



Technion IIT  
Department of Computer Science  
Spring 2008–9

Course 236610:

**RESEARCH LABORATORY IN  
THEORETICAL COMPUTER SCIENCE**

Eli Ben-Sasson

## Preface

In our research laboratory we plan to study selected results from the field of additive combinatorics and some of their applications to theoretical computer science. A tentative list of results includes

- Roth's Theorem about the existence of arithmetic progressions of length 3 in dense subsets of  $\mathbb{Z}_p^m$ .
- The Gowers norm and Samorodnitsky's Theorem about testing quadratic functions over  $\mathbb{F}_2$ .
- The Bogdanov-Viola pseudorandom generator for low-degree polynomials.
- Sum-product theorems in finite fields and extractors for independent sources.

# Table of Contents

## Lecture 1: Introduction

---

## Lecture 2: Fourier Analysis over $\mathbb{Z}_n^m$

|     |   |     |
|-----|---|-----|
| 2.1 | Characters and Fourier representation . . . . .     | 2-1 |
| 2.2 | Fourier analysis of a homomorphism tester . . . . . | 2-2 |

---

## Lecture 3: Roth's Theorem

|     |   |     |
|-----|---|-----|
| 3.1 | The Fourier representation of $A$ . . . . .   | 3-1 |
| 3.2 | Density increment . . . . .                   | 3-2 |
| 3.3 | Completing the proof of Theorem 3.1 . . . . . | 3-3 |

---

## Lecture 4: Fooling low-degree polynomials

|     |                                   |     |
|-----|-----------------------------------|-----|
| 4.1 | Fooling sets . . . . .            | 4-1 |
| 4.2 | $\epsilon$ -biased sets . . . . . | 4-2 |

---

## Lecture 5: Fooling low-degree polynomials part II

|     |   |     |
|-----|---|-----|
| 5.1 | Directional derivatives . . . . .                                   | 5-1 |
| 5.2 | Fooling polynomials with large bias . . . . .                       | 5-2 |
| 5.3 | Fooling polynomials with small bias . . . . .                       | 5-2 |
| 5.4 | Completing the proof of Theorem 5.1 . . . . .                       | 5-3 |
| 5.5 | Approximating biased functions by majority of derivatives . . . . . | 5-3 |

---

## Lecture 6: On sets that are roughly closed under pairwise addition

|     |  |     |
|-----|--|-----|
| 6.1 | Global and local properties . . . . .        | 6-1 |
| 6.2 | The Balog-Szemerédi-Gowers Theorem . . . . . | 6-2 |

---

## Lecture 7: On sets that are roughly closed under addition, part II

|     |   |     |
|-----|---|-----|
| 7.1 | Bounding the size of $4A$ . . . . .                   | 7-1 |
| 7.2 | Bounding the size of $kA$ for arbitrary $k$ . . . . . | 7-2 |

---

**Lecture 8: Samorodnitsky’s quadratic low-degree test**

|     |   |     |
|-----|---|-----|
| 8.1 | A “low-degreeness” test . . . . .                           | 8–1 |
| 8.2 | The Fourier spectrum of derivatives of quadratics . . . . . | 8–2 |
| 8.3 | From Fourier analysis to additive combinatorics . . . . .   | 8–3 |

---

**Lecture 9: Sum-Product Theorems over finite fields and independent-source extractors**

|     |   |     |
|-----|---|-----|
| 9.1 | Sum-Product Theorems . . . . .                      | 9–1 |
| 9.2 | An explicit expanding rational expression . . . . . | 9–2 |

---

**Lecture 10: Sum-Product Theorems over finite fields and independent-source extractors**

|      |   |      |
|------|---|------|
| 10.1 | Bounding expressions that involve an operator and its inverse . . . . . | 10–1 |
| 10.2 | Bounding expressions that involve addition and multiplication . . . . . | 10–2 |

---

**References**

## LECTURE 1

### INTRODUCTION

MARCH 16TH, 2009

---

LECTURER: Eli Ben-Sasson

SCRIBE: Not applicable

In this lecture we surveyed some of the basic questions addressed in additive combinatorics, such as sum-product estimates and the existence of arithmetic progressions in dense subsets of the integers. A good reference point for this lecture (and for the remainder of the course) is the book titled Additive Combinatorics [Tao and Vu \[2006\]](#) as well as the minicourse on [Additive Combinatorics and Computer Science](#) given at Princeton University. The webpage for this course includes lecture notes, powerpoint presentations and online videos of the lectures given there.

In this lecture we briefly surveyed the basics of Fourier analysis over  $\mathbb{Z}_n^m$  and sketched the proof of the homomorphism test of Blum et al. [1990] using Fourier analysis. Below we recount the essential information regarding the Fourier transform.

## 2.1 Characters and Fourier representation

Throughout this lecture let  $Z = \mathbb{Z}_n^m$  for  $n \geq 2$  and  $m \geq 1$ . We are interested in studying functions from  $Z$  to the complex numbers  $\mathbb{C}$ . The set of all such functions forms a  $|Z|$ -dimensional vector space over  $\mathbb{C}$ , namely, the space  $\mathbb{C}^Z$ . We now define a set of functions that will later on be shown to be a basis for  $\mathbb{C}^Z$ . In what follows we let  $\omega_n$  denote the primitive  $n^{\text{th}}$  complex root of unity  $e^{2\pi i/n}$ .

**Definition 2.1** (Character). For  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{Z}_n^m$  let  $\chi_\alpha \in \mathbb{C}^Z$  be the function defined by

$$\chi_\alpha(z) = \omega_n^{\sum_{j=1}^m \alpha_j z_j}, \quad z = (z_1, \dots, z_m) \in \mathbb{Z}_n^m.$$

The character  $\chi_0 = 1^Z$  is called the *zero* character (sometimes known as the *trivial* character) and all other characters are said to be nonzero (or *nontrivial*).

One useful property of characters is that each one of them is a homomorphism from the additive group  $(\mathbb{Z}_n^m, +)$  to the cyclic multiplicative group  $(\langle \omega_n \rangle, \times)$  where  $\langle \omega_n \rangle = \{\omega_n^0, \dots, \omega_n^{n-1}\}$  is the set of complex roots of unity of order  $n$ . Another important property is that they form an orthonormal basis for  $\mathbb{C}^Z$  with respect to the inner-product  $\langle \cdot, \cdot \rangle$  defined for  $f, g \in \mathbb{C}^Z$  as

$$\langle f, g \rangle = \mathbf{E}_{x \in Z} [f(x) \cdot \overline{g(x)}],$$

where the expectation is taken with respect to the uniform distribution over  $Z$ . Recall that such an inner-product induces a *norm* over  $\mathbb{C}^Z$  given by  $\|f\| = \sqrt{\langle f, f \rangle}$ .

**Proposition 2.2** (Orthonormality of characters). *The characters of  $Z$  form an orthonormal set, i.e.,*

$$\langle \chi_\alpha, \chi_\beta \rangle = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases}.$$

>From this claim, a few useful properties of characters emerge.

**Inversion formula** Since there are  $|Z|$  distinct characters, **Proposition 2.2** implies the characters form a basis for  $\mathbb{C}^Z$ , which means that every  $f \in \mathbb{C}^Z$  can be represented uniquely

as a linear combination of characters:

$$f = \sum_{\alpha \in Z} \hat{f}_\alpha \cdot \chi_\alpha, \quad \text{where } \hat{f}_\alpha \in \mathbb{C}. \quad (1)$$

The previous equation is known as the *Fourier inversion formula* and the coefficient  $\hat{f}_\alpha$  is called the  $\alpha$ -*Fourier coefficient* of  $f$ , it is computed by “projecting” the vector  $f$  onto the vector  $\chi_\alpha$ :

$$\hat{f}_\alpha = \langle f, \chi_\alpha \rangle.$$

We let  $\hat{f} \in \mathbb{C}^Z$  denote the vector of Fourier coefficients of  $f$  which is also known as the *Fourier representation* of  $f$ .

**Plancherel’s equality** Moving from a “standard” representation of functions  $f, g \in \mathbb{C}^Z$  to their Fourier representations amounts to representing the same vectors under two different orthonormal bases. Thus, we don’t expect the angle between the vectors to change under this change of basis. Indeed, *Plancherel’s equality* says that for all  $f, g \in \mathbb{C}^Z$  we have

$$\langle f, g \rangle = |Z| \cdot \langle \hat{f}, \hat{g} \rangle. \quad (2)$$

**Parseval’s equality** In particular, taking  $g = f$  in [Equation 2](#) we also see that the norm of  $f$  does not change under the two representations. Formally, *Parseval’s equality* says that for all  $f \in \mathbb{C}^Z$ ,

$$\mathbf{E}_{x \in Z} [ |f(x)|^2 ] = \sum_{\alpha \in Z} | \hat{f}_\alpha |^2. \quad (3)$$

In particular, if the range of  $f$  is  $\langle \omega_n \rangle$  (as will soon be the case) then this equality implies  $\sum_{\alpha \in Z} | \hat{f}_\alpha |^2 = 1$ .

