Background on Higher-Order Fourier Analysis

FOCS '14 Workshop

Arnab Bhattacharyya
Indian Institute of Science

October 18, 2014

Plan for the day

- **Talk 1**: Mathematical primer (me)
- **Talk 2**: Polynomial pseudorandomness (P. Hatami)
- **Talk 3**: Algorithmic h.o. Fourier analysis (Tulsiani)
- **Talk 4**: Applications to property testing (Yoshida)
- **Talk 5**: Applications to coding theory (Bhowmick)
- **Talk 6**: A different generalization of Fourier analysis and application to communication complexity (Viola)

Teaser

Given a quartic polynomial $P\colon \mathbb{F}_2^n \to \mathbb{F}_2$, can we decide in poly($n$) time whether:

$$P = Q_1 Q_2 + Q_3 Q_4$$

where Q_1, Q_2, Q_3, Q_4 are quadratic polys?

Yes! [B. '14, B.-Hatami-Tulsiani '15]

Some Preliminaries

Setting

$\mathbb{F}$ = finite field of **fixed prime** order

- For example, $\mathbb{F} = \mathbb{F}_2$ or $\mathbb{F} = \mathbb{F}_{97}$
- Theory simpler for fields of large (but fixed) size

Functions

Functions are always multivariate, on n variables

$$f\colon \mathbb{F}^n \to \mathbb{C} \qquad (|f| \leq 1)$$

and $$P\colon \mathbb{F}^n \to \mathbb{F}$$

Current bounds aim to be efficient wrt *n*

Polynomial

Polynomial of degree $\boldsymbol{d}$ is of the form:

$$\sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n}$$

where each $c_{i_1,\ldots,i_n} \in \mathbb{F}$ and $i_1 + \cdots + i_n \leq d$

Phase Polynomial

Phase polynomial of degree $\boldsymbol{d}$ is a function $f: \mathbb{F}^n \to \mathbb{C}$ of the form $f(x) = \mathrm{e}(P(x))$ where:

1. $P: \mathbb{F}^n \to \mathbb{F}$ is a polynomial of degree d
2. $\mathrm{e}(k) = e^{2\pi i k/|\mathbb{F}|}$

Inner Product

The **inner product** of two functions $f, g\colon \mathbb{F}^n \to \mathbb{C}$ is:

$$\langle f, g \rangle = \mathbb{E}_{x \in \mathbb{F}^n}\left[f(x) \cdot \overline{g(x)}\right]$$

Magnitude captures correlation between f and g

Derivatives

Additive derivative in direction

$h \in \mathbb{F}^n$ of function $P\colon \mathbb{F}^n \to \mathbb{F}$ is:

$$D_h P(x) = P(x+h) - P(x)$$

Derivatives

Multiplicative derivative in direction $h \in \mathbb{F}^n$ of function $f\colon \mathbb{F}^n \to \mathbb{C}$ is:

$$\Delta_h f(x) = f(x+h) \cdot \overline{f(x)}$$

Polynomial Factor

Factor of degree d and order m is a tuple of polynomials $\mathcal{B} = (P_1, P_2, \dots, P_m)$, each of degree d.

As shorthand, write:

$$\mathcal{B}(x) = (P_1(x), \dots, P_m(x))$$

Fourier Analysis over $\mathbb{F}$

Fourier Representation

Every function $f\colon\ \mathbb{F}^n \to \mathbb{C}$ is a linear combination of **linear** phases:

$$f(x) = \sum_{\alpha \in \mathbb{F}^n} \hat{f}(\alpha)\, \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

Linear Phases

- The inner product of two linear phases is:

$$\left\langle \mathrm{e}\left(\sum_i \alpha_i x_i\right), e\left(\sum_i \beta_i x_i\right)\right\rangle = \mathbb{E}_x\left[\mathrm{e}\left(\sum_i (\alpha_i - \beta_i)x_i\right)\right] = 0$$

if $\alpha \neq \beta$ and is 1 otherwise.

- So:

$\hat{f}(\alpha) = \langle f, \mathrm{e}(\sum_i \alpha_i x_i)\rangle$= correlation with linear phase

Random functions

With high probability, a **random** function $f\colon \mathbb{F}^n \to \mathbb{C}$ with $|f| = 1$ has each $\hat{f}(\alpha) \to 0$.

