Sum-product estimates in finite fields

Liangpan Li

Department of Mathematics, Shanghai Jiao Tong University
Department of Mathematics, Texas State University

October 2, 2009
Texas State University

Part I: A brief introduction to the sum-product estimates **in the reals**

"Paul Erdős presumably learnt his multiplication tables rather more rapidly than the other students, and was left wondering: How many distinct integers are there in the N-by-N multiplication table?"

– Andrew Granville

$$A = \{1, 2, 3, 4, 5\},$$

$$AA \doteq \{ab : a, b \in A\} = \{1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 16, 20, 25\},$$

$$|AA| = 14 < 25 = |A|^2.$$

- Erdős and Szemerédi (1983) conjectured that for any $\epsilon > 0$ one should have[1]

$$\max\{|A+A|,\ |AA|\} \succeq |A|^{2-\epsilon},$$

where A is any finite subset of $\mathbb{Z}$.

- They proved that there exists $\gamma > 1$ such that

$$\max\{|A+A|,\ |AA|\} \succeq |A|^{\gamma}$$

holds for all finite subsets $A \subset \mathbb{Z}$.

- Question: Find explicit estimates on the expansion index γ.

[1] Throughout this talk $f(A) \succeq g(A)$ means there exists a constant $C > 0$ independent of A such that $f(A) \geq Cg(A)$ holds for all A.

Three supporting examples of the Erdős-Szemerédi conjecture:

- Arithmetic progressions (only the product-sets are large)

$$1, 2, 3, 4, 5, 6, 7, 8, 9, \cdots$$

- Geometric progressions (only the sum-sets are large)

$$1, 2, 4, 8, 16, 32, 64, \cdots$$

- Farey sequences (both the sum-sets and the product-sets are large)

$$F_n = \Big\{\frac{a}{q} : 1 \leq a \leq q \leq n, gcd(a, d) = 1\Big\}.$$

Three well-known results:

- Elekes (1997) applied the Szemeredi-Trotter incidence theorem to obtain

$$\max\{|A+A|,\ |AA|\} \succeq |A|^{5/4}.$$

  Until now this result was cited 30 times (MathSciNet).

- Solymosi (announced in 2008, published in 2009) ingeniously obtained

$$\max\{|A+A|,\ |AA|\} \succeq |A|^{4/3-\epsilon}.$$

  This result was discussed in Netz Hawk Katz's and Izabella Laba's blogs.

- Elekes-Nathanson-Ruzsa (2000) applied the generalized Szemeredi-Trotter incidence theorem to obtain

$$|S+T| \succeq |S|^{1.5}|T|^{0.5},$$

where $S \subset \mathbf{R}^2$ is any finite subset of a strictly convex curve, $T \subset \mathbf{R}^2$ is arbitrary.
This result shows that the union of translations of a strictly convex curve would expand.
The main purpose of this talk is to show the power of the finite field version of the Elekes-Nathanson-Ruzsa estimate.

Part II: A brief introduction to the sum-product estimates **in prime fields**

"Finite field analogs of classical problems in harmonic analysis, geometric measure and combinatorics have received much recent attention due the relative technical transparency afforded by the discrete setting and the presence of fascinating arithmetic considerations."

– Alex Iosevich and Misha Rudnev

"Thomas Wolff had formulated the finite field version of the Kakeya problem in 1996, and had observed that there appeared to be a connection between that conjecture and the finite field version of the Erdős-Szemerédi sum-product problem in 2000."

– Terence Tao

Theorem (Bourgain, Katz, Tao 2003 GAFA + Bourgain, Glibichuk, Konyagin 2006 JLMS)
Let F_p be a field of prime order p and let A be a subset of F_p with $|A| \leq p^{1-\delta}$ for some $\delta > 0$. Then one has a bound of the form

$$\max\{|A+A|,\ |AA|\} \geq C_\delta |A|^\gamma$$

for some $\gamma = \gamma(\delta) > 1$.

Question: Find explicit estimates on the expansion index γ.

Finite fields not of prime order?

