What is a Regularity Lemma?

Arithmetic Regularity, Stable Regularity & Test Problems

Ming Ng

m.ng@qmul.ac.uk

Queen Mary, University of London

Retracing our Steps

Manin's Conjecture

Setup:

- X is a suitable[a] Fano variety over a number field k.
- $\mathcal{L}$ is an ample line bundle on X, and $H_{\mathcal{L}}\colon X(k) \to \mathbb{R}_{>0}$ is the height function associated to $\mathcal{L}$.
- For any subset $Q \subset X(k)$, define the counting function

$$N_{\mathcal{L}}(Q, B) := \#\{x \in Q \mid H_{\mathcal{L}}(x) \leq B\}.$$

[a] Smooth, projective, $X(k)$ is Zariski dense in X.

Manin's Conjecture

Setup:

- X is a suitable[a] Fano variety over a number field k.
- $\mathcal{L}$ is an ample line bundle on X, and $H_{\mathcal{L}} : X(k) \to \mathbb{R}_{>0}$ is the height function associated to $\mathcal{L}$.
- For any subset $Q \subset X(k)$, define the counting function

$$N_{\mathcal{L}}(Q, B) := \#\{x \in Q \,|\, H_{\mathcal{L}}(x) \leq B\}.$$

Then, there exists an exceptional set $Z \subset X(k)$ such that

$$N_{\mathcal{L}}(X(k) \setminus Z, B) \sim cB^{a} \log B^{b-1}, \qquad B \to \infty,$$

where c is Peyre's constant.

[a] Smooth, projective, $X(k)$ is Zariski dense in X.

Manin's Conjecture & Equidistribution

In particular, Manin's Conjecture splits into two sub-problems:

Sub-problem #1: Identify the exceptional set

Sub-problem #2: Bound the growth of the remaining points

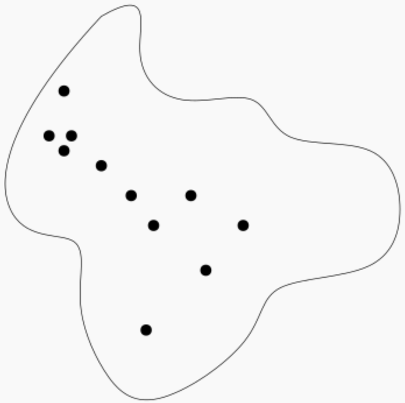
In fact, following Peyre [Pey95], one can reformulate Manin's Conjecture as holding that rational points of Fano varieties are equidistributed outside the exceptional thin set.

Analysing Distribution of Rational Points

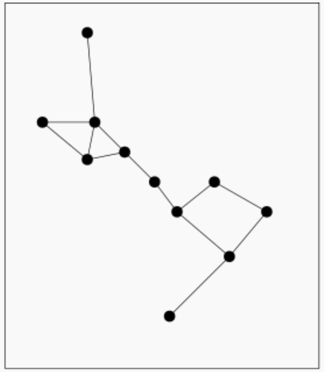
Original Idea

Let us consider the set of rational points $X(k)$ as vertices of a graph:

- Associate a vertex to each rational point.
- Draw an edge between two vertices if their corresponding rational points are “close” to each other.



Rational Points on a Variety



Graph

On “Closeness”

Two Basic Questions

1) What exactly does it mean for rational points to be close to each other?
2) What can this notion of closeness tell us about the distribution of rational points?

Discussion.

On “Closeness”

Two Basic Questions

1) What exactly does it mean for rational points to be close to each other?
2) What can this notion of closeness tell us about the distribution of rational points?

Discussion. McKinnon [McK11] proved that rational points tend to repel each other, and used this to prove Batryev-Manin’s Conjecture for $K3$ surfaces ...

On “Closeness”

Two Basic Questions

1) What exactly does it mean for rational points to be close to each other?
2) What can this notion of closeness tell us about the distribution of rational points?

Discussion. McKinnon [McK11] proved that rational points tend to repel each other, and used this to prove Batryev-Manin’s Conjecture for $K3$ surfaces ... but we can’t use this repulsion principle directly since it only applies to varieties with Kodaira dim ≥ 0.

On “Closeness”

Two Basic Questions

1) What exactly does it mean for rational points to be close to each other?
2) What can this notion of closeness tell us about the distribution of rational points?

Discussion. **Week 5 (Local Distribution):**

On "Closeness"

Two Basic Questions

1) What exactly does it mean for rational points to be close to each other?
2) What can this notion of closeness tell us about the distribution of rational points?

Discussion. **Week 5 (Local Distribution)**: Fixing some rational point **P**, what is the distribution of other rational points (of bounded height) close to **P** ...?

On “Closeness”

Two Basic Questions

1) What exactly does it mean for rational points to be close to each other?
2) What can this notion of closeness tell us about the distribution of rational points?

Discussion. **Week 5 (Local Distribution)**: Fixing some rational point **P**, what is the distribution of other rational points (of bounded height) close to **P** ...?

- We saw how a fixed rational point **P** repels other points, but more work needs to be done to quantify how *any* two rational points **P**, **Q** cannot get too close.
- We also don't have a clear understanding of how the local distribution changes depending on whether **P** is the exceptional set or not.

A Different Angle

Question

What are the key features of the regularity lemma?

A Different Angle

Question

What are the key features of the regularity lemma?

General Proof Strategy.

Step 1: Start with a suitable notion of pseudo-randomness (“regularity”) for the problem. Then, partition the object into pieces of the same size.

Step 2: Continue refining this partition so that we obtain an “essentially regular” partition of the object, possibly with a small percentage of irregularity.

