i)-n) \\ = \; & \mathbf{E}_{m_j\in[M_j]}\textstyle\prod_{i\in e} b_{e,i,j}((v_s+w_s-n\delta_{s,l+1})_{s\in e\backslash\{i\}}, \sum_{s\in e}(v_s+w_s)-n+m_j) \end{aligned}$$

where $v_{l+1} := -\sum_{i=1}^{l} v_i$, where $w_{l+1} := -\sum_{i=1}^{l} w_i$.

The key feature here is that $\|w\| \leq O(L) \leq O(M_k^{\frac{1}{2}}) \leq O(M_j^{\frac{1}{2}})$, $n \leq F(M) \leq F(M^{**}) \simeq M_k^{\frac{1}{4}}$ so we can replace $\sum_{s\in e}(v_s+w_s)-n+m_j$ by $\sum_{s\in e} v_s - n$ in the last coordinate and get an error term at most $O(\frac{|\sum_{s\in e} w_s - n|}{M_j}) \leq O(M_K^{-\frac{1}{2}})$ as most terms remained in the summation after the shift. Then for a fixed v, the $b_{e,i,j}$ factor depends only on the coordinates in $e\backslash\{i\}$. Thus the product of such factors are of complexity $d-1$.

The rest of the proof adds up these products and is straightforward. The probability space Ω represents averages over products of intervals of the form $[M_j]$ and the $O_\epsilon(1)$ in the complexity expression comes from the coefficients of the Ψ_e's. $\square$

$$
\begin{aligned}
& \|\triangle_N h - \triangle_{N'} h\|_{L^2(\mathbb{Z}_P^l)} \\
= \; & (\mathbf{E}_{v\in\mathbb{Z}_P^l}\mathbf{E}_{w\in[L]^l}|(\mathbf{E}_{n\in[N]} - \mathbf{E}_{n\in[N']})h(v+w, -\textstyle\sum_{i=1}^l (v_i+w_i) - n)|^2)^{\frac{1}{2}} \\
= \; & \big(\mathbf{E}_{v\in\mathbb{Z}_P^l}\mathbf{E}_{w\in[L]^l}|(\mathbf{E}_{n\in[N]} - \mathbf{E}_{n\in[N']}) \circ \mathbf{E}_{\omega\in\Omega}\phi_{v,\omega}(w, -\textstyle\sum_{i=1}^l w_i - n)|^2\big)^{\frac{1}{2}} + O_\epsilon(M_k^{-\frac{1}{4}}) \\
= \; & \big(\mathbf{E}_{v\in\mathbb{Z}_P^l,\omega\in\Omega}\mathbf{E}_{w\in[L]^l}|(\mathbf{E}_{n\in[N]} - \mathbf{E}_{n\in[N']}) \circ \phi_{v,\omega}(w, -\textstyle\sum_{i=1}^l w_i - n)|^2\big)^{\frac{1}{2}} + O_\epsilon(M_k^{-\frac{1}{4}})
\end{aligned}
$$

Consider $\mathbb{Z}_P^l \times \Omega$ as an extrinsic probability space and identify $[L]^l$ in the obvious way with a subset of $\mathbb{Z}_Q^l$ where $Q = (l+1)L$, still denoted by $[L]^l$. Like in Theorem 3, we construct a function $\tilde{\phi}_{v,\omega}$ on the new space $\mathbb{Z}_Q^{l+1}$: let $\tilde{\phi}_{v,\omega}(w_1, \cdots, w_{l+1})$ equal $\phi_{v,\omega}(w_1, \cdots, w_{l+1})$ if $(w_1, \cdots, w_{l+1}) \in [L]^l$ and $w_{l+1} \in [-Q, -1]$, equal 0 otherwise. Then $\tilde{\phi}_v, m$ is an elementary function of complexity $(d-1, O_\epsilon(1))$ as well as ϕ_v, m.

$$
\|\triangle_N h - \triangle_{N'} h\|_{L^2(\mathbb{Z}_P^l)} \leq \|\triangle_N \tilde{\phi}_{v,\omega} - \triangle_{N'} \tilde{\phi}_{v,\omega}\|_{L^2(\mathbb{Z}_Q^l \times \Omega)} + O_\epsilon(M_k^{-\frac{1}{4}}).
$$

Now we can apply the inductive hypothesis (Theorem 6) to $\tilde{\phi}_{v,\omega}$ on $\mathbb{Z}_Q^{l+1} \times \Omega$. There is a constant $O_{\epsilon,F,M_{**}}(1)$ such that

$$
\|\triangle_N \tilde{\phi}_{v,\omega} - \triangle_{N'} \tilde{\phi}_{v,\omega}\|_{L^2(\mathbb{Z}_Q^l \times \Omega)} \leq \frac{\epsilon}{4}, \forall N, N' \in [M, F(M)]
$$

for some $M_{**} \leq M \leq O_{\epsilon,F,M_{**}}(1)$. When $\tilde{F}$ grows rapidly enough, $M^{**} = \lceil F^{-1}(\tilde{F}(M_{k-1})^{\frac{1}{4}})\rceil$ is larger than the constant C and $O_\epsilon(M_k^{-\frac{1}{4}})$ is smaller then $\frac{\epsilon}{4}$, which implies (6) is satisfied and completes the proof of the main theorem.

References

[1] Furstenberg, H., *Ergodic behavior of diagonal measures and a theorem of Szeméredi on arithmetic progressions.* J. d'Analyse Math. 31 (1977), 204-256

[2] Tao, T., *Norm Convergence of Multiple Ergodic Averages for Commuting Transformations.* arxiv:0707.1117.

Zhiren Wang, Princeton University
email: **zhirenw@math.princeton.edu**