Let F_q be a finite field with order q and characteristic p. It is well-known that there exists an $m \in \mathbf{N}$ such that $q = p^m$ and F_q contains a subfield A of order p.

- $m = 1$. Go to Bourgain, Katz, Tao, Glibichuk, Konyagin.
- $m \geq 2$. Since A itself is a field, we have

$$A + A = AA = A$$

and

$$\max\{|A + A|,\ |AA|\} = |A|.$$

This shows that the expanding phenomenon fails for not too large sets in fields not of prime order. How about large sets?

Small sets in finite fields of prime order

F_p: a prime field of order p
A: a subset with $|A| \leq p^{0.5}$.

1. Moubariz Z. Garaev (2007): $\gamma = \frac{15}{14} + o(1)$.
   Remark: $o(1)$ is caused by logarithmic terms.
2. N. H. Katz and Chun-Yen Shen (2008): $\gamma = \frac{14}{13} + o(1)$.
3. J. Bourgain and M. Z. Garaev (2007): $\gamma = \frac{13}{12} + o(1)$.
4. Chun-Yen Shen (2008): $\gamma = \frac{13}{12} + o(1)$.
5. Liangpan Li (2009): $\gamma = \frac{13}{12}$. For any $\oplus \in \{+, -\}$, $\otimes \in \{\times, \div\}$, one has

$$|A \oplus A|^8 \cdot |A \otimes A|^4 \succeq |A|^{13}.$$

How about large sets?

No matter a finite field is prime or not, we are left with large sets.

To be introduced in Part III and Part IV.

Applications

- A wonderful application of the sum-product estimates in prime fields is a work of Harald Helfgott published in Ann. of Math. (2008), who confirmed a special case of a big conjecture in ALGEBRA.
- From the work of Bourgain-Katz-Tao (2003 GAFA) we know that another application of the sum-product estimates in prime fields is to deduce good results towards the finite field Kakeya problem.
  (Zeev Dvir completely confirmed the the finite field Kakeya conjecture in 2008 using different approach.)
- Jean Bourgain could tell us more applications.

Part III: Spectral graph method vs Fourier analytic method

"There are several examples where one can choose either the Fourier method or a proof based on eigenvalues."

– Jozsef Solymosi

Via Fourier analytic method

F_q: a field of order q
A: a subset with $|A| \geq q^{0.5}$.

- Hart, Iosevich and Solymosi (2007) proved

$$\max\{|A+A|,\ |AA|\} \geq C\min\{\frac{|A|^{1.5}}{q^{0.25}},\ |A|^{2/3}q^{1/3}\}$$

via Kloosterman sums

$$K_{a,b} = \sum_{x\in F_p^*} \exp(2\pi i(ax + \frac{b}{x})).$$

This is a pioneer work.

Via Fourier analytic method

- Garaev (2008) improved the HIS estimate to

$$\max\{|A+A|,\ |AA|\} \geq C \min\{\frac{|A|^2}{q^{0.5}},\ |A|^{1/2}q^{1/2}\}.$$

♣ When $|A| \geq q^{2/3}$ this result is optimal up to constant:

$$\max\{|A+A|,\ |AA|\} \geq C|A|^{1/2}q^{1/2}.$$

♣ In PRIME fields when $|A| \leq p^{2/3}$ Garaev conjectured

$$\boxed{\max\{|A+A|,\ |AA|\} \geq C_\epsilon \min\{|A|^{2-\epsilon},\ |A|^{1/2}p^{1/2-\epsilon}\}.}$$

Via spectral graph method

F_q: a field of order q
A: a subset with $|A| \geq q^{0.5}$.

- Le Anh Vinh (2008) reproved the Garaev estimate.
- Van Vu (2008) proved that if P is a **non-degenerate**[2] polynomial then

$$\max\{|A+A|, |P(A,A)|\} \geq C_{deg(P)} \min\{\frac{|A|^{1.5}}{q^{0.25}}, \ |A|^{2/3}q^{1/3}\}.$$

This result characterizes the polynomials for which $\boxed{\square}$ hold.

Particularly, taking $P(x_1, x_2) = x_1x_2$ yields the HIS estimate.