We end by stating the very useful Cauchy-Schwarz inequality, that will be used time and again in future sessions. It says that for all vectors  $u, v$  of an inner product space,

$$| \langle u, v \rangle |^2 \leq \langle u, u \rangle \cdot \langle v, v \rangle. \quad (4)$$

## 2.2 Fourier analysis of a homomorphism tester

In the second part of this session we discussed the “homomorphism testing” problem introduced in [Blum et al. \[1990\]](#) and its connections to program checking, probabilistically checkable proofs and property testing. We sketched the main steps in the Fourier-analytic approach to the problem, due to [Bellare et al. \[1995\]](#). For explicit details of the simplest case of linearity testing over  $\mathbb{Z}_2^m$  we refer the reader to the original paper [Bellare et al. \[1995\]](#) as well as [\[Ben-Sasson, Fall 2006, Lectures 3–4\]](#). For the case of  $\mathbb{Z}_p^m, p > 2$  discussed in class see [Håstad and Wigderson \[2003\]](#); [Grigorescu et al. \[2006\]](#).

Roth's theorem says that "dense" subsets of the integers contain arithmetic progressions of length 3. Formally, it says that for every  $\delta > 0$  and sufficiently large  $n$  the following holds. Every subset  $A$  of  $\mathbb{Z}_n$  that has density  $|A|/n \geq \delta$  must contain an arithmetic progression of length 3. In this session we shall prove the following analogous statement, due to Meshulam [1995], for the simpler case of  $A \subseteq \mathbb{Z}_p^m$  where  $p$  is an odd prime. In what follows, a *arithmetic progression* of length  $k$  is a set  $\{a + jr \mid j = 0, \dots, k - 1\}$  where  $r \neq 0$ .

**Theorem 3.1** (Roth's theorem for  $\mathbb{Z}_p^m$ ). *Let  $p$  be an odd prime and  $m$  an integer. Every set  $A \subseteq \mathbb{Z}_p^m$  of size at least  $3p^m/m$  contains an arithmetic progression of length 3.*

The bound stated above is better than what is known for the case of integers in the sense that the minimal density required of  $A$  for ensuring it contains a length 3 progression is  $1/m = 1/\log_p |\mathbb{Z}_p^m|$ . In comparison, the smallest density for which Roth's theorem is known to be true over the integers, as shown in Bourgain [1999], is  $\sqrt{\log \log |\mathbb{Z}_n| / \log |\mathbb{Z}_n|}$ . Our proof will closely follow the exposition in [Tao and Vu, 2006, Section 10.2]. Two interesting aspects of the proof are the crucial role played by the Fourier representation of functions and the "density-increment" argument discussed in Section 3.2.

### 3.1 The Fourier representation of $A$

In this section we show that if  $A$  does not contain an arithmetic progression of length 3 then the indicator function of the set  $A$  will contain a nontrivial Fourier coefficient of large magnitude. We start by defining the *indicator function* of a set  $B \subseteq \mathbb{Z}_p^m$  as the function

$$\mathbf{1}_B : \mathbb{Z}_p^m \rightarrow \{0, 1\}, \quad \mathbf{1}_B(x) = \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

**Problem 3.1.** Letting  $B = 2A = \{2a \mid a \in A\}$ , notice that the oddness of  $p$  implies  $|B| = |A|$ . Prove:

$$\Pr_{x,r \in \mathbb{Z}_p^m} [x, x+r, x+2r \in A] = \mathbf{E}_{x,y \in \mathbb{Z}_p^m} [\mathbf{1}_A(x) \cdot \mathbf{1}_B(x+y) \cdot \mathbf{1}_A(y)].$$

**Problem 3.2.** Let  $\mu = |A|/p^m$  denote the density of  $A$  and let  $\epsilon$  denote the magnitude of the largest nontrivial Fourier coefficient of  $\mathbf{1}_A$ , i.e.,  $\epsilon = \max \left\{ \left| (\widehat{\mathbf{1}_A})_\alpha \right| \mid \alpha \neq 0 \right\}$ . Prove: If  $A$  contains no arithmetic progression of length 3, then there must exist a nonzero Fourier

coefficient of magnitude at least  $\mu^2 - p^{-m}$ . To this end, prove the following inequality

$$\left| \mathbf{E}_{x,y \in \mathbb{Z}_p^m} [\mathbf{1}_A(x) \cdot \mathbf{1}_B(y) \cdot \mathbf{1}_A(-(x+y))] - \mu^3 \right| \leq \epsilon \mu. \quad (5)$$

Hints regarding the proof of the inequality:

- Express all functions mentioned inside the expectation using the inversion formula given in [Equation 1](#). What is  $(\widehat{\mathbf{1}_A})_0$ ?
- Use [Proposition 2.2](#) to show that this expectation equals  $\sum_{\alpha \in \mathbb{Z}_p^m} (\widehat{\mathbf{1}_A})_\alpha^2 \cdot (\widehat{\mathbf{1}_B})_\alpha$ .
- Use Parseval's equality ([Equation 3](#)) to complete the proof.

## 3.2 Density increment

The *bias* of a function  $f : S \rightarrow \mathbb{R}$  is defined to be  $|\mathbf{E}_{x \in \mathbb{Z}_p^m} [f(x)]| = |\widehat{f}_0|$  and a function with bias 0 is said to be *unbiased*. The heart of the proof of Roth's theorem is the observation that a unbiased function  $f$  with large nonzero Fourier coefficient, say, of magnitude greater than  $\epsilon$ , must have "large" bias  $\geq \epsilon/2$  on a subspace of  $\mathbb{Z}_p^m$  of codimension 1. This observation will be used in the next section to argue, by way of contradiction, that if  $A$  is dense but contains no arithmetic progressions of length 3, then  $A$  must be even more dense in a subspace of  $\mathbb{Z}_p^m$  of codimension 1.

**Lemma 3.2** (Density increment). *Let  $f : \mathbb{Z}_p^m \rightarrow \mathbb{R}$  be a function satisfying  $\widehat{f}_0 = 0$  and  $\max \{ |\widehat{f}_\alpha| \mid \alpha \in \mathbb{Z}_p^m \setminus 0 \} = \epsilon > 0$ . Then there exists a affine subspace  $Z'$  of  $\mathbb{Z}_p^m$  of codimension 1 such that  $\mathbf{E}_{x \in Z'} [f(x)] \geq \epsilon/2$ .*

**Problem 3.3.** Prove [Lemma 3.2](#). Hints:

- Show there exists  $\xi \in \mathbb{C}, |\xi| = 1$  and  $\alpha \in \mathbb{Z}_p^m, \alpha \neq 0$  such that  $\langle f, \xi \cdot \chi_\alpha \rangle = \epsilon$ .
- Let  $g = \xi \cdot \chi_\alpha + 1$ . Show, using the assumption  $\widehat{f}_0 = 0$  and [Proposition 2.2](#), that  $\langle f, g \rangle = \epsilon$ .
- Recalling  $\langle \alpha, \beta \rangle_{\mathbb{Z}_p^m} = \sum_{j=1}^m \alpha_j \beta_j$ , let  $\{\alpha\}^\perp = \{ \beta \in \mathbb{Z}_p^m \mid \langle \alpha, \beta \rangle_{\mathbb{Z}_p^m} = 0 \}$  be the space orthogonal to  $\{\alpha\}$ . Notice  $\{\alpha\}^\perp$  has co-dimension 1 and  $\mathbb{Z}_p^m$  can be partitioned into  $p$  affine subspaces that are each a coset of  $\{\alpha\}^\perp$ . Let  $\bigcup_{c \in \mathbb{Z}_p} Z'_c$  denote this partition. Observe that  $g$  is constant on each space  $Z'_c$ .
- To complete the proof, argue that there exist real numbers  $0 \leq \gamma_c \leq 2$  such that

$$\langle f, g \rangle = \mathbf{E}_{x \in \mathbb{Z}_p^m} [f(x) \overline{g(x)}] = \frac{1}{p} \sum_{c \in \mathbb{Z}_p} \gamma_c \cdot \mathbf{E}_{x \in Z'_c} [f(x)].$$

### 3.3 Completing the proof of **Theorem 3.1**

We can now complete the proof of **Theorem 3.1**. The two main ingredients are the observation, taken from **Problem 3.2**, that lack of arithmetic progressions implies a large nonzero Fourier coefficient, and the density increment argument of **Lemma 3.2**.

**Problem 3.4.** Prove **Theorem 3.1** by induction on  $m \geq 3$ . For the inductive step,

- Assume  $A$  has density  $\mu \geq 3/m$  but contains no arithmetic progressions of length 3.
- Let  $f = \mathbf{1}_A - \mu$ .
- Use **Equation 5** from **Problem 3.2** to bound from below the largest magnitude of a nonzero Fourier coefficient of  $f$ .
- Apply **Lemma 3.2** to conclude that there exists a subspace  $Z'$  of  $\mathbb{Z}_p^m$  of codimension 1 in which  $A$  has large density  $\mu' = \frac{|A \cap Z'|}{|Z'|} \geq \frac{3}{n-1}$ .
- Apply the inductive hypothesis regarding  $A \cap Z'$  to conclude the existence of an arithmetic progression of length 3 in  $A$ .