Decomposition Theorem

$$f(x) = g(x) + h(x)$$

where:

$$g(x) = \sum_{\alpha:\, \hat{f}(\alpha) \geq \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

$$h(x) = \sum_{\alpha:\, \hat{f}(\alpha) < \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

Decomposition Theorem

$$g(x) = \sum_{\alpha:\, \hat{f}(\alpha) \geq \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

$$h(x) = \sum_{\alpha:\, \hat{f}(\alpha) < \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

Every Fourier coefficient of h is less than ϵ, so h is "pseudorandom".

Decomposition Theorem

$$g(x) = \sum_{\alpha:\, \hat{f}(\alpha) \geq \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

$$h(x) = \sum_{\alpha:\, \hat{f}(\alpha) < \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

g has only $1/\epsilon^2$ nonzero Fourier coefficients

Decomposition Theorem

$$g(x) = \sum_{\alpha:\, \hat{f}(\alpha) \geq \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

$$h(x) = \sum_{\alpha:\, \hat{f}(\alpha) < \epsilon} \hat{f}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i x_i\right)$$

The nonzero Fourier coefficients of g can be found in poly time **[Goldreich-Levin ‘89]**

Elements of Higher-Order Fourier Analysis

Higher-order Fourier analysis is the interplay between three different notions of pseudorandomness for functions and factors.

1. Bias
2. Gowers norm
3. Rank

Bias

Bias

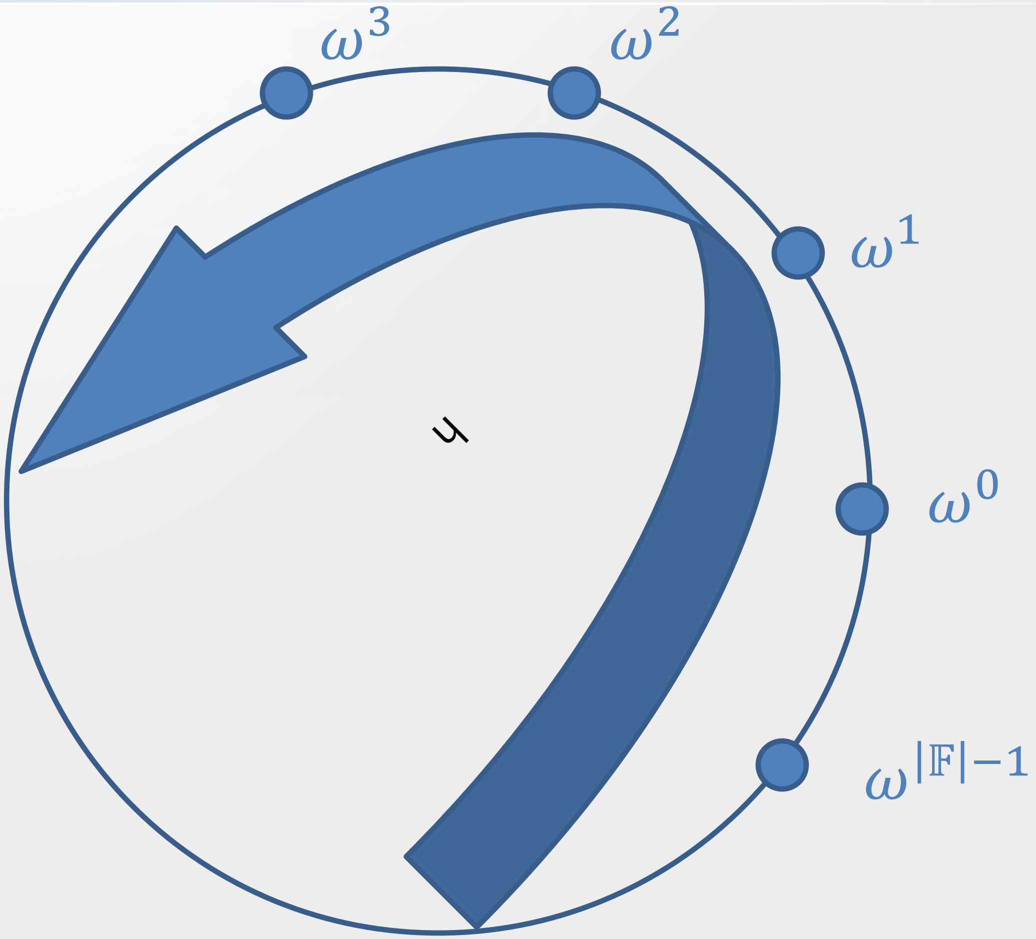
For $f\colon \mathbb{F}^n \to \mathbb{C}$,

$$\text{bias}(f) = |\mathbb{E}_x[f(x)]|$$

For $P\colon \mathbb{F}^n \to \mathbb{F}$,

$$\text{bias}(P) = \left|\mathbb{E}_x\left[\mathrm{e}\left(P(x)\right)\right]\right|$$

[…, Naor-Naor ‘89, …]



How well is P equidistributed?