[2] P is said to be **degenerate** if it is of the form $Q \circ L$ where $Q \in F_q[x]$ and $L \in F_q[x_1, x_2]$ is a linear form in x_1, x_2.

Why degenerate polynomials not expand?

P is said to be **degenerate** if it is of the form $Q \circ L$ where $Q \in F_q[x]$ and $L \in F_q[x_1, x_2]$ is a linear form in x_1, x_2. Hence

$$|P(A, A)| = |Q(L(A, A))| \leq |L(A, A)|.$$

For example, let $L(x_1, x_2) = 2x_1 + 3x_2$ and $A = \{1, 2, 3, \ldots, n\}$ where $n < \frac{p}{5}$. Then

$$\max\{|A + A|, |P(A, A)|\} \leq 5|A|.$$

Via spectral graph method

- Jozsef Solymosi (2008) gave a similar bound for a general class of functions $f : F_q \to F_q$ of which polynomials of integer coefficients and degrees greater than one are members. A typical result of Solymosi is

$$|A + f(A)| \geq C_{deg(f)} \min\{\frac{|A|^2}{q^{0.5}}, \ |A|^{1/2}q^{1/2}\}.$$

One may compare this with the Garaev-Vinh estimate

$$\max\{|A + A|, \ |AA|\} \geq C \min\{\frac{|A|^2}{q^{0.5}}, \ |A|^{1/2}q^{1/2}\}.$$

Comparison

- As to the sum-product estimates in finite fields, it seems that the spectral graph method is more stronger than the Fourier analytic method.
- But Jozsef Solymosi himself said "There are several examples where one can choose either the Fourier method or a proof based on eigenvalues."
- Thanks to Jozsef Solymosi's prediction, the presenter along with Derrick Hart and Chun-Yen Shen.....
  Fourier analysis and expanding phenomena in finite fields
  To appear on arXiv.

Part IV: The finite field version of the Elekes-Nathanson-Ruzsa estimate and its consequences

Where we do Fourier analysis?

1. Let $\mathbb{G}^2 = G_1 \times G_2$ where $G_i \in \{F_q, F_q^*\}$ and the group operation $\odot$ is inherited from each coordinate group.
2. A character[3] χ on a finite group $(G, +)$ is a homophormism from $(G, +)$ to $(S^1, \times)$.
3. A character χ on G is said to be trivial if $\chi(g) \equiv 1$ for all $g \in G$.
4. Any character χ on $(\mathbb{Z}_p, +)$ is of the form $\chi(a) = \chi_\xi(a) = \exp(2\pi i \xi a)$ for some $\xi \in \mathbb{Z}_p$.
5. A character χ on the group $(F_q, +)$ is called an additive character of F_q.
6. A character ψ on the group $(F_q^*, \times)$ is called a multiplicative character of F_q.

[3] Pronunciation: $\chi \sim$ chi $\sim$ kai $\quad \psi \sim$ psi $\sim$ psai

Definition of Fourier transform

- Define the Fourier transform of any given function $f : \mathbb{G}^2 \to \mathbb{C}$ by

$$\widehat{f}(\chi_1, \chi_2) = \frac{1}{|\mathbb{G}^2|} \sum_{(x_1,x_2)\in\mathbb{G}^2} f(x_1, x_2) \cdot \overline{\chi_1(x_1)} \cdot \overline{\chi_2(x_2)},$$

  where χ_j denotes the additive or multiplicative character corresponding to G_j is additive or multiplicative.
- Standard example $\mathbb{G}^2 = F_p \times F_p$:

$$\begin{aligned}\widehat{f}(\xi_1, \xi_2) &= \frac{1}{p^2} \sum_{(x_1,x_2)\in\mathbb{F}_p\times F_p} f(x_1, x_2) \cdot \overline{\chi_{\xi_1}(x_1)} \cdot \overline{\chi_{\xi_2}(x_2)} \\ &= \frac{1}{p^2} \sum_{(x_1,x_2)\in\mathbb{F}_p\times F_p} f(x_1, x_2) \cdot \exp(-2\pi i(x_1\xi_1 + x_2\xi_2)).\end{aligned}$$