Our next topic comes from the field of *derandomization* which is devoted to the study of randomness in computation and the ways by which randomness, viewed as a computational resource, can be minimized. In particular, we will follow the work initiated in [Bogdanov and Viola \[2007\]](#) and further developed in [Lovett \[2008\]](#); [Viola \[2008\]](#), that uses notions from additive combinatorics to present  $\epsilon$ -fooling sets for low-degree polynomials.

## 4.1 Fooling sets

Let us first define what a fooling set is. Recall our definition of  $\omega_p$  as the primitive  $p^{\text{th}}$  complex root of unity of order  $p$  given by  $\omega_p = e^{2\pi i/p}$ . In what follows let  $\mathbb{F}_q$  denote the finite field of size  $q$ .

**Definition 4.1** (Fooling set). Let  $\mathcal{F}$  be a set of functions from  $\mathbb{F}_p^n$  to  $\mathbb{F}_p$  for prime  $p$ . We say that a multiset  $S \subseteq \mathbb{F}_p^n$  is an  $\epsilon$ -fooling set for  $\mathcal{F}$  if for all  $f \in \mathcal{F}$  we have

$$\left| \mathbf{E}_{x \in \mathbb{F}_p^n} [\omega_p^{f(x)}] - \mathbf{E}_{s \in S} [\omega_p^{f(s)}] \right| \leq \epsilon.$$

Our starting point for the study of  $\epsilon$ -fooling sets is the observation given in the following problem. For simplicity we shall study fooling sets only for  $p = 2$  (in which case  $\omega_2 = (-1)$ ) and point out that all results regarding fooling sets can be generalized to larger prime  $p$  as well.

**Problem 4.1.** Suppose  $\mathcal{F}$  is a set of functions from  $\mathbb{F}_2^n$  to  $\mathbb{F}_2$ . Prove that for any  $\epsilon > 0$  there exists a  $\epsilon$ -fooling set  $S$  of size at most  $|S| \leq O\left(\frac{\log |\mathcal{F}|}{\epsilon^2}\right)$ . Hint: Use the Chernoff bound which says: Let  $X_1, \dots, X_m$  be  $m$  independent identically distributed (iid) random variables each with expectation  $\mu$  and let  $X = \sum_{i=1}^m X_i$ . Then, for all  $0 < \delta < 1$  we have

$$\Pr \left[ \left| \frac{X}{m} - \mu \right| > \delta \right] \leq 2e^{-\delta^2 \mu m / 3}. \quad (6)$$

The previous problem implies that any class of functions that is of size  $2^{n^{O(1)}}$  can be  $\epsilon$ -fooled by a set of size  $n^{O(1)}/\epsilon^2$ . A central question in derandomization is that of finding an *explicit*  $\epsilon$ -fooling set for *interesting* sets of functions. By *explicit* we roughly mean that there exists a deterministic machine that enumerates  $S$  and runs in time polynomial in  $|S|$  and  $n$ . As an example of an *interesting* set of functions take  $\mathcal{F}$  to be the set of functions that can be

computed by circuits of polynomial size in  $n$ . It can be shown that  $|\mathcal{F}| = 2^{n^{O(1)}}$ . Finding a  $\frac{1}{3}$ -fooling set of size  $n^{O(1)}$  for this class is a major open problem and its solution implies  $\mathbf{P} = \mathbf{BPP}$ .

As said above, in the next few lectures we shall study  $\epsilon$ -fooling sets for the interesting set of *low-degree polynomials*. Fix a degree bound  $d$  and let  $\mathcal{F}$  be the set of functions that are evaluations of  $n$ -variate polynomials over  $\mathbb{F}_2$  of total degree at most  $d$ . We start with the simplest case of  $d = 1$  where  $\mathcal{F}$  is the set of *affine linear* functions.

## 4.2 $\epsilon$ -biased sets

A  $\epsilon$ -fooling set for affine linear functions is also known as an  $\epsilon$ -*biased set*. The study of these sets was initiated in [Naor and Naor \[1993\]](#) which defined these sets, presented a construction of them and discussed several important applications of them. By now,  $\epsilon$ -biased sets have found numerous applications in the study of lower bounds, constructions of PCPs, and, of course, derandomization, to name a few (see [Bogdanov and Viola \[2007\]](#); [Ben-Sasson et al. \[2003\]](#) for more details.)

We now present one explicit construction of  $\epsilon$ -biased sets out of three distinct constructions presented in [Alon et al. \[1992\]](#).

*Example 4.2* (An explicit  $\epsilon$ -biased set). Let  $\mathbb{F}_{2^k}$  be the finite field of size  $2^k$ . Fix a basis for  $\mathbb{F}_{2^k}$  over  $\mathbb{F}_2$  and let  $\bar{\alpha} = (\bar{\alpha}_1, \dots, \bar{\alpha}_k) \in \mathbb{F}_2^k$  be the representation of  $\alpha \in \mathbb{F}_{2^k}$  in this basis. Let  $\langle \bar{\alpha}, \bar{\beta} \rangle_{\mathbb{F}_2} = \sum_{i=1}^k \bar{\alpha}_i \bar{\beta}_i$ , where addition and multiplication is carried out in  $\mathbb{F}_2$ . Let

$$S = \left\{ \left( \langle \bar{\alpha}^0, \bar{\beta} \rangle_{\mathbb{F}_2}, \langle \bar{\alpha}^1, \bar{\beta} \rangle_{\mathbb{F}_2}, \dots, \langle \bar{\alpha}^{n-1}, \bar{\beta} \rangle_{\mathbb{F}_2} \right) \mid \alpha, \beta \in \mathbb{F}_{2^k} \right\}.$$

**Problem 4.2.** Prove that the set  $S$  in [Example 4.2](#) is  $\epsilon$ -biased for a suitable choice of  $\epsilon$ . What is the dependence of  $\epsilon$  on  $k$  and  $n$ ? What is the size of  $S$  and how does this compare with the bound shown in [Problem 4.1](#)?

Next are a couple of applications of  $\epsilon$ -biased sets.

**Problem 4.3.** Given  $S \subseteq \mathbb{F}_2^n$  consider the graph  $G$  over vertex set  $\mathbb{F}_2^n$  with edge set

$$E = \{(x, y) \mid \exists s \in S \text{ such that } x = y + s\}.$$

Such a graph is known as a *Cayley graph*. Prove:  $S$  is  $\epsilon$ -biased if and only if the normalized second eigenvalue of the adjacency matrix of  $G$  is at most  $\epsilon$  (i.e.,  $G$  is an expander graph).

**Problem 4.4.** Given  $S \subseteq \mathbb{F}_2^n$  let  $M$  be the  $|S| \times n$  matrix whose rows are the elements of  $S$ . Prove that if  $S$  is  $\epsilon$ -biased then the span of the columns of  $M$ , i.e.,  $C = \{M \cdot x \mid x \in \mathbb{F}_2^n\}$ , forms a linear code with minimal distance  $n(\frac{1}{2} - \epsilon)$ .

We end by stating a well-known open problem.

**Problem 4.5.** Construct an explicit  $\epsilon$ -biased set of size  $O(\frac{n}{\epsilon^2})$ .

In this session we construct explicit  $\epsilon$ -fooling sets for polynomials of degree  $d > 1$ . In particular, we shall prove the following theorem from Viola [2008] which simplifies and improves previous results of Bogdanov and Viola [2007]; Lovett [2008].

**Theorem 5.1** (The sum-set of  $d$   $\epsilon$ -biased sets fools degree  $d$  polynomials). *Let  $S_1, \dots, S_d$  be a sequence of  $\epsilon_1$ -biased sets (as defined in Section 4.2) for  $\mathbb{F}_2^n$ . Then the multiset*

$$S = S_1 + \dots + S_d = \{s_1 + \dots + s_d \mid s_i \in S_i\}$$

*is a  $\epsilon_d$ -fooling set for  $n$ -variate polynomials of degree  $\leq d$  over  $\mathbb{F}_2$ , where*

$$\epsilon_d = \left(16 \cdot \epsilon_1^{1/2^{d-1}}\right).$$

The rest of this session is devoted to the proof of this theorem. The proof goes by induction on  $d$ . Let  $f$  be a degree  $d$  polynomial. We view a random  $s \in S$  as a sum of  $d$  independent random variables  $s_1 + \dots + s_d$ . Roughly speaking, we use the random variable  $s_1$  to reduce the degree of  $f$  to  $d-1$  and then argue that  $s_2 + \dots + s_d$  fools the lower-degree polynomials. To make this strategy work we will employ *directional derivatives*, defined next.