A Different Angle

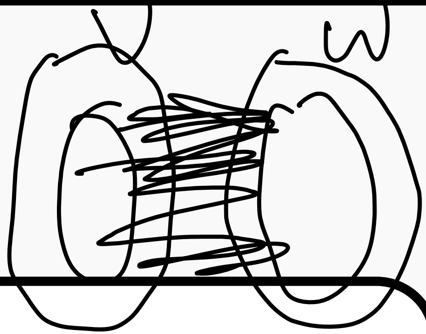
Question

What are the key features of the regularity lemma?

What is regularity? Given a set X, define a measure μ on subsets of X.

Say that X is **regular** if for any good subset $X' \subseteq X$, the expected measure of X' is close to the actual measure of X'.

Week 6 (Graph Regularity)



ϵ-regular

Let G be a graph, and $U, W \subseteq V(G)$. We call (U, W) an ϵ-regular pair if for all ϵ-**good** subsets[a] $A \subseteq U$ and $B \subseteq W$, we have

$$|d(A, B) - d(U, W)| \leq \epsilon.$$

$$d(A,B) = \frac{e(A,B)}{|A||B|}$$

[a]That is, $|A| \geq \epsilon|U|$ and $|B| \geq \epsilon|W|$.

Week 2 (Equidistribution of Rational Points)

Peyre's Equidistribution

Let V be a Fano variety, ω_V^{-1} be the anti-canonical divisor, and **h** be a height associated to it. For any open subset $U \subseteq V$,define

$$\underline{n_{h,W}} := \#\{x \in U(k) \cap W : \mathbf{h}(x) \leq H\}.$$

We say that the rational points of V are *equidistributed on U* if there is a height function such that for all **good** opens[a] $W \subseteq U(\mathbb{A}_k)$

$$\lim_{H \to \infty} \frac{n_{\mathbf{h},W}(H)}{n_U(H)} = \frac{\omega_{\mathbf{h}}(\overline{V(k)} \cap W)}{\omega_{\mathbf{h}}(\overline{V(k)})}$$

normalised

Tamagawa measure.

[a]Call an open $W \subseteq V(\mathbb{A}_k)$ **good** if there is a height **h** s.t. $\omega_{\mathbf{h}}(\partial W) = 0$.

A New Test Problem

Geometric Regularity

Prove an analogue of the Regularity Lemma in the setting of rational points on Fano varieties.

Discussion.

A New Test Problem

Geometric Regularity

Prove an analogue of the Regularity Lemma in the setting of rational points on Fano varieties.

Discussion.

- No need to define when two rational points are close!

A New Test Problem

Geometric Regularity

Prove an analogue of the Regularity Lemma in the setting of rational points on Fano varieties.

Discussion.

- No need to define when two rational points are close!
- Some technicalities (e.g. what exactly does a partition mean in this setting?), but natural to ask: can we adapt Szemerédi's proof strategy here?

A New Test Problem

Geometric Regularity

Prove an analogue of the Regularity Lemma in the setting of rational points on Fano varieties.

Discussion.

- No need to define when two rational points are close!
- Some technicalities (e.g. what exactly does a partition mean in this setting?), but natural to ask: can we adapt Szemerédi's proof strategy here?
- Morally, geometric regularity should say: $X(k)$ can be broken up into pieces which are *locally equidistributed*, up to some small error …

A New Test Problem

Geometric Regularity

Prove an analogue of the Regularity Lemma in the setting of rational points on Fano varieties.

Discussion.

- No need to define when two rational points are close!
- Some technicalities (e.g. what exactly does a partition mean in this setting?), but natural to ask: can we adapt Szemerédi's proof strategy here?
- Morally, geometric regularity should say: $X(k)$ can be broken up into pieces which are *locally equidistributed*, up to some small error ... relation to local distribution? To global equidistribution (Manin-Peyre Conjecture)?

Geometric Regularity

Prove an analogue of the Regularity Lemma in the setting of rational points on Fano varieties.

Discussion. The idea of proving a regularity lemma for objects which are not graphs is not new: arthimetic regularity.

Arithmetic Regularity

Question: What is arithmetic regularity, and how can it help us count solutions?

Szemerédi's Theorem (Finitary Version)

Define

$$[N] := \{1, 2, \dots, N\}$$

Fix $\delta > 0$ and positive integer k. Then, there exists a positive integer $N = N(k, \delta)$ such that every subset of $[N]$ of size at least δN contains an arithmetic progression of length k.

Szemerédi's Theorem

Szemerédi's Theorem (Finitary Version)

Define

$$[N] := \{1, 2, \dots, N\}$$

Fix $\delta > 0$ and positive integer k. Then, there exists a positive integer $N = N(k, \delta)$ such that every subset of $[N]$ of size at least δN contains an arithmetic progression of length k.

Observation. Roth's Theorem deal with the case of 3-term arithmetic progressions in ~~dense~~ subsets $A \subset [N]$: x, $x + d$, $x + 2d$.

Szemerédi's Theorem (Finitary Version)

Define

$$[N] := \{1, 2, \ldots, N\}$$

Fix $\delta > 0$ and positive integer k. Then, there exists a positive integer $N = N(k, \delta)$ such that every subset of $[N]$ of size at least δN contains an arithmetic progression of length k.

Observation. Roth's Theorem deal with the case of 3-term arithmetic progressions in dense subsets $A \subset [N]$: $x,\ x + d,\ x + 2d$. But notice this is equivalent to finding solutions in A to the linear equation $x + z - 2y = 0$.