Concept of Salem sets

- Let Z be a subset of $\mathbb{G}^2$ and $\mathbb{I}_Z$ be its indicator function. Z is said to be a **Salem** set with constant C if

$$\|Z\|_U \doteq \max_{(\chi_1,\chi_2)\neq(\text{trivial, trivial})} |\widehat{\mathbb{I}_Z}(\chi_1,\chi_2)| \geq C\frac{\sqrt{|Z|}}{|\mathbb{G}^2|}.$$

$\|Z\|_U$ is called the *linear bias* or *Fourier bias* of Z.
To compare,

$$|\widehat{\mathbb{I}_Z}(\text{trivial, trivial})| = \frac{|Z|}{|\mathbb{G}^2|} \quad \text{is much bigger.}$$

- Informal statements: **A straight line is not a Salem set, while a "curve" is Salem.**

How to understand the concept of "Salem"?

An informal example in the Euclidean space

$$\mathbb{S}^{n-1} = \left\{(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n : x_1^2 + x_2^2 + \cdots + x_n^2 = 1\right\}.$$

- The Housdorff dimension of $\mathbb{S}^{n-1}$ is $n-1$.
- The Fourier dimension of $\mathbb{S}^{n-1}$ is the supremum of $\alpha \geq 0$ such that

$$|\widehat{\mathbb{I}_{\mathbb{S}^{n-1}}}(\xi)| \leq C_\alpha |\xi|^{-\alpha/2} \quad (\forall \xi \in \mathbf{R}^n),$$

where

$$\widehat{\mathbb{I}_{\mathbb{S}^{n-1}}}(\xi) = \int_{\mathbb{S}^{n-1}} \exp(-ix \cdot \xi) d\mu(x).$$

From Fourier analysis (E. M. Stein and G. Weiss 1970) we know that $\mathbb{S}^{n-1}$ has Fourier dimension $n-1$.

A concrete set that is Salem (**Basic tool: Gaussian sums**)

- $\mathbb{G}^2 = F_p \times F_p$. Then $\boxed{\boxed{Z \doteq \{(x, x^2) : x \in F_p\}}}$ is Salem.

  Suppose $(\xi_1, \xi_2) \neq (0,0) = (\text{trivial}, \text{trivial})$. Then

$$\widehat{\mathbb{I}_Z}(\xi_1, \xi_2) = \frac{1}{p^2} \sum_{(x_1,x_2)\in F_p\times F_p} \mathbb{I}_Z(x_1, x_2) \cdot \exp\left(\frac{-2\pi i(\xi_1 x_1 + \xi_2 x_2)}{p}\right)$$

$$= \frac{1}{p^2} \sum_{x\in F_p} \exp\left(\frac{-2\pi i(\xi_1 x + \xi_2 x^2)}{p}\right)$$

  Case 1: $\xi_2 \neq 0$: Gaussian sums $\Rightarrow |\widehat{\mathbb{I}_Z}(\xi_1, \xi_2)| = \frac{\sqrt{p}}{p^2} = \frac{\sqrt{|Z|}}{p^2}$.

  Case 2: $\xi_2 = 0$. $\xi_1 \neq 0$. Then $|\widehat{\mathbb{I}_Z}(\xi_1, \xi_2)| = 0$.

Gaussian sums:

$$\Big| \sum_{x \in F_p} \exp(2\pi i a x^2) \Big| = \sqrt{p},$$

where $a \in F_p^*$.

More Salem sets (**Basic tool: Weil's bound**)

Let p be the characteristic of $\mathbb{F}_q$. Suppose $f, g \in \mathbb{F}_q[x]$ with $M \doteq deg(f) + deg(g) < p$ and define

$$\boxed{F = \{(f(x), g(x)) \in \mathbb{G}^2\}.}$$

1. Let $G^2 = \mathbb{F}_q \times \mathbb{F}_q$ and suppose $1 \leq deg(f) < deg(g)$. Or,
2. Let $G^2 = \mathbb{F}_q \times \mathbb{F}_q^*$ and suppose $gcd(deg(g), q-1) = 1$. Or,
3. Let $G^2 = \mathbb{F}_q^* \times \mathbb{F}_q^*$. Suppose f contains some irreducible factors that are not factors of g such that the great common divisor of the powers of these factors in the canonical factorization of f is 1, and vice versa.