## 5.1 Directional derivatives

The following notion plays a crucial role in the proof of Theorem 5.1 as well as in many other results in additive combinatorics and in applications of additive combinatorics to computer science. Two crucial properties of directional derivatives are obtained in Problem 5.1.

**Definition 5.2** (Directional derivative). Let  $f : \mathbb{F}^n \rightarrow \mathbb{F}$  be a function over a field  $\mathbb{F}$ . The *directional derivative* of  $f$  in direction  $a \in \mathbb{F}^n$  is the function  $f_a$  defined by  $f_a(x) = f(a+x) - f(x)$ .

**Problem 5.1.** Prove: The transformation  $f \mapsto f_a$  is a linear transformation, i.e.,  $(\alpha f + \beta g)_a = \alpha f_a + \beta g_a$  for all functions  $f, g$  and constants  $\alpha, \beta \in \mathbb{F}$ . Furthermore, this transformation strictly reduces the degree of  $f$  as a multivariate polynomial, i.e., if  $\deg(f) > 0$  then  $\deg(f_a) < \deg(f)$ .

## 5.2 Fooling polynomials with large bias

The inductive proof of [Theorem 5.1](#) splits into two cases based on the *bias* of  $f$ ,

$$\text{bias}(f) = \left| \mathbf{E}_{x \in \mathbb{F}_p^n} [\omega_p^{f(x)}] \right| = |\widehat{f}_0|. \quad (7)$$

We start with the simpler case of large bias.

**Lemma 5.3** (Large bias). *Let  $S$  be a  $\epsilon_{d-1}$ -fooling set for degree  $d-1$  polynomials. Then for any degree  $d$  polynomial  $f$ ,*

$$\left| \mathbf{E}_{s \in S} [(-1)^{f(s)}] - \text{bias}(f) \right| \cdot \text{bias}(f) \leq \epsilon_{d-1} \quad (8)$$

**Problem 5.2.** Prove [Lemma 5.3](#). Hints:

- Write [Equation 8](#) as the difference of two terms, where each term is an expectation over *two* independent random variables.
- Fix one of the independent random variables and notice the residual term is a directional derivative.
- Apply induction using [Problem 5.1](#).

## 5.3 Fooling polynomials with small bias

To prove the next lemma we will use directional derivatives and the Cauchy-Schwartz inequality (as described in [Problem 5.3](#)) to reduce the degree of  $f$ .

**Lemma 5.4** (Small bias). *Let  $S$  be a  $\epsilon_{d-1}$ -fooling set for degree  $d-1$  polynomials and let  $T$  be a  $\epsilon_1$ -biased set (i.e., a  $\epsilon_1$ -fooling set for degree 1 polynomials). Then for any degree  $d$  polynomial  $f$ ,*

$$\left| \mathbf{E}_{s \in S, t \in T} [(-1)^{f(s+t)}] \right| \leq \sqrt{(\text{bias}(f))^2 + \epsilon_1^2 + \epsilon_{d-1}} \quad (9)$$

**Problem 5.3.** To prove [Lemma 5.4](#) we need the following basic statement from probability. Let  $X, Y$  be independent random variables and let  $g$  be a function of  $X, Y$  taking on complex values. Prove, using the Cauchy-Schwartz inequality ([Equation 4](#)):

$$\left| \mathbf{E}_{X, Y} [g(X, Y)] \right|^2 \leq \mathbf{E}_X \left[ \mathbf{E}_{Y, Y'} [g(X, Y) \cdot \overline{g(X, Y')}] \right],$$

where  $Y'$  is an independent random variable distributed as  $Y$ .

**Problem 5.4.** Prove [Lemma 5.4](#). Hints:

- Square both sides of [Equation 9](#).
- Expand the (squared) left hand side using [Problem 5.3](#).
- Use [Problem 5.1](#) to replace  $S$  with  $\mathbb{F}_2^n$ , incurring a “small” additive error of  $\epsilon_{d-1}$ , reaching the inequality

$$\left| \mathbf{E}_{s \in S, t \in T} [(-1)^{f(s+t)}] \right|^2 \leq \epsilon_{d-1} + \left| \mathbf{E}_{x \in \mathbb{F}_2^n, t, t' \in T} [(-1)^{f(x+t)+f(x+t')}] \right|^2 \quad (10)$$

- Bound the rightmost term in [Equation 10](#) by  $\epsilon_1^2 + (\text{bias}(f))^2$  by using the Fourier representation of  $f$ .

## 5.4 Completing the proof of [Theorem 5.1](#)

Let us complete the proof of the main theorem of this session.

**Problem 5.5.** Prove [Theorem 5.1](#). Hints: by induction on  $d$ , splitting into two cases based on whether  $\text{bias}(f) > \sqrt{\epsilon_{d-1}}$  or not.

We end the discussion of [Theorem 5.1](#) by stating the following important open problem.

**Problem 5.6.** Design an explicit  $\epsilon$ -fooling set (say, for  $\epsilon \leq 1/4$ ) for the set of  $n$  variate polynomials over  $\mathbb{F}_2$  of degree  $\log n$ .

## 5.5 Approximating biased functions by majority of derivatives

The following lemma from [Bogdanov and Viola \[2007\]](#) is another interesting application of the method of directional derivatives to the study of low-degree polynomials. This lemma has been very useful in the study of error correcting codes based on low-degree polynomials (the so-called Reed-Muller codes) [Kaufman and Lovett \[2008\]](#); [Lovett and Kaufman \[2008\]](#) and in additive combinatorics [Green and Tao \[2007\]](#).

**Lemma 5.5** (Approximation by majority of derivatives). *Let  $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$  be a function with  $\text{bias}(f) = \beta > 0$ . Then for any integer  $k$  there exist  $a_1, \dots, a_k \in \mathbb{F}_2^n$  such that*

$$\Pr_{x \in \mathbb{F}_2^n} [f(x) \neq \text{majority}(f_{a_1}(x), \dots, f_{a_k}(x))] \leq \exp(-\Omega(\beta^2 k))$$

**Problem 5.7.** Prove [Lemma 5.5](#). Hints: Use the Chernoff bound.

**Problem 5.8.** Prove: For all  $\beta, \epsilon > 0$  the following holds. Suppose  $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$  is a degree  $d$  polynomial that agrees with a degree 1 polynomial on a  $\left(\frac{1+\beta}{2}\right)$ -fraction of inputs. Then there exists a function  $g : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$  that is the majority of  $k = O\left(\frac{\log 1/\epsilon}{\beta^2}\right)$  many degree  $(d-1)$ -polynomials, such that  $f$  equals  $g$  on a  $(1-\epsilon)$ -fraction of inputs.

## 6.1 Global and local properties

A *global* property of a set  $A \subseteq \mathbb{F}_2^n$  is a property that depends on *all* elements of  $A$ , i.e., deciding whether  $A$  has this property requires knowing the identity of all elements of  $A$ . A *local* property is one that can be decided by querying a small number of elements of  $A$ . The interplay between local and global properties is a fascinating subject, arising within theoretical computer science in the study of probabilistically checkable proofs (PCPs), hardness of approximation, property testing and coding theory. In the next few sessions we will study the “global vs. local” phenomena in the context of additive combinatorics and see applications of this phenomena to questions in theoretical computer science.

The property of  $A$  being a  $\mathbb{F}_2$ -linear space is a global one. However, this global property implies a multitude of local properties. Namely, if  $A$  is a linear space then *for all*  $a, a' \in A$  we also have  $a + a' \in A$ . The question we shall study next is: Suppose that a *constant fraction* of pairs  $a, a' \in A$  have the local property  $a + a' \in A$ . Can we say that  $A$  is “close” to being a linear space? The following theorem gives a sounding “yes” answer to our question and the rest of this session will be devoted to proving this theorem. Later on we will see one application of this theorem, due to Samorodnitsky [2007], to questions regarding locally testable codes. Our exposition closely follows the presentation in Viola [2007].

**Theorem 6.1.** *Balog and Szemerédi [1994]; Gowers [1998]; Ruzsa [1999]* For every  $\epsilon > 0$  there exists  $\epsilon' > 0$  such that the following holds for all sufficiently large  $n$ . If  $A \subseteq \mathbb{F}_2^n$  satisfies

$$\Pr_{a, a' \in A} [a + a' \in A] \geq \epsilon, \quad (11)$$

then there exists  $A' \subseteq A, |A'| \geq \epsilon'|A|$  such that

$$\frac{|A'|}{|\text{span}(A')|} \geq \epsilon'.$$

The proof of this theorem is combined of two parts. First, we show that Equation 11 implies that some relatively large subset  $A' \subset A$  is roughly closed under pairwise addition, i.e.,  $|A' + A'| = O(|A'|)$ . This part of the proof was shown in Balog and Szemerédi [1994] and later improved in Gowers [1998]. Second, we show that if  $A'$  is roughly closed under pairwise addition, then it is also roughly closed under repeated addition, i.e.,  $|\text{span}(A')| = O(|A'|)$ .