General Problem

Problem

Let $A \subset [N]$, with density $\delta := |A|/N$. How do we count solutions in A:

$$\sum_{c_1x_1+\dots c_sx_s=b} 1_A(x_1)\dots 1_A(x_s)\,?$$

General Problem

Problem

Let $A \subset [N]$, with density $\delta := |A|/N$. How do we count solutions in A:

$$\sum_{c_1x_1+\ldots c_sx_s=b} 1_A(x_1)\ldots 1_A(x_s)\,?$$

Here's a general heuristic:

- A is Fourier uniform $\implies$ obtain asymptotic count of solutions on A
- A is an arbitrary dense set $\implies$ obtain a local asymptotic count.
- To count solutions globally, use arithmetic regularity.

Fourier Uniformity & Counting

Warm-up Exercise

Given $A \subset [N]$ with density $\delta := |A|/N$, what size is

$$\sum_{x+y=z} 1_A(x)1_A(y)1_A(z)?$$

Warm-up Exercise

Given $A \subset [N]$ with density $\delta := |A|/N$, what size is

$$\sum_{x+y=z} 1_A(x)1_A(y)1_A(z)?$$

$x+y-z=0$

Equidistribution Heuristic. For $x+y-z=c$, is there anything special about the case $c=0$? Perhaps some differences emerge when c gets to the extreme ends of the interval $(-N,2N)$, but for most $c \in (-N,2N)$, we expect:

$$\sum_{x+y-z=0} 1_A(x)1_A(y)1_A(z) \asymp \sum_{x+y-z=c} 1_A(x)1_A(y)1_A(z).$$

Warm-up Exercise

Given $A \subset [N]$ with density $\delta := |A|/N$, what size is

$$\sum_{x+y=z} 1_A(x)1_A(y)1_A(z)?$$

Equidistribution Heuristic. Tidying this up:

$$\begin{aligned}\sum_{x+y-z=c} 1_A(x)1_A(y)1_A(z) &\asymp \mathop{\mathbb{E}}_{c\in(-N,2N)} \sum_{x+y-z=c} 1_A(x)1_A(y)1_A(z)\\ &\asymp \frac{1}{N}\sum_{x,y,z} 1_A(x)1_A(y)1_A(z)\\ &= |A|^3/N = \delta^3 N^2\end{aligned}$$

Random Sets Obey Heuristic

Let $A \subset [N]$, $\mathbb{P}(n \in A) = \delta$. Thus, $\mathbb{E}|A| = \delta N$.

What is the expected number of solutions in A?

Random Sets Obey Heuristic

Let $A \subset [N]$, $\mathbb{P}(n \in A) = \delta$. Thus, $\mathbb{E}|A| = \delta N$.

What is the expected number of solutions in A?

$z \geq 1$

# of all pairs (x, y) s.t. $x + y = z$

$\sim |z - 1|$.

$$\mathbb{E}\left(\sum_{x+y=z} 1_A(x)1_A(y)1_A(z)\right) \approx \sum_{x+y=z} \mathbb{P}(x, y, z \in A)$$

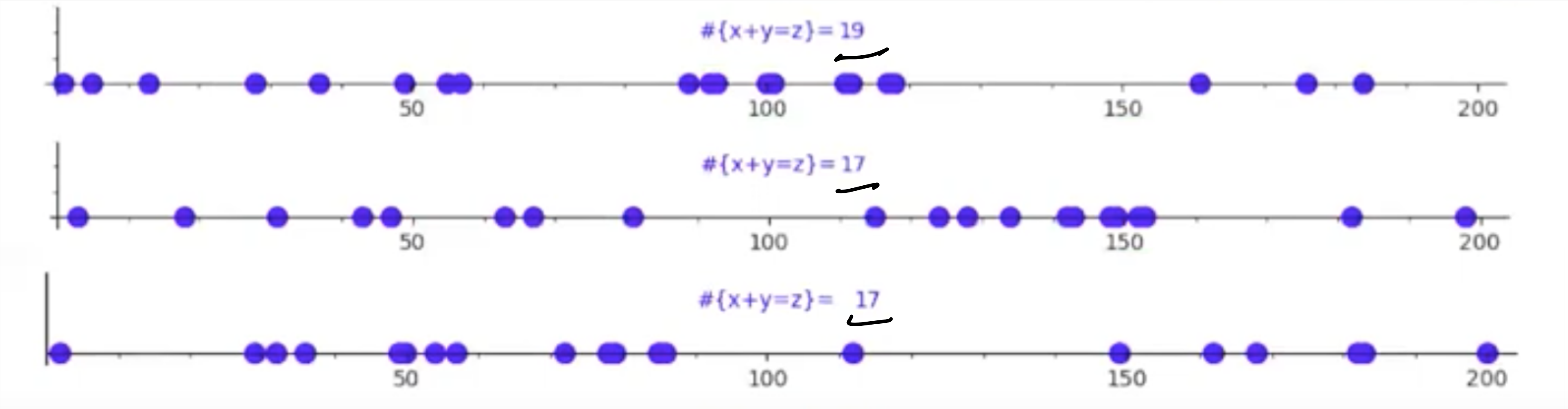
$$\approx \delta^3 \sum_{x+y=z} 1_{[N]}(x)1_{[N]}(y)1_{[N]}(z)$$

$$\approx \delta^3 \times \frac{1}{2}N^2(1 + o(1))$$

$\frac{N(N-1)}{2}$

Random Sets Obey Heuristic

Let $N = 200$. Let $|A| = 20$. So $\delta = \frac{1}{10}$. Naive heuristic predicts that we get around $(\frac{1}{10})^3 \times (200)^2 = 20$ solutions.