Then F is a Salem set with constant M.

Weil's bound

Let χ be a non-trivial additive character of $\mathbb{F}_q$ and ψ be a non-trivial multiplicative character of $\mathbb{F}_q$ of order s.

1. Suppose that $f \in \mathbb{F}_q[x]$ satisfies $gcd(deg(f), q) = 1$. Then we have
$$\Big| \sum_{x \in \mathbb{F}_q} \chi(f(x)) \Big| \leq (deg(f) - 1)\sqrt{q}.$$
2. Suppose that $g \in \mathbb{F}_q[x]$ is not, up to a nonzero multiplicative constant, an s-th power of a polynomial in $\mathbb{F}_q[x]$. Then for any $f \in \mathbb{F}_q[x]$ we have
$$\Big| \sum_{x \in \mathbb{F}_q} \chi(f(x))\psi(g(x)) \Big| \leq (deg(f) + d - 1)\sqrt{q},$$
where d is the number of distinct roots of g in its splitting field over $\mathbb{F}_q$.

Main Result

- The Elekes-Nathanson-Ruzsa estimate in the reals:

$$|S+T| \succeq |S|^{1.5}|T|^{0.5},$$

where $S \subset \mathbf{R}^2$ is a finite subset of a strictly convex curve, $T \subset \mathbf{R}^2$ is arbitrary.

- The Hart-Li-Shen estimate in finite fields:

$$\boxed{|X \odot Y| \succeq \min(|\mathbb{G}^2|\frac{|X|}{|Z|}, \frac{|X|^2|Y|}{|Z|}),}$$

where $X \subset (\mathbb{G}^2, \odot)$ is a finite subset of a Salem set Z, $Y \subset (\mathbb{G}^2, \odot)$ is arbitrary.

- Informal statement: **strictly convex curve $\sim$ Salem sets**

A typical application of the Hart-Li-Shen estimate

$(\mathbb{G}^2, \odot) = (F_p, +) \times (F_p, +)$.
$Z \doteq \{(x, x^2) : x \in F_p\}$ is a Salem set.
$X \doteq \{(x, x^2) : x \in A\}$ is a subset of the Salem set Z.
$Y \doteq A^2 \times A$, where $A^2 = \{a^2 : a \in A\}$.

$$|X \odot Y| \succeq \min(|\mathbb{G}^2|\frac{|X|}{|Z|}, \frac{|X|^2|Y|}{|Z|}) \Longrightarrow$$

$$\Longrightarrow |A + A^2|^2 \succeq \min(p^2\frac{|A|}{p}, \frac{|A|^2 \cdot |A||A^2|}{p}) \overset{|A^2| \sim |A|}{\Longrightarrow}$$

$$\Longrightarrow |A + A^2| \succeq \min(p^{0.5}|A|^{0.5}, \frac{|A|^2}{p^{0.5}}).$$

Do you remember this result? Go to Page 22.

Since we have lots of Salem sets, we have lots of sum-product estimates in finite fields.

1. Suppose $1 \leq deg(g) < deg(f) < p$. Then

$$\boxed{|f(A)+g(A)| \geq C_{deg(f)} \min(p^{0.5}|A|^{0.5}, \quad \frac{|A|^2}{p^{0.5}})}$$

2. Suppose f, g satisfy some conditions. Then

$$\boxed{|f(A)g(A)| \geq C_{f,g} \min(p^{0.5}|A|^{0.5}, \quad \frac{|A|^2}{p^{0.5}}).}$$

3. Suppose f, g satisfy some conditions. Then

$$\max(|f(A)+B|, |g(A)C|) \geq C_{f,g} \min(p^{0.5}|A|^{0.5}, \frac{|A||B|^{0.5}|C|^{0.5}}{p^{0.5}}).$$

Another application: reprove a result of Van Vu

F_q: a field of order q
A: a subset with $|A| \geq q^{0.5}$.