This part of the proof is due to Ruzsa [1999] and improves upon an earlier result of Freiman [1973]. We prove these two parts consecutively.

## 6.2 The Balog-Szemerédi-Gowers Theorem

The following result originally appeared in Balog and Szemerédi [1994] and its quantitative bounds were improved in Gowers [1998].

**Theorem 6.2.** *Suppose  $A \subseteq \mathbb{F}_2^n$  satisfies  $\Pr_{a,a' \in A}[a + a' \in A] \geq \epsilon$ . Then there exists  $A' \subseteq A, |A'| \geq \frac{\epsilon|A|}{3}$  such that  $|A' + A'| \leq \left(\frac{6}{\epsilon}\right)^8 \cdot |A|$ .*

We shall prove Theorem 6.2 by reducing it to the following graph-theoretic question. This method of proof which is due to Sudakov et al. [2005] relies on the following key lemma.

**Lemma 6.3** (Graph density implies multiple length-4 paths). *Let  $G$  be a simple undirected graph over a vertex set  $A$  of size  $m$  and an edge set  $E$  of size  $\epsilon m^2$ . Then there is a set  $A' \subseteq A, |A'| \geq \epsilon m$  such that for all  $a, b \in A'$  there are at least  $(\epsilon/2)^8 \cdot m^3$  paths of length 4 in  $G$  from  $a$  to  $b$ .*

**Problem 6.1.** Prove that Lemma 6.3 implies Theorem 6.2. Hints:

- Construct a graph over vertex set  $A$  and argue the number of edges in it is at least  $\epsilon|A|^2/3$ .
- Apply Lemma 6.3 to conclude every  $c \in A' + A'$  is the sum of many distinct quadruples  $(x, y, z, w) \in A^4$ .

The proof of Lemma 6.3 is a remarkable application of the first moment method, i.e., it uses (only) the linearity of expectation. The magic is due to a careful selection of the elements we are counting.

**Problem 6.2.** Prove Lemma 6.3 by showing there exists  $A' \subseteq A$  such that every  $a \in A'$  shares many ( $\epsilon m$ ) neighbors with most  $((1 - \epsilon')$ -fraction of) nodes in  $A'$ . Hints:

- For  $v \in V$  let  $N(v)$  be the neighborhood of  $v$ , i.e.,  $N(v) = \{u \in V \mid (v, u) \in E\}$ . A pair  $(u, w)$  is said to be a *bad pair* if  $|N(u) \cap N(w)| \leq \epsilon^3 m$ . A vertex  $u \in V$  is *bad for*  $v$  if  $u$  forms a bad pair with at least  $\epsilon^2 m$  other nodes in  $N(v)$ . Let  $S(v)$  denote the set of vertices in  $N(v)$  that are bad for  $v$ .
- Show  $\mathbf{E}_{v \in A}[|S(v)|] \leq \epsilon m$  by giving two bounds on the expected number of bad pairs in  $N(v)$ : (i) a lower bound as a function of  $|S(v)|$ , and (ii) an upper bound using linearity of expectation.
- Conclude  $\mathbf{E}_{v \in A}[|N(v) \setminus S(v)|] \geq \epsilon m$  and fix  $A' = N(v) \setminus S(v)$  for some  $v$  satisfying  $|N(v) \setminus S(v)| \geq \epsilon m$ .
- Prove  $A'$  satisfies the conditions of Lemma 6.3 by showing that every  $a, b \in A'$  form a good pair with almost all  $c \in A'$ .

In this session we complete the proof of [Theorem 6.1](#) by proving the following theorem from [Ruzsa \[1999\]](#) which improves upon [Freiman \[1973\]](#). The starting point of this theorem is that  $A$  does not expand significantly under addition and its conclusion is that  $A$  is roughly a large subset of a linear space. From here on let  $2A$  denote  $A + A$  and inductively define  $kA$  to be  $(k - 1)A + A$  for  $k > 2$ . The notation  $A \times B$  will be reserved for the cartesian product of  $A$  and  $B$ .

**Theorem 7.1.** *For every  $c \geq 1$  there exists  $c' \geq 1$  such that the following holds for all sufficiently large  $n$ . If  $A \subseteq \mathbb{F}_2^n$  satisfies  $|A + A| \leq c \cdot |A|$ , then  $|\text{span}(A)| \leq c' \cdot |A|$ .*

**Problem 7.1.** Prove [Theorem 6.1](#) using [Theorem 6.2](#) and [Theorem 7.1](#).

## 7.1 Bounding the size of $4A$

The proof of [Theorem 7.1](#) has two parts. The “bootstrapping” part, discussed in this section, shows that if  $2A$  is small then  $4A$  is also small. The second, “inductive” part, shows that  $kA$  is small for all  $k \geq 1$ , thus completing the proof. We start the first part by finding a small set of shifts of  $A$  that intersects significantly with  $b + A$  for all  $b \in A$ .

**Problem 7.2.** Prove: If  $|2A| \leq c|A|$  then there exists  $X \subseteq A, |X| \leq 2c$  such that

$$\forall b \in A, |(X + A) \cap (b + A)| \geq |A|/2.$$

Hint: Use a greedy algorithm to construct  $X$ .

**Problem 7.3.** Prove, using the previous problem:

$$\text{If } |2A| \leq c|A| \text{ then } |4A| \leq 16c^5|A|. \tag{12}$$

This is done by showing that every  $w \in 4A$  can be written *in many different ways* as a sum of elements from sets that are smaller than  $4A$ .

- Show that for every  $z \in 2A$  there are at least  $|A|/2$  different triples  $(c, a, x) \in 2A \times A \times X$  such that  $z = c + a + x$ .
- Conclude that for every pair  $(z, z') \in 2A \times 2A$  there exist at least  $|A|^2/4$  *distinct* quintuples  $(c, c', a + a', x, x') \in 2A \times 2A \times 2A \times X \times X$  such that  $z + z' = c + a + x + c' + a' + x'$ .
- Noticing  $z + z' \in 4A$ , conclude the proof.

## 7.2 Bounding the size of $kA$ for arbitrary $k$

We now proceed to the second part of the proof of [Theorem 7.1](#).

**Problem 7.4.** Prove, as in [Problem 7.2](#), that the right hand side of [Equation 12](#) implies there exists  $Y \subseteq 3A$ ,  $|Y| \leq 16c^5$  such that for every  $b \in 3A$  we have  $(Y + A) \cap (b + A) \neq \emptyset$ .

**Problem 7.5.** Complete the proof of [Theorem 7.1](#). Hint: prove by induction on  $k \geq 3$  that  $kA \subseteq (k - 2)Y + 2A$ . Notice that the bound we get on  $|\text{span}(A)|$  is exponential in  $c$ .

We end with an important open problem.

**Problem 7.6.** Is the exponential dependence of  $\epsilon'$  on  $\epsilon$  in [Theorem 6.1](#), which arises from the argument in [Problem 7.5](#), necessary? Can it be replaced by a polynomial dependence? (I.e., can [Theorem 6.1](#) be stated with the added claim that  $\epsilon' = \epsilon^d$  for some absolute constant  $d$  that is independent of  $\epsilon$  and  $n$ ?)

In this session we follow Samorodnitsky [2007] which showed that if a function  $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$  passes a certain *quadracity test* with nonnegligible probability then  $f$  has nonnegligible agreement with a quadratic polynomial. We will focus on showing one of the main observations in Samorodnitsky [2007], namely, the reduction from the above-mentioned problem to a question in additive combinatorics that we have already discussed (in Theorem 6.1). We start by describing the test analyzed in Samorodnitsky [2007].

## 8.1 A “low-degreeness” test

The quadracity test considered by Samorodnitsky is based on the realization that the third iterated directional derivative of a quadratic polynomial is the constant zero function.

**Definition 8.1** (Iterated directional derivative). For integer  $k > 1$ , the  $k^{\text{th}}$  iterated (directional) derivative of a function  $f : \mathbb{F}^n \rightarrow \mathbb{F}$  in directions  $a_1, \dots, a_k \in \mathbb{F}_2^n$  is denoted by  $f_{a_1, \dots, a_k}$  and defined inductively by  $f_{a_1, \dots, a_k} = (f_{a_1, \dots, a_{k-1}})_{a_k}$ , where  $g_a$  is the directional derivative of  $g$  in direction  $a$  as given in Definition 5.2.

In what follows, let  $\mathcal{A}(b; a_1, \dots, a_k)$  denote the affine space that is the additive  $b$ -coset of  $\text{span}(a_1, \dots, a_k)$ ,

$$\mathcal{A}(b; a_1, \dots, a_k) = \left\{ b + \sum_{i \in I} a_i \mid I \subseteq \{1, \dots, k\} \right\}. \quad (13)$$

The *dimension* of  $\mathcal{A}(b; a_1, \dots, a_k)$  is  $\dim(\text{span}(a_1, \dots, a_k))$ .