Bias of Factor

A factor $\mathcal{B} = (P_1, \dots, P_k)$ is ***α*-unbiased** if every nonzero linear combination of $P_1, \dots, P_k$ has bias less than α:

$$\text{bias}\left(\sum_{i=1}^{k} c_i P_i\right) < \alpha \quad \forall (c_1, \dots, c_k) \in \mathbb{F}^k \setminus \{0\}$$

Bias implies equidistribution

Lemma: If $\mathcal{B}$ is α-unbiased and of order k, then for any $c \in \mathbb{F}^k$:

$$\Pr[\mathcal{B}(x) = c] = \frac{1}{|\mathbb{F}|^k} \pm \alpha$$

Bias implies equidistribution

Lemma: If $\mathcal{B}$ is α-unbiased and of order k, then for any $c \in \mathbb{F}^k$:

$$\Pr[\mathcal{B}(x) = c] = \frac{1}{|\mathbb{F}|^k} \pm \alpha$$

Corollary: If $\mathcal{B}$ is α-unbiased and $\alpha < \frac{1}{|\mathbb{F}|^k}$, then $\mathcal{B}$ maps onto $\mathbb{F}^k$.

Gowers Norm

Gowers Norm

Given $f\colon\ \mathbb{F}^n \to \mathbb{C}$, its **Gowers norm of order d** is:

$$U^d(f) = \left|\mathbb{E}_{x,h_1,h_2,\ldots,h_d}\Delta_{h_1}\Delta_{h_2}\cdots\Delta_{h_d}f(x)\right|^{1/2^d}$$

[Gowers '01]

Gowers Norm

Given $f\colon\ \mathbb{F}^n \to \mathbb{C}$, its **Gowers norm of order d** is:

$$U^d(f) = |\mathbb{E}_{x,h_1,h_2,\ldots,h_d} \Delta_{h_1}\Delta_{h_2}\cdots\Delta_{h_d} f(x)|^{1/2^d}$$

Observation: If $f = \mathrm{e}(P)$ is a phase poly, then:

$$U^d(f) = |\mathbb{E}_{x,h_1,h_2,\ldots,h_d} \mathrm{e}\big(D_{h_1}D_{h_2}\cdots D_{h_d}P(x)\big)|^{1/2^d}$$

Gowers norm for phase polys

- If f is a phase poly of degree d, then:

$$U^{d+1}(f) = 1$$

- Converse is true when $d < |\mathbb{F}|$.

Other Observations

- $U^1(f) = \sqrt{|\mathbb{E}[f]|^2} = \text{bias}(f)$

- $U^2(f) = \sqrt[4]{\sum_\alpha \hat{f}^4(\alpha)}$

- $U^1(f) \leq U^2(f) \leq U^3(f) \leq \cdots$ (C.-S.)

Pseudorandomness

- For random $f\colon\ \mathbb{F}^n \to \mathbb{C}$ and fixed d,
$$U^d(f) \to 0$$

- By monotonicity, low Gowers norm implies low bias and low Fourier coefficients.

Correlation with Polynomials

Lemma: $U^{d+1}(f) \geq \max |\langle f, \mathrm{e}(P) \rangle|$ where max is over all polynomials P of degree d.

Proof: For any poly P of degree d:

$$\left|\mathbb{E}\left[f(x) \cdot \mathrm{e}\big(-P(x)\big)\right]\right| = U^1\big(f \cdot \mathrm{e}(-P)\big)$$
$$\leq U^{d+1}\big(f \cdot \mathrm{e}(-P)\big)$$
$$= U^{d+1}(f)$$

Gowers Inverse Theorem

Theorem: If $d < |\mathbb{F}|$, for all $\epsilon > 0$, there exists $\delta = \delta(\epsilon, d, \mathbb{F})$ such that if $U^{d+1}(f) > \epsilon$, then $|\langle f, \mathrm{e}(P)\rangle| > \delta$ for some poly P of degree d.