- Van Vu (2008) proved that if P is a **non-degenerate**[4] polynomial then

$$\max\{|A+A|, |P(A,A)|\} \geq C_{deg(P)} \min\{\frac{|A|^{1.5}}{q^{0.25}}, \ |A|^{2/3}q^{1/3}\}.$$

This result characterizes the polynomials for which $\boxed{\ }$ hold.

[4] P is said to be **degenerate** if it is of the form $Q \circ L$ where $Q \in F_q[x]$ and $L \in F_q[x_1, x_2]$ is a linear form in x_1, x_2.

An informal proof

Step 1: Decompose $F_q \times F_q$ into the level sets of P:

$$F_q \times F_q = \bigcup_{z \in F_q} P^{-1}(z)$$

Assume for all $z \in F_q$: $P^{-1}(z)$ is a Salem set with $|P^{-1}(z)| \sim q$.

Step 2:

$$|A|^2 = \sum_{z \in P(A,A)} \left| P^{-1}(z) \cap (A \times A) \right|$$

$$\stackrel{\exists z^*}{\Longrightarrow} \frac{|A|^2}{|P(A,A)|} \leq \left| P^{-1}(z^*) \cap (A \times A) \right|.$$

Step 3: $Z = P^{-1}(z^*)$, $X = P^{-1}(z^*) \cap (A \times A)$, $Y = A \times A$.

Applying the Hart-Li-Shen estimate yields exactly Van Vu's result.

This is a wrong proof since we couldn't assume that for all $z \in F_q$, $P^{-1}(z)$ is a Salem set. (Example: $P(x_1, x_2) = x_1 x_2$.)

But a correct proof is 90% identical to the above one since one can prove that for most of z, $P^{-1}(z)$ is a Salem set.

Lemma 1

Suppose that $Z, Y, P \subset (\mathbb{G}^2, \odot)$. *Then*

$$|\{(z,y) \in Z \times Y : z \odot y \in P\}| \leq \frac{|Z||Y||P|}{|\mathbb{G}^2|} + \|Z\|_U \sqrt{|Y||P|}|\mathbb{G}^2|.$$

Proof:

$$|\{(z,y) \in Z \times Y : z \odot y \in P\}| = \sum_w (Z * Y)(w)P(w) \overset{\text{Plancherel}}{=}$$

$$= |\mathbb{G}^2| \sum_\chi \widehat{Z * Y}(\chi)\overline{\widehat{P}(\chi)} \overset{\text{Convolution}}{=} |\mathbb{G}^2|^2 \sum_\chi \widehat{Z}(\chi)\widehat{Y}(\chi)\overline{\widehat{P}(\chi)}$$

$$= \frac{|Z||Y||P|}{|\mathbb{G}^2|} + |\mathbb{G}^2|^2 \|Z\|_U \sum_{\chi \neq \text{trivial}} \widehat{Y}(\chi)\overline{\widehat{P}(\chi)},$$

which in turn by CS and Plancherel proves the lemma.

Proof of the Hart-Li-Shen estimate

Let Z be a Salem set with constant C and X be one of its subset. Choose $P = X \odot Y$. Then

$$\begin{aligned} |X||Y| &= |\{(x,y) \in X \times Y : x \odot y \in P\}| \\ &\leq |\{(z,y) \in Z \times Y : z \odot y \in P\}| \\ &\leq \frac{|Z||Y||P|}{|\mathbb{G}^2|} + \|Z\|_U \sqrt{|Y||P|}|\mathbb{G}^2|, \end{aligned}$$

which implies

$$\begin{aligned} |X \odot Y| = |P| &\geq 0.5 \min\{|\mathbb{G}^2|\frac{|X|}{|Z|}, \ \frac{|X|^2|Y|}{\|Z\|_U^2|\mathbb{G}^2|^2}\} \\ &\geq \frac{0.5}{C^2} \min\{|\mathbb{G}^2|\frac{|X|}{|Z|}, \ \frac{|X|^2|Y|}{|Z|}\}. \end{aligned}$$

Thank you for your attention