Structured Sets do not obey Heuristic

10 20 30 200

5 15 25 ....

However, the same is not true for structured sets.

- $A = \{x \in [N] | 10b,\ b \in \mathbb{N}\}$. We get 190 solutions to $x + y = z$.
- $A = \{x \in [N] | 10b + 5, b \in \mathbb{N}\}$. We get 0 solutions.
- $A = \{1, 2, \dots, 20\}$. We get 190 solutions.
- $A = \{91, \dots, 110\}$. We get 0 solutions.

What is a “structured set”?

For us, a structured set is a **Bohr translate**.

What is a “structured set”?

For us, a structured set is a **Bohr translate**.

Start by embedding $[N]$ homomorphically into a circle:

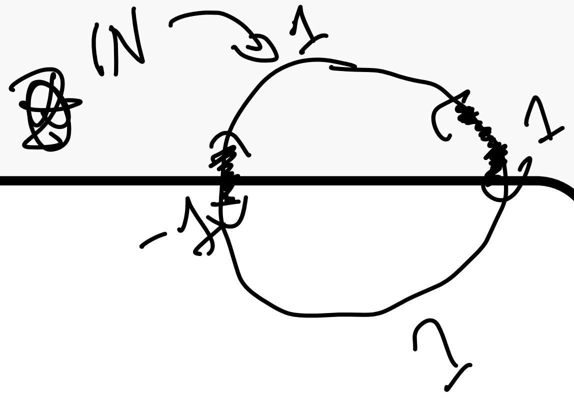
- Write $S^1 := \{z \in \mathbb{C} : |z| = 1\}$ for the unit circle in plane, a group under multiplication.
- Any homomorphism from $(\mathbb{Z}, +) \to S^1$ is of the form $n \mapsto e(\alpha n)$ for some $\alpha \in \mathbb{T}$, where $e(\alpha) := e^{2\pi i \alpha}$ is the complex exponential, and $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ is the torus.

Definition

For any $\alpha \in \mathbb{R}$, define $||\alpha||_{\mathbb{T}}$ to give the distance to the nearest integer:

$$||\alpha||_{\mathbb{T}} := \min_{n \in \mathbb{Z}} |\alpha - n|.$$

What is a “structured set”?



Bohr Set & Bohr Translate

Let $S \subset \mathbb{T}$ and $\epsilon > 0$.

- A 1-dim Bohr set pulls back an arc around 1.

$$B(\alpha, \epsilon) := \{x \in \mathbb{Z} : ||\alpha x||_{\mathbb{T}} \leq \epsilon\}$$

- A 1-dim Bohr translate takes a general arc, not necessarily around 1.

$$B(\alpha, I) := \{x \in \mathbb{Z} : n\alpha \in I\}$$

What is a "structured set"?

Bohr Set & Bohr Translate

Let $S \subset \mathbb{T}$ and $\epsilon > 0$.

- A 1-dim Bohr set pulls back an arc around 1.

$$B(\alpha, \epsilon) := \{x \in \mathbb{Z} : ||\alpha x||_{\mathbb{T}} \leq \epsilon\}$$

- A 1-dim Bohr translate takes a general arc, not necessarily around 1.

$$B(\alpha, I) := \{x \in \mathbb{Z} : n\alpha \in I\}$$

What is a "structured set"?

Bohr Set & Bohr Translate

Let $S \subset \mathbb{T}$ and $\epsilon > 0$.

- A d-dim Bohr set:

$$B(\vec{\alpha}, \epsilon) := B(\alpha_1, \epsilon) \cap \cdots \cap B(\alpha_d, \epsilon)$$

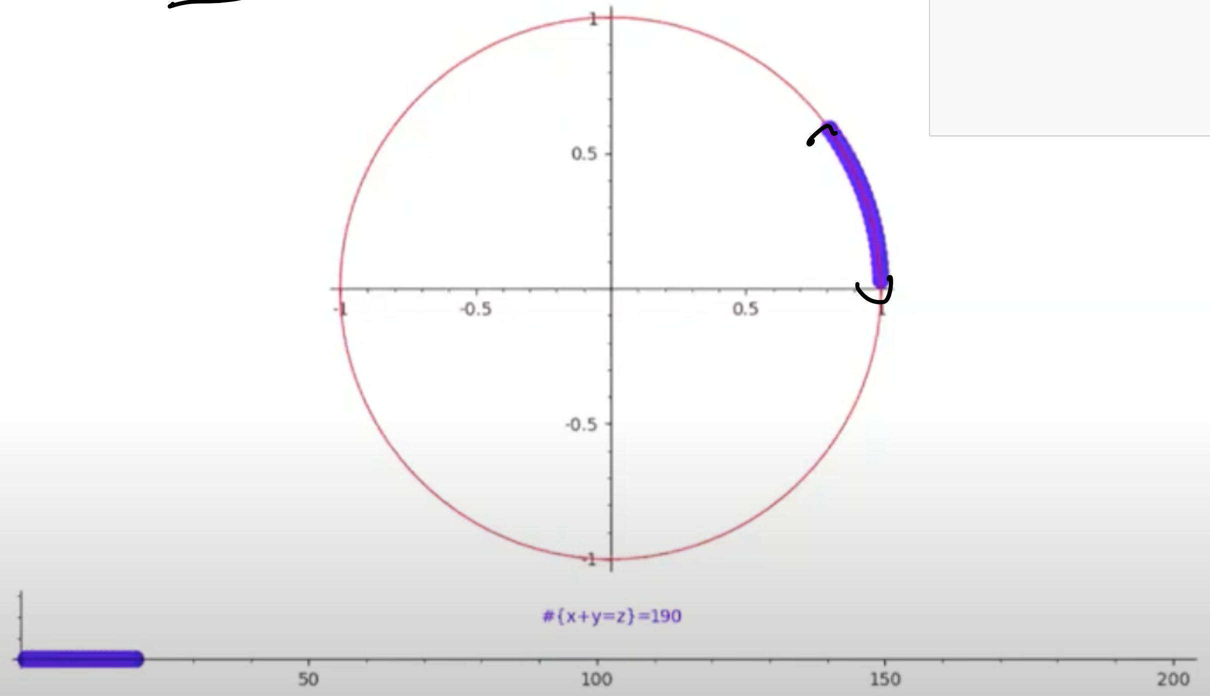
- A d-dim Bohr translate translate takes a general arc:

$$B(\alpha, I) := \{x \in \mathbb{Z} : n\alpha \in I\}$$

  does this for $I = I_1 \times \cdots \times I_d \subset \mathbb{T}^d$ is a product of intervals and $\vec{\alpha} = (\alpha_1, \dots, \alpha_d)$.

Bohr Translates

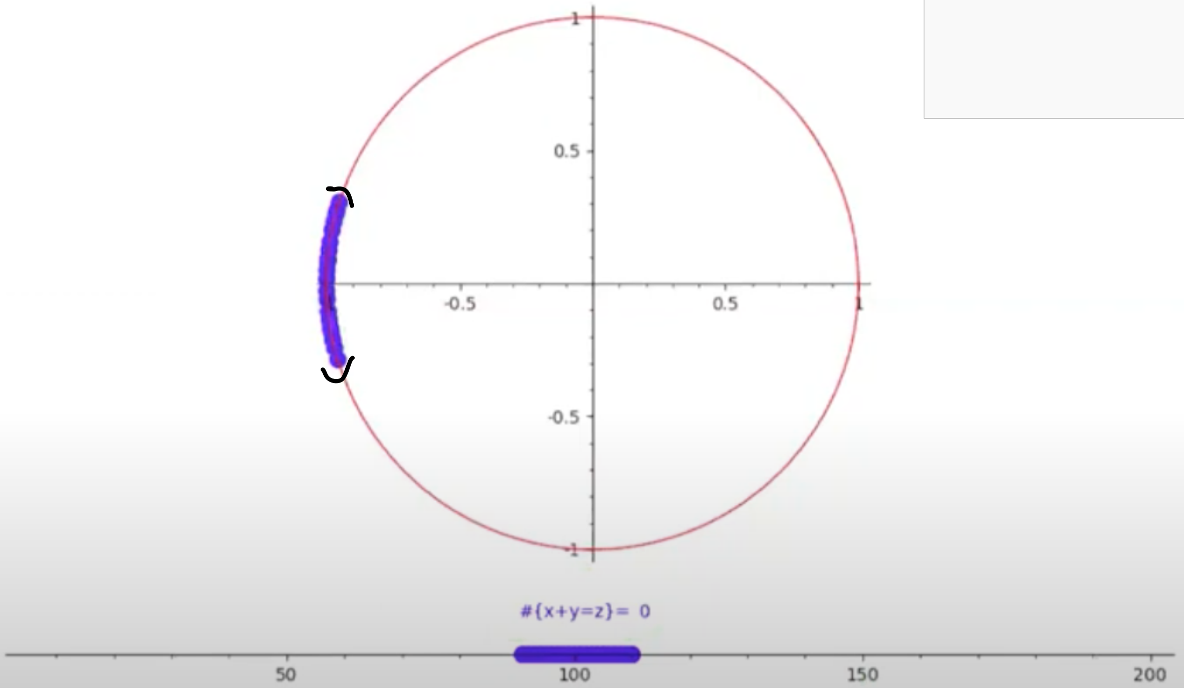
$n \mapsto e(n/200),$



$\{n \in [200] : \|n/200\|_{\mathbb{T}} \leqslant 1/10\}$

Bohr Translates

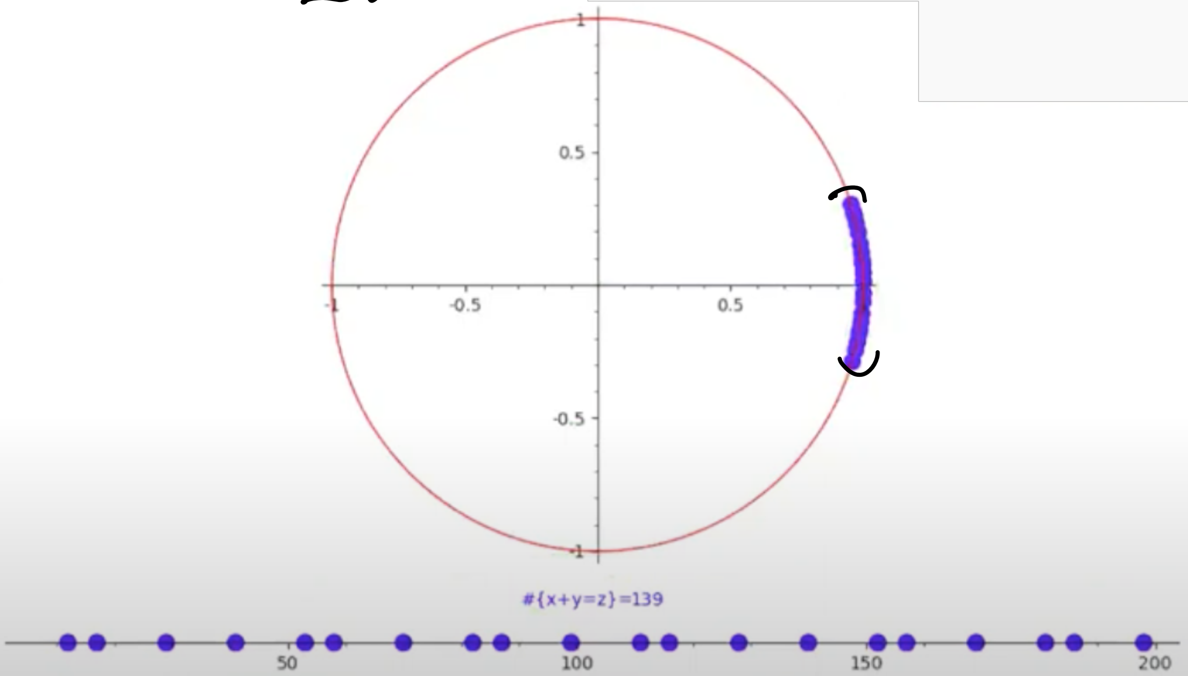
$n \mapsto e(n/200),$



$\{n \in [200] : \|n/200\|_{\mathbb{T}} \geqslant 9/20\}$

Bohr Translates

$n \mapsto e(n\sqrt{2})$,

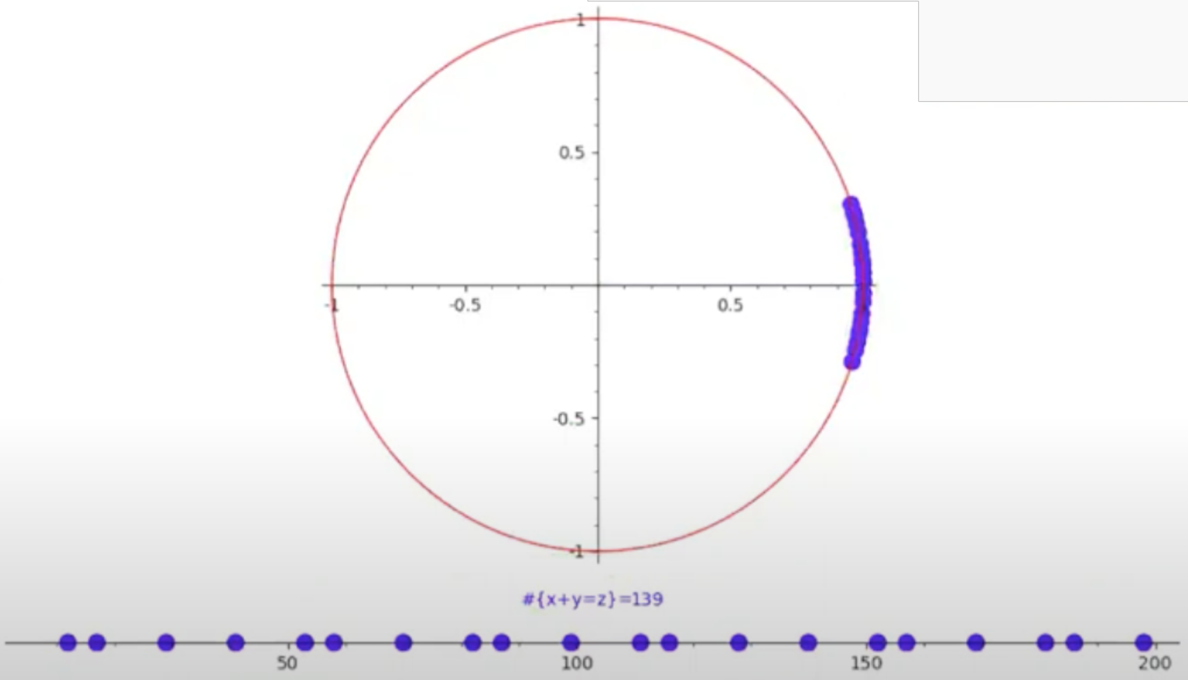


$\{n \in [200] : \|n\sqrt{2}\|_{\mathbb{T}} \leqslant 1/20\}$

190 solns

Bohr Translates

$n \mapsto e(n\sqrt{2}),$