Proof:

- **[Green-Tao '09]** Combinatorial for phase poly f (c.f. Madhur's talk later).
- **[Bergelson-Tao-Ziegler '10]** Ergodic theoretic proof for arbitrary f.

Small Fields

Consider $f\colon \mathbb{F}_2^1 \to \mathbb{C}$ with:

$$f(0) = 1$$
$$f(1) = i$$

f not a phase poly but $U^3(f) = 1$!

Small fields: worse news

Consider $f = \mathrm{e}(P)$ where $P: \mathbb{F}_2^n \to \mathbb{F}_2$ is symmetric polynomial of degree 4.

$$U^4(f) = \Omega(1)$$

but:

$$|\langle f, \mathrm{e}(C)\rangle| = \exp(-n)$$

for all cubic poly C.

[Lovett-Meshulam-Samorodnitsky ’08, Green-Tao ‘09]

Nevertheless…

Just ***define*** **non-classical phase polynomials of degree $\boldsymbol{d}$** to be functions $f\colon \mathbb{F}^n \to \mathbb{C}$ such that $|f| = 1$ and

$$\Delta_{h_1}\Delta_{h_2}\cdots\Delta_{h_{d+1}}f(x) = 1$$

for all $x, h_1, \dots, h_{d+1} \in \mathbb{F}^n$

Inverse Theorem for small fields

Theorem: For all $\epsilon > 0$, there exists $\delta = \delta(\epsilon, d, \mathbb{F})$ such that if $U^{d+1}(f) > \epsilon$, then $|\langle f, g \rangle| > \delta$ for some non-classical phase poly g of degree d.

Proof:

- [Tao-Ziegler] Combinatorial for phase poly f .
- [Tao-Ziegler] Nonstandard proof for arbitrary f.

Pseudorandomness & Counting

Theorem: If $L_1, \ldots, L_m$ are m linear forms $(L_j(X_1, \ldots, X_k) = \sum_{i=1}^{k} \ell_{i,j} X_i)$, then:

$$\mathbb{E}_{X_1,\ldots,X_k \in \mathbb{F}^n}\left[\prod_{j=1}^{m} f(L_j(X_1, \ldots, X_k))\right] \le U^t(f)$$

if $f\colon\ \mathbb{F}^n \to \mathbb{C}$ and t is the *complexity* of the linear forms $L_1, \ldots, L_m$.

[Gowers-Wolf '10]

Examples

- If $f\colon\ \mathbb{F}^n \to \{0,1\}$ indicates a subset and we want to count the number of 3-term AP's:

$$\mathbb{E}_{X,Y}[f(X) \cdot f(X+Y) \cdot f(X+2Y)] \leq \sum_{\alpha} \hat{f}^3(\alpha)$$

- Similarly, number of 4-term AP's controlled by 3rd order Gowers norm of f.

- More in Pooya's upcoming talk!

Rank

Rank

Given a polynomial $P\colon\ \mathbb{F}^n \to \mathbb{F}$ of degree d, its **rank** is the smallest integer r such that:

$$P(x) = \Gamma(Q_1(x), \dots, Q_r(x)) \quad \forall x \in \mathbb{F}^n$$

where $Q_1, \dots, Q_r$ are polys of degree $d-1$ and $\Gamma\colon\ \mathbb{F}^r \to \mathbb{F}$ is arbitrary.

Pseudorandomness

- For random poly P of fixed degree d,
$$\mathrm{rank}(P) = \omega(1)$$

- High rank is pseudorandom behavior

Rank & Gowers Norm

If $P\colon\ \mathbb{F}^n \to \mathbb{F}$ is a poly of degree d, P has high rank **if and only if** $\mathrm{e}(P)$ has low Gowers norm of order d!

Low rank implies large Gowers norm

Lemma: If $P(x) = \Gamma(Q_1(x), \ldots, Q_k(x))$ where $Q_1, \ldots, Q_k$ are polys of deg $d-1$, then $U^d\big(\mathrm{e}(P)\big) \geq \frac{1}{|\mathbb{F}|^{k/2}}$.

Low rank implies large Gowers norm

Lemma: If $P(x) = \Gamma(Q_1(x), \dots, Q_k(x))$ where $Q_1, \dots, Q_k$ are polys of deg $d-1$, then $U^d(\mathrm{e}(P)) \geq \frac{1}{|\mathbb{F}|^{k/2}}$.