**Problem 8.1.** Prove: If  $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$  is of degree  $d$  then for any  $b, a_1, \dots, a_{d+1} \in \mathbb{F}_2^n$  we have  $\sum_{\alpha \in \mathcal{A}(b; a_1, \dots, a_{d+1})} f(\alpha) = 0$ .

Consider the following “low-degreeness” test: Pick a random  $d + 1$ -dimensional affine space  $\mathcal{A}$  and sum the value of  $f$  on all points in  $\mathcal{A}$ ; Accept if and only if the sum (over  $\mathbb{F}_2$ ) is 0. The previous problem shows that if  $\deg(f) \geq d$  then the test will pass with probability 1. What about the converse? Can we say anything about  $f$  if we know it passes the test with “good” probability?

This question was first considered for  $d = 1$  by Blum et al. [1990] and an improved, Fourier-based, analysis was given in Bellare et al. [1995]. For  $d > 1$  the question was first discussed in Alon et al. [2005] which showed that if  $f$  passes the test with very high probability (say, .99) then  $f$  is very close (say, .98-close) to a degree  $d$  polynomial. The following theorem from Samorodnitsky [2007], which is the focus of our study, shows that for  $d = 2$ , if  $f$

passes the aforementioned degree-2 test with nonnegligible probability (say, .51), then  $f$  is nonnegligibly close (say, .50001-close) to a quadratic polynomial. We point out that very recently, a result in this spirit was shown for degree  $d > 2$  in [Tao and Ziegler \[2008\]](#), using an ergodic-theory based proof.

**Theorem 8.2** (Quadracity local test). *For every  $\epsilon > 0$  there exists  $\epsilon' > 0$  such that the following holds for all sufficiently large  $n$ . If  $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$  satisfies*

$$\mathbf{E}_{b, a_1, a_2, a_3} \left[ \prod_{z \in \mathcal{A}(b, a_1, a_2, a_3)} (-1)^{f(z)} \right] \geq \epsilon, \quad (14)$$

*then there exists a quadratic  $n$ -variate polynomial  $g$  over  $\mathbb{F}_2$  such that*

$$\Pr_{x \in \mathbb{F}_2^n} [f(x) = g(x)] \geq \frac{1 + \epsilon'}{2}. \quad (15)$$

## 8.2 The Fourier spectrum of derivatives of quadratics

The starting point of the proof of [Theorem 8.2](#) is that when  $f$  is a quadratic function, the Fourier coefficients of the derivatives of  $(-1)^f$  are “nicely structured”. As customary when using Fourier analysis of Boolean functions, from here on we will assume that the range of  $f$  is  $\{-1, 1\}$  by considering  $f'(z) = (-1)^{f(z)}$ . We also abuse notation and denote the  $b$ -Fourier coefficient of the function  $f'$  by  $\widehat{f}'(b)$ . This will allow us to consider such objects as the  $b$ -Fourier coefficient of  $f'_a$ , which will be denoted by  $\widehat{f}'_a(b)$ .

**Problem 8.2.** If  $f$  is a quadratic polynomial then there exists a symmetric matrix  $B \in \mathbb{F}_2^{n \times n}$  with zero diagonal (i.e.,  $B_{ii} = 0$  for  $i = 1, \dots, n$ ) such that for every  $a \in \mathbb{F}_2^n$ , the  $|\widehat{f}'_a(B \cdot a)| = 1$ . (Notice this implies, by Parseval’s equality ([Equation 3](#)) that all other Fourier coefficients of  $f'_a$  are 0). Hint: Express  $f(x)$  as  $x^\top Ax + \alpha$  for  $A \in \mathbb{F}_2^{n \times n}$  and  $\alpha \in \mathbb{F}_2$ .

Thus, if  $f$  is a quadratic polynomial then the *choice function* that maps  $a$  to the largest Fourier coefficient of  $f'_a$ , is a *linear* function. Next, we show that to prove [Theorem 8.2](#) it is sufficient to find a *linear* choice function that has nonnegligible success probability in choosing large Fourier coefficients of derivatives of  $f$ .

**Problem 8.3.** Prove: If  $B \in \mathbb{F}_2^{n \times n}$  is symmetric with zero diagonal and satisfies

$$\mathbf{E}_a \left[ \widehat{f}'_a(Ba)^2 \right] \geq \epsilon > 0 \quad (16)$$

then there exists a quadratic polynomial  $g$  such that [Equation 15](#) holds for  $f, g$  and some  $\epsilon' > 0$  that depends only on  $\epsilon$ . Hints:

- Let  $h(x) = x^\top Ax$  for  $A$  satisfying  $A + A^\top = B$ .
- Use Plancherel’s equality ([Equation 2](#)) and [Problem 8.2](#) to show  $\mathbf{E}_a [\langle f'_a, h'_a \rangle^2] \geq \epsilon$ .

- Show, using the Fourier representation of  $f'$  and  $h'$  and the orthonormality of characters ([Proposition 2.2](#)), that  $\mathbf{E}_a[\langle f'_a, h'_a \rangle^2] = \sum_{c \in \mathbb{F}_2^n} (\widehat{f+h})'(c)^4$ , where  $(f+h)'(x) = (-1)^{f(x)+h(x)}$ .
- Use Parseval's equality ([Equation 3](#)) to conclude there exists  $c \in \mathbb{F}_2^n, c' \in \mathbb{F}_2$  such that  $g(x) = h(x) + \langle c, x \rangle_{\mathbb{F}_2} + c'$  is the quadratic satisfying [Equation 15](#).

### 8.3 From Fourier analysis to additive combinatorics

We now arrive at the crux of the argument in the proof of [Theorem 8.2](#). Our starting point is the following lemma from [Samorodnitsky \[2007\]](#). We omit its proof, which is done by the standard method of Fourier-analysis (which is similar to the way we solved [Problem 8.3](#)

**Lemma 8.3.** *If [Equation 14](#) holds then there exists  $\tilde{\epsilon} > 0$  that depends only on  $\epsilon$  such that*

$$\mathbf{E}_{a,b} \left[ \sum_{\alpha, \beta} \widehat{f'_a}^2(\alpha) \widehat{f'_b}^2(\beta) \widehat{f'_{a+b}}^2(\alpha + \beta) \right] \geq \tilde{\epsilon}. \quad (17)$$

In the following problem we show that [Lemma 8.3](#) implies the existence of a choice function that displays “linear-like” behavior.

**Problem 8.4.** Prove, using [Lemma 8.3](#), that there exists a “choice function”  $\psi : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$  satisfying

$$\Pr_{x,y} \left[ \psi(x) + \psi(y) = \psi(x+y) \text{ and } \min \left\{ \widehat{f'_x}^2(\psi(x)), \widehat{f'_y}^2(\psi(y)), \widehat{f'_{x+y}}^2(\psi(x+y)) \right\} \geq \tilde{\epsilon}/6 \right] \geq \tilde{\epsilon}/2. \quad (18)$$

Hints:

- Define a distribution on functions  $\phi : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$  by selecting  $\phi(x) = \alpha$  with probability  $\widehat{f'_x}(\alpha)$ , independently for each  $x \in \mathbb{F}_2^n$ .
- Let  $L(\phi)$  be the random variable defined by

$$L(\phi) = \Pr_{x,y} \left[ \phi(x) + \phi(y) = \phi(x+y) \text{ and } \min \left\{ \widehat{f'_x}^2(\phi(x)), \widehat{f'_y}^2(\phi(y)), \widehat{f'_{x+y}}^2(\phi(x+y)) \right\} \geq \tilde{\epsilon}/6 \right]$$

- Show that  $\mathbf{E}_\phi[L(\phi)] \geq \tilde{\epsilon}/2$ .

Next, we use additive combinatorics to argue that  $\phi$  is really a linear transformation defined on a affine subspace of  $\mathbb{F}_2^n$  of constant co-dimension.

**Problem 8.5.** Prove: if [Equation 14](#) holds then there exists an affine subspace  $\mathcal{A}$  of co-dimension  $c$ , where  $c$  depends only on  $\epsilon$  (and is independent of  $n$ ), and a affine transformation  $D : \mathcal{A} \rightarrow \mathbb{F}_2^n$  such that  $\mathbf{E}_{x \in \mathcal{A}} \left[ \widehat{f'_x}^2(D(x)) \right] \geq \epsilon'$  for some  $\epsilon' > 0$  that depends only on  $\epsilon$ . Hints: Use [Problem 8.3](#) and [Theorem 6.1](#).

To sum up, we have showed that [Equation 14](#) implies the existence of a affine choice function  $D$  that works well on a affine subspace of constant co-dimension. To complete the proof of [Theorem 8.2](#), via [Problem 8.2](#), we need to extend  $D$  to all of  $\mathbb{F}_2^n$  and furthermore show that it is a a linear transformation described by a symmetric function with zero diagonal. Details of this part can be found in [Samorodnitsky \[2007\]](#).