$\{n \in [200] : \|n\sqrt{2}\|_{\mathbb{T}} \leqslant 1/20\}$

How do we detect Bohr structures?

A Bohr structure is essentially when you pullback an arc along your homomorphism $\mathbb{Z} \to S^1$. We want a way of describing when we have dense clustering in S^1.

How do we detect Bohr structures?

A Bohr structure is essentially when you pullback an arc along your homomorphism $\mathbb{Z} \to S^1$. We want a way of describing when we have dense clustering in S^1.

Fourier Transform

Given a finitely supported[a] function $f : \mathbb{Z} \to \mathbb{C}$ define its Fourier transform by

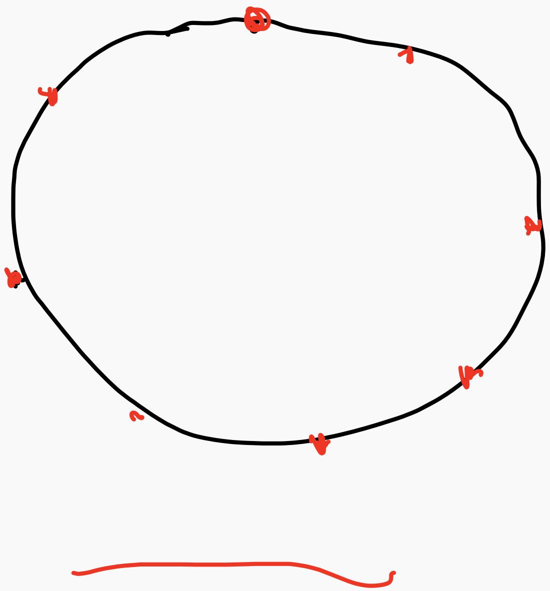
$$\hat{f}(\alpha) := \sum_x f(x)e(\alpha x).$$