Proof: By (linear) Fourier analysis:

$$\mathrm{e}(P(x)) = \sum_{\alpha} \hat{\Gamma}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i \cdot Q_i(x)\right)$$

Therefore:

$$|\mathbb{E}_x \sum_{\alpha} \hat{\Gamma}(\alpha) \cdot \mathrm{e}\left(\sum_i \alpha_i \cdot Q_i(x) - P(x)\right)| = 1$$

Then, there's an α such that

$$\langle \mathrm{e}(P), \mathrm{e}(\textstyle\sum_i \alpha_i Q_i) \rangle \geq |\mathbb{F}|^{-k/2}.$$

Inverse theorem for polys

Theorem: For all ϵ and d, there exists $R = R(\epsilon, d, \mathbb{F})$ such that if P is a poly of degree d and $U^d\big(\mathrm{e}(P)\big) > \epsilon$, then rank($P$)$< R$.

[Tao-Ziegler '11]

Bias-rank theorem

Theorem: For all ϵ and d, there exists $R = R(\epsilon, d, \mathbb{F})$ such that if P is a poly of degree d and bias$(P) > \epsilon$, then rank$(P)< R$.

[Green-Tao '09, Kaufman-Lovett '08]

Regularity of Factor

A factor $\mathcal{B} = (P_1, \dots, P_k)$ is ***R*-regular** if every nonzero linear combination of $P_1, \dots, P_k$ has rank more than R.

An Example

Claim: If a factor $\mathcal{B} = (P_1, \ldots, P_k)$ of degree d is sufficiently regular, then for any poly Q of degree d, there can be at most one P_i that is ϵ-correlated with Q.

An Example

Claim: If a factor $\mathcal{B} = (P_1, \ldots, P_k)$ of degree d is sufficiently regular, then for any poly Q of degree d, there can be at most one P_i that is ϵ-correlated with Q.

Proof:

Q ϵ-correlated with $P_i \Rightarrow \mathrm{bias}(Q - P_i) > \epsilon$

Q ϵ-correlated with $P_j \Rightarrow \mathrm{bias}\left(Q - P_j\right) > \epsilon$

So, $\mathrm{rank}(Q - P_i), \mathrm{rank}(Q - P_j)$ bounded. But then $\mathrm{rank}(P_i - P_j)$ bounded, a contradiction.

Things I didn't talk about

- Decomposition theorem
  - For any d, R, ϵ, given function $f\colon \mathbb{F}^n \to \mathbb{C}$, can find functions f_S and f_R such that $f = f_S + f_R$, $U^{d+1}(f_R) < \epsilon$, and $f_S = \Gamma(\mathcal{B})$ for a factor $\mathcal{B}$ of rank R and constant order.
- Gowers' proof of Szemeredi's theorem
- Ergodic-theoretic aspects

Plan for the day

- ~~**Talk 1**: Mathematical primer (me)~~
- **Talk 2**: Polynomial pseudorandomness (P. Hatami)
- **Talk 3**: Algorithmic h.o. Fourier analysis (Tulsiani)
- **Talk 4**: Applications to property testing (Yoshida)
- **Talk 5**: Applications to coding theory (Bhowmick)
- **Talk 6**: A different generalization of Fourier analysis and application to communication complexity (Viola)

A final example

Claim: Let $\mathcal{B} = (P_1, \dots, P_m)$ be a sufficiently regular factor of degree d. Define:

$$F(x) = \Gamma(P_1(x), \dots, P_m(x))$$

Then, for any $Q_1, \dots, Q_m$ with $\deg(Q_j) \leq \deg(P_j)$, if

$$G(x) = \Gamma(Q_1(x), \dots, Q_m(x))$$

it holds that: $\deg(G) \leq \deg(F)$.

[B.-Fischer-Hatami-Hatami-Lovett '13]

Sketch of Proof

- Suppose $D = \deg(F)$.
- Using (standard) Fourier analysis, write:

$$\mathrm{e}(F(x)) = \sum_{\alpha} c_\alpha \mathrm{e}\left(\sum_i \alpha_i P_i(x)\right)$$

- Now, differentiate above expression $D + 1$ times to get 1. But all the derivatives of $\omega^{\sum_i \alpha_i P_i(x)}$ are linearly independent. So, all coefficients of these derivatives cancel formally.
- Can expand out the derivative of $\omega^{G(x)}$ in the same way to get that it too equals 1.