In the last part of this course we will see how *sum-product* theorems over finite fields can be used to construct so-called *extractors for independent sources*. In particular, we will follow [Barak et al. \[2006\]](#) who were the first to notice the application of sum-product theorems to constructing pseudorandom objects. As done earlier in the course, we start by describing the mathematical tools to be used and later on will elaborate on their applications to theoretical computer science.

## 9.1 Sum-Product Theorems

A fundamental theorem in additive combinatorics, due to Erdos and Szemerédi [ES](#), says that no set of integers can be *simultaneously* closed under addition and under multiplication. Formally, there exists  $\epsilon > 0$  such that for any  $A \subset \mathbb{N}$  we have  $\max\{|A + A|, |A \cdot A|\} \geq |A|^{1+\epsilon}$ , i.e., at least one of the sum-set or product-set of  $A$  must be significantly larger than  $A$ . (Yet another way of thinking of this is that sets that are “arithmetic-progression-like” in the sense that  $|A + A| \approx O(|A|)$  cannot be “geometric-progression-like”, i.e., they cannot have  $|A \cdot A| \approx O(|A|)$  as well.)

Recently, similar sum-product theorems were obtained for subsets of finite fields that have no large subfields (examples of such fields are  $\mathbb{F}_p$  and  $\mathbb{F}_{2^p}$  where  $p$  is a prime). The first such result is the following theorem of [Bourgain et al. \[2003\]](#).

**Theorem 9.1.** *For any  $\delta > 0$  there exists  $\epsilon > 0$  such that the following holds for all sufficiently large finite field that does not contain a subfield of size  $\geq |\mathbb{F}|^\delta$ . For every*

$$A \subset \mathbb{F}, \quad |\mathbb{F}|^\delta \leq |A| \leq |\mathbb{F}|^{1-\delta} \tag{19}$$

*we have  $\max\{|A + A|, |A \cdot A|\} \geq |A|^{1+\epsilon}$ .*

*Remark.* Notice that the requirement placed on  $\mathbb{F}$  is essential. If  $\mathbb{F}$  contains a subfield of size  $|\mathbb{F}|^\delta$  then taking  $A$  to be this subfield one sees that  $|A + A|, |A \cdot A| = |A|$ . Similarly, the size of  $A$  must be bounded from above because taking  $A = \mathbb{F}$  gives a set with no expansion under addition or multiplication.

The proof of the above theorem as well as followups and improvements to it [Konyagin \[2002\]](#); [Bourgain et al. \[2006\]](#) follows in two steps. To explain these steps, we need the following definition.

**Definition 9.2** (Rational expression). A *rational expression* is an expression involving the four basic arithmetic operations (addition, subtraction, multiplication and division) and distinct variables denoted  $x_1, x_2, \dots$ . Given an expression  $R(x_1, \dots, x_k)$  and a set  $A \subseteq \mathbb{F}$  let  $R(A) = \{R(a_1, \dots, a_k) \mid a_i \in A\}$ .

For instance, when  $R(x_1, x_2) := x_1 + x_2$  then  $R(A) = A + A$  and when  $R(x_1, x_2, x_3) := x_1 \cdot x_2 + x_3$  then  $R(A) = (A \cdot A) + A$ . Using [Definition 9.2](#) we describe the proof strategy of [Theorem 9.1](#) as follows.

1. Pick a specific rational expression  $R$  and show that there exists  $\epsilon > 0$  such that for all  $A$  satisfying [Equation 19](#) we have  $|R(A)| \geq |A|^{1+\epsilon}$ .
2. Show that if one assumes  $|A+A|, |A \cdot A| < |A|^{1+\epsilon'}$  then there exists  $A' \subset A, |A'| > |A|^\gamma$  (for  $\gamma > 0$ ) such that  $|R(A')| < |A'|^{1+\epsilon}$ , contradicting part 1.

Next we elaborate on these two parts.

## 9.2 An explicit expanding rational expression

As an example of the first part of the proof of [Theorem 9.1](#) we use a rational expression suggested and analyzed in [Barak et al. \[2006\]](#). In particular, they proved:

**Lemma 9.3.** *There exists a rational expression over 16 variables such that for every  $\delta > 0$  and every field  $\mathbb{F}$  such that  $\mathbb{F}$  does not contain a subfield of size  $\geq |\mathbb{F}|^\delta$  the following holds. If  $|A| \geq |\mathbb{F}|^\delta$  then  $R(A) \geq \min\{|A|^{1+\delta}, |\mathbb{F}|\}$ .*

To prove this lemma [Barak et al. \[2006\]](#) showed that the following rational expression expands.

**Lemma 9.4.** *Let  $A \subseteq \mathbb{F}$  be such that  $|\mathbb{F}|^{\frac{1}{k}} < |A| \leq |\mathbb{F}|^{\frac{1}{k-1}}$  for integer  $k \geq 2$ . Then*

$$\left| \frac{A - A}{A - A} \right| \geq |\mathbb{F}|^{\frac{1}{k-1}}.$$

**Problem 9.1.** Prove [Lemma 9.4](#). Hints:

- Argue by way of contradiction. Let  $B = \frac{A-A}{A-A}$  and suppose  $|B| < |\mathbb{F}|^{\frac{1}{k-1}}$ .
- Show the existence of a sequence  $s_1, \dots, s_{k-1} \in \mathbb{F}$  such that for each  $i = 1, \dots, k-1$  we have  $s_i \notin B + \sum_{j < i} s_j B$ .
- Consider the function  $f : A^k \rightarrow \mathbb{F}$  given by  $f(x_0, \dots, x_{k-1}) = x_0 + \sum_{i < k} s_i x_i$ .
- Argue the existence of distinct  $x, y \in A^k$  such that  $f(x) = f(y)$ .
- Conclude, using such  $x, y$ , that the sequence  $s_1, \dots, s_{k-1}$  as defined above cannot exist.

**Problem 9.2.** Prove [Lemma 9.3](#). Hints: use [Lemma 9.4](#).

- Let  $R'$  be the rational expression stated in [Lemma 9.4](#).
- Let  $R$  be the composition of  $R'$  with itself.
- Choose  $k$  such that  $|\mathbb{F}|^{1/k} < |A| \leq |\mathbb{F}|^{1/(k-1)}$ .
- Notice that  $|\mathbb{F}|^{1/k-1}$  is not an integer (because  $\mathbb{F}$  has no large subfields) so the inequality in the conclusion of [Lemma 9.4](#) is strict.

In our next session we continue with the proof of [Theorem 9.1](#) by attending to its second part as described at the end of [Section 9.1](#). Recall that we need to show that if both the size  $|A+A|$  and of  $|A \cdot A|$  are small relative to the size of  $A$  then the size of  $R(A)$  is also small, where  $R$  is a fixed rational expression. We have already seen in [Problem 7.3](#) that when  $|A+A| = O(|A|)$  then  $|\ell A| = O(|A|)$  as well. The very same result (stated for a multiplicative group) implies that  $|A^\ell|$  is small when  $|A \cdot A|$  is small. However, such results discuss only a single arithmetic operation, whereas we need to bound the size of sets involving all four of them and this is what we do next.

## 10.1 Bounding expressions that involve an operator and its inverse

The following theorem bounds the size of rational expressions involving a single group operation and its inverse. We state the theorem for the pair of addition–subtraction but the same proof applies to multiplication–division as well. The initial proof of this theorem is from [Plünnecke \[1969\]](#) and we give the simpler proof from [Ruzsa \[1996\]](#).

**Theorem 10.1** (Plünnecke-Ruzsa iterated sums theorem). *There exist constants  $c_1, c_2$  such that the following holds for any finite subsets  $A, B, |A| = |B|$  of an abelian group  $G$  (written additively). If  $|A+B| \leq K|A|$  then  $|\ell A - \ell A + \ell B - \ell B| \leq c_1 K^{c_2} |A|$ .*

To prove the theorem we start with the following lemma from [Ruzsa \[1996\]](#).

**Lemma 10.2** (Ruzsa’s inequality). *For any finite subsets  $C, D, E$  of a group  $G$  (written additively) we have*

$$|C - D| \cdot |E| \leq |C - E| \cdot |E - D|. \quad (20)$$

**Problem 10.1.** Prove [Lemma 10.2](#). Hints:

- Define a function  $\phi : (C - E) \times D \rightarrow (C - E) \times (D - E)$  by fixing for each  $\lambda \in C - E$  a pair  $c_\lambda \in C, d_\lambda \in D$  such that  $c_\lambda - d_\lambda = \lambda$  and letting  $\phi(\lambda, e) = (c_\lambda - e, e - d_\lambda)$ .
- Prove  $\phi$  is injective.

To prove [Theorem 10.1](#) we also need the result stated in the next problem.

**Problem 10.2.** Prove: If  $|A| = |B|$  and  $|A + B| \leq K|A|$  then there exists  $S \subseteq A + B$ ,  $|S| \geq |A|/2$  such that  $|A + B + S| \leq 2K^3|A|$ . Hints:

- Let  $S \subseteq A + B$  be the set of elements that have at least  $|A|/2K$  representations as sums of pairs from  $A, B$ . Argue  $|S| \geq |A|/2$ .
- To prove  $|A + B + S| \leq 2K^3|A|$  notice that every element  $x \in A + B + S$  can be written in  $|S|/2K$  distinct ways as a sum of two elements from  $A + B$ .