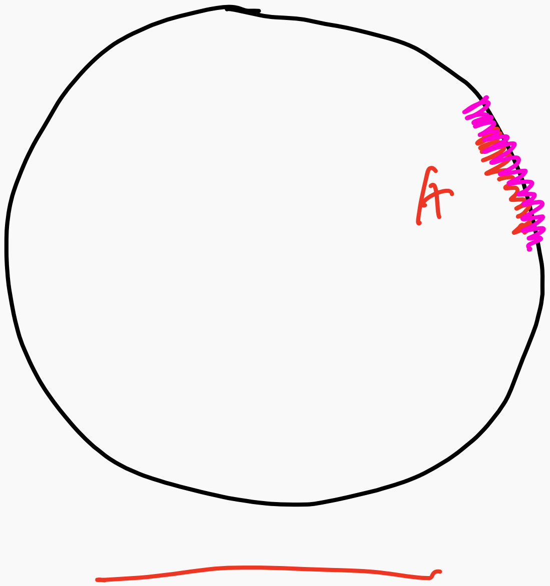
$f :=$ "1_A"

[a] $f(x) \neq 0$ for only finitely many integers.

Let $f = 1_A$ for A a finite set of integers A. Then, the Fourier transform $\hat{1}_A(\alpha)$ is an average of A embedded into S^1.

$[N] \hookrightarrow S^1$





How do we detect Bohr structures?

Let $f = 1_A$ for A a finite set of integers A. Then, the Fourier transform $\hat{1}_A(\alpha)$ is an average of A embedded into S^1.

Fourier Uniformity

A set $A \subset [N]$ of density $\delta := |A|/N$ is ϵ-*Fourier uniform* on $[N]$ if

$$||\hat{1}_A - \delta \hat{1}_{[N]}||_\infty \leq \epsilon ||\hat{1}_A||_\infty$$

Equivalently,

$$\forall \alpha \in \mathbb{T} \qquad |\mathop{\mathbb{E}}_{n \in A} e(\alpha n) - \mathop{\mathbb{E}}_{n \in [N]} e(\alpha n)| \leq \epsilon$$

Key Takeaways

Fourier non-uniformity $\implies$ Bohr Bias

Let $A \subset [N]$. Suppose A has density $\delta := |A|/N$ and A is not ϵ-Fourier uniform. Then, there exists a 1-d Bohr translate B s.t.

$$|A \cap B| \geq \frac{|A||B|}{N} + \frac{\epsilon}{\pi}|A|.$$

Fourier Uniform $\implies$ Naive Heuristic

Let $A \subset [N]$. Let $b, c_1, \dots, c_s \in \mathbb{Z}$ with $c_i \neq 0$ and $s \geq 3$. Suppose A has density $\delta := |A|/N$, A is ϵ-Fourier uniform. Then,

$$\sum_{c_1x_1+\cdots+c_sx_s=b} 1_A(x_1)\dots 1_A(x_s) = \delta^s \sum_{c_1x_1+\cdots+c_sx_s=b} 1_{[N]}(x_1)\dots 1_{[N]}(x_s) + o(\epsilon N^{s-1})$$

Local Asymptotics for Arbitrary Sets

Question. Under what conditions do subsets of $A \subset [N]$ exhibit 3APs $\{x, x+d, x+2d\}$?

Roth's Theorem

Fix $\delta > 0$. Then, either:

- $\forall A \subset [N]$ with $|A| \geq \delta N$, $\exists 3AP \in A$ $(d \neq 0)$
- or $N \ll_\delta 1$.

Fourier-Uniform Sets contain 3APs

Recall: (x, y, z) is a 3AP iff $x - 2y + z = 0$.

Fourier-Uniform Sets contain 3APs

Recall: (x, y, z) is a 3AP iff $x - 2y + z = 0$. Therefore, if $A \subset [N]$ has density δ and is ϵ-Fourier Uniform, the number of 3APs satisfies the asymptotic.

$$\sum_{x,d} 1_A(x)1_A(x+d)1_A(x+2d) = \delta^3 \sum_{x,d} 1_{[N]}(x)1_{[N]}(x+d)1_{[N]}(x+2d) + o(\epsilon N^2). \qquad (1)$$

Observation

$$x, z, \boxed{\frac{1}{2}(x+z)} \in [N] \iff x, z \in [N] \,\&\, x \equiv z \pmod 2.$$

Fourier-Uniform Sets contain 3APs

Recall: (x, y, z) is a 3AP iff $x - 2y + z = 0$. Therefore, if $A \subset [N]$ has density δ and is ϵ-Fourier Uniform, the number of 3APs satisfies the asymptotic.

$$\sum_{x,d} 1_A(x)1_A(x+d)1_A(x+2d) = \delta^3 \sum_{x,d} 1_{[N]}(x)1_{[N]}(x+d)1_{[N]}(x+2d) + o(\epsilon N^2). \tag{1}$$

Observation

$$x, z, \frac{1}{2}(x+z) \in [N] \iff x, z \in [N] \,\&\, x \equiv z \pmod 2.$$

And thus:

$$\sum_{x,d} 1_{[N]}(x)1_{[N]}(x+d)1_{[N]}(x+2d) = \lfloor N/2 \rfloor^2 + \lceil N/2 \rceil^2 = \frac{1}{2}N^2(1+o(1)).$$

Local Fourier Uniformity

Question. Not all dense subsets $A \subset [N]$ are Fourier uniform. What do we do then?

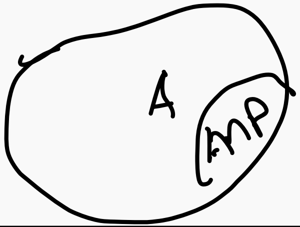
Local Fourier Uniformity

Question. Not all dense subsets $A \subset [N]$ are Fourier uniform. What do we do then?

Answer. We restrict to a subset $P \subset [N]$ for which $A \cap P$ is essentially Fourier uniform.

2 Key Lemmas

key behind AP



Lemma 1: Local Fourier Uniformity Exists

Given any $\epsilon > 0$ and any $A \subset [N]$, there exists a function $\omega_\epsilon(N) \to \infty$ and a progression $P \subset [N]$ s.t.

- *A is as dense on P as it is on* $[N]$. $\mathbb{E}_P(1_A) \geq \mathbb{E}_{[N]}(1_A)$;
- *Fourier Uniformity.* $||\hat{1}_{A\cap P} - \mathbb{E}_P(1_A)\hat{1}_P||_\infty \leq \epsilon||\hat{1}_P||_\infty$
- *Arbitrary Length.* $|P| \geq \omega_\epsilon(N)$.

$P = a + q[N]$

$A \cap P$ is "Fourier uniform"

$[1, 2, 3, \ldots, N]$

Lemma 2: Local Fourier Uniformity $\implies$ Naive Heuristic

If

- $A \subset P$
- A has density $\delta := |A|/|P|$
- $||\hat{1}_A - \delta\hat{1}_P||_\infty \leq \epsilon|\hat{1}_P(0)|$.

Then,

$$\sum_{x-2y+z} 1_A(x)1_A(y)1_A(z) = \delta^3 \sum_{x-2y+z=0} 1_{[P]}(x)1_{[P]}(y)1_{[N]}(z) + o(\epsilon|P|^2).$$

Key Observation

Let $P = a + q \cdot [M] = a + \{q, 2q, \dots, qM\}$. Notice 3APs are translation-dilation invariant, i.e.

$$(a + qx) - 2(a + qy) + (a + qz) = 0 \iff x - 2y + z = 0$$

Key Observation

Let $P = a + q \cdot [M] = a + \{q, 2q, \dots, qM\}$. Notice 3APs are translation-dilation invariant, i.e.

$$(a + qx) - 2(a + qy) + (a + qz) = 0 \iff x - 2y + z = 0$$

Hence,

$$\sum_{x-2y+z=0} 1_P(x)1_P(y)1_P(z) = \sum_{x-2y+z=0} 1_{[M]}(x)1_{[M]}(y)1_{[M]}(z).$$

The same argument as before gives the following lower bound

$$\sum_{x-2y+z=0} 1_P(x)1_P(y)1_P(z) \geq \frac{1}{2}|P|^2.$$

Let $A \subset [N]$ with $|A| = \delta N$.

Step 1. By the Local Fourier Uniformity Lemma 1, we obtain a progression $P = a + q \cdot [N_1]$ on which $A \cap P$ is essentially Fourier-uniform & $\omega_\epsilon(N) \le |P|$.

Proving Roth's Theorem

Let $A \subset [N]$ with $|A| = \delta N$.

Step 1. By the Local Fourier Uniformity Lemma 1, we obtain a progression $P = a + q \cdot [N_1]$ on which $A \cap P$ is essentially Fourier-uniform & $\omega_\epsilon(N) \leq |P|$.

Step 2. Apply Lemma 2 and Key Observation to get

$$\sum_{x-2y+z=0} 1_{A\cap P}(x)1_{A\cap P}(y)1_{A\cap P}(z) \geq \delta^3 \sum_{x-2y+z=0} 1_{[P]}(x)1_{[P]}(y)1_{[P]}(z) - o(\epsilon N^2)$$

$$\geq \frac{1}{2}\delta^3|P|^2 - o(\epsilon|P|^2)$$

So if we take ϵ sufficiently small with respect to δ, we obtain a lower bound.

$$\sum_{x-2y+z=0} 1_{A\cap P}(x)1_{A\cap P}(y)1_{A\cap P}(z) \gg \delta^3|P|^2.$$

Proving Roth's Theorem

Let $A \subset [N]$ with $|A| = \delta N$.

Step 2. So if we take ϵ sufficiently small with respect to δ, we obtain a lower bound.

$$\sum_{x-2y+z=0} 1_{A\cap P}(x)1_{A\cap P}(y)1_{A\cap P}(z) \gg \delta^3|P|^2.$$

Step 3. We're almost done, except it's possible that this contains only trivial 3APs – e.g. (x, x, x).

Proving Roth's Theorem

Let $A \subset [N]$ with $|A| = \delta N$.

Step 2. So if we take ϵ sufficiently small with respect to δ, we obtain a lower bound.

$\omega_\delta(N) \leq |P|$

$$\sum_{x-2y+z=0} 1_{A\cap P}(x) 1_{A\cap P}(y) 1_{A\cap P}(z) \gg \delta^3 |P|^2.$$

Step 3. We're almost done, except it's possible that this contains only trivial 3APs – e.g. (x, x, x). In which case, impose a crude bound

$1 \gg \delta^3 |P|$

$$|P| \geq \left(\sum_{x-2y+z=0} 1_{A\cap P}(x) 1_{A\cap P}(y) 1_{A\cap P}(z) \right) \gg \delta