**Problem 10.3.** Prove, using [Lemma 10.2](#) and [Problem 10.2](#):

1. If  $|A + A| \leq |A|^{1+\epsilon}$  then  $|A - A| \leq |A|^{1+2\epsilon}$ .
2. If  $|A| = |B|$  and  $|A + B| \leq K|A|$  then  $|A - A + B - B| \leq 8K^6|A|$ .
3. [Theorem 10.1](#) for the case of  $\ell$  being a power of 2.

## 10.2 Bounding expressions that involve addition and multiplication

Finally, we discuss how to bound expressions involving addition and multiplication. Following [[Barak et al., 2007](#), Section 4.4], we next state a “dream version” of [Theorem 6.2](#), i.e., a version that is not known to be true, and use it to bound a certain expression involving subtraction and multiplication. Similar techniques are used to bound the size of an arbitrary rational expression (see, e.g., [[Barak et al., 2007](#), Section 4.4] for details).

**Lemma 10.3** (“Dream version” of the Balog-Szemerédi-Gowers theorem). *If  $|A - A| \leq K|A|$  then  $\exists A' \subset A$ ,  $|A'| \geq |A|/K^c$  such that every  $b \in A' - A'$  has  $|A|/K^c$  representations as  $b = x - y$ , where  $x, y \in A'$ .*

**Problem 10.4.** Prove that [Lemma 10.3](#) implies that if  $|A + A|, |A \cdot A| \leq K|A|$  then there exists  $B \subset A$ ,  $|B| \geq |A|/K^c$  such that  $|B \cdot B - B \cdot B| \leq |A|/K^c$ . Hints:

- Bound the size of  $(B - B) \cdot B$  by arguing every element in this set has many representations as a difference of two elements from  $B \cdot B$ .
- Notice that  $b - b' \in B \cdot B - B \cdot B$  can be written in  $\approx |B|$  ways as a sum of two elements from  $(B - B) \cdot B$ .

## References

1. On sums and products of integers. In *Studies in Pure Mathematics. To the memory of Paul Turán*.
2. Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple construction of almost  $k$ -wise independent random variables. *Random Struct. Algorithms*, 3(3):289–304, 1992.
3. Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing reed-muller codes. *IEEE Transactions on Information Theory*, 51(11):4032–4039, 2005. URL <http://doi.ieeecomputersociety.org/10.1109/TIT.2005.856958>.
4. Antal Balog and Endre Szemerédi. A statistical theorem of set addition. *Combinatorica*, 14(3):263–268, 1994.
5. Boaz Barak, Russell Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. *SIAM J. Comput.*, 36(4):1095–1118, 2006. URL <http://dx.doi.org/10.1137/S0097539705447141>.
6. Boaz Barak, Luca Trevisan, and Avi Wigderson. A mini course on additive combinatorics. Online, October 2007. URL <http://www.cs.princeton.edu/theory/index.php/Main/AdditiveCombinatoricsMinicourse>.
7. Mihir Bellare, Don Coppersmith, Johan Håstad, Marcos A. Kiwi, and Madhu Sudan. Linearity testing in characteristic two. In *FOCS '95: Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, pages 432–441, Washington, DC, USA, 1995. IEEE Computer Society. ISBN 0-8186-7183-1.
8. Eli Ben-Sasson. Research laboratory in foundations of computer science (236602). Course notes, Fall 2006. URL [http://www.cs.technion.ac.il/~eli/courses/2006\\_Fall/](http://www.cs.technion.ac.il/~eli/courses/2006_Fall/).
9. Eli Ben-Sasson, Madhu Sudan, Salil P. Vadhan, and Avi Wigderson. Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. In *STOC*, pages 612–621, 2003.
10. Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. In *STOC '90: Proceedings of the twenty-second annual ACM symposium on Theory of computing*, pages 73–83, New York, NY, USA, 1990. ACM Press. ISBN 0-89791-361-2.
11. Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. In *FOCS*, pages 41–51. IEEE Computer Society, 2007. URL <http://doi.ieeecomputersociety.org/10.1109/FOCS.2007.56>.
12. J. Bourgain, A. A. Glibichuk, and S. V. Konyagin. Estimates for the number of sums and products and for exponential sums in fields of prime order. *J. London Math. Soc.*, 73(2):380–398, 2006.
13. Jean Bourgain. On triples in arithmetic progressions. *GAF*, 9:968–984, 1999.
14. Jean Bourgain, Nets Katz, and Terence Tao. A sum-product estimate in finite fields, and applications, March 10 2003. URL <http://arxiv.org/abs/math/0301343>. Comment: 29 pages. The distance set result needs to be restricted to the case when  $-1$  is not a square.
15. G. A. Freiman. *Foundations of a structural theory of set addition*, volume 37. American Mathematical Society, 1973.

16. W. T. Gowers. A new proof of szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998.
17. Ben Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the gowers norms, November 20 2007. URL <http://arxiv.org/abs/0711.3191>. Comment: 33 pages.
18. Elena Grigorescu, Swastik Kopparty, and Madhu Sudan. Local decoding and testing for homomorphisms. In Josep Díaz, Klaus Jansen, José D. P. Rolim, and Uri Zwick, editors, *APPROX-RANDOM*, volume 4110 of *Lecture Notes in Computer Science*, pages 375–385. Springer, 2006. ISBN 3-540-38044-2. URL [http://dx.doi.org/10.1007/11830924\\_35](http://dx.doi.org/10.1007/11830924_35).
19. Johan Hästad and Avi Wigderson. Simple analysis of graph tests for linearity and PCP. *Random Struct. Algorithms*, 22(2):139–160, 2003. URL <http://dx.doi.org/10.1002/rsa.10068>.
20. Tali Kaufman and Shachar Lovett. Worst case to average case reductions for polynomials. In *FOCS*, pages 166–175. IEEE Computer Society, 2008. URL <http://dx.doi.org/10.1109/FOCS.2008.17>.
21. S. V. Konyagin. Estimates for trigonometric sums and for gaussian sums. In *IV International conference on 'Modern problems of number theory and its applications' Part 3*, pages 86–114. Moscow State University, 2002.
22. Shachar Lovett. Unconditional pseudorandom generators for low degree polynomials. In Richard E. Ladner and Cynthia Dwork, editors, *STOC*, pages 557–562. ACM, 2008. ISBN 978-1-60558-047-0. URL <http://doi.acm.org/10.1145/1374376.1374455>.
23. Shachar Lovett and Tali Kaufman. The list-decoding size of reed-muller codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 15(111), 2008. URL <http://eccc.hpi-web.de/eccc-reports/2008/TR08-111/index.html>.
24. Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *J. Comb. Theory, Ser. A*, 71(1):168–172, 1995.
25. Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM Journal on Computing*, 22(4):838–856, August 1993. ISSN 0097-5397 (print), 1095-7111 (electronic).
26. H. Plünnecke. Eigenschaften und abschätzungen von wirkungsfunktionen. *Berichte der Gesellschaft für Mathematik und Datenverarbeitung*, 22, 1969.
27. Imre Ruzsa. Sums of finite sets. In M. Nathanson D. Chudnovsky, G. Chudnovsky, editor, *Number Theory: New York Seminar*. Springer-Verlag, 1996.
28. Imre Z. Ruzsa. An analog of freiman's theorem in groups. *Astérisque*, 258:323–326, 1999.
29. Alex Samorodnitsky. Low-degree tests at large distances. In David S. Johnson and Uriel Feige, editors, *STOC*, pages 506–515. ACM, 2007. ISBN 978-1-59593-631-8. URL <http://doi.acm.org/10.1145/1250790.1250864>.
30. B. Sudakov, E. Szemerédi, and V. H. Vu. On a question of erdős and moser, July 15 2005. URL <http://ProjectEuclid.org/getRecord?id=euclid.dmj/1121448866>.

31. Terence Tao and Van Vu. *Additive Combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2006.
32. Terence Tao and Tamar Ziegler. The inverse conjecture for the gowers norm over finite fields via the correspondence principle. *Arxiv*, 2008. URL <http://arxiv.org/abs/0810.5527v1>.
33. Emanuele Viola. The sum of  $d$  small-bias generators fools polynomials of degree  $d$ . In *IEEE Conference on Computational Complexity*, pages 124–127. IEEE Computer Society, 2008. ISBN 978-0-7695-3169-4. URL <http://doi.ieeecomputersociety.org/10.1109/CCC.2008.16>.
34. Emanuele Viola. Selected results in additive combinatorics: An exposition. *Electronic Colloquium on Computational Complexity (ECCC)*, 14(103), 2007. URL <http://eccc.hpi-web.de/eccc-reports/2007/TR07-103/index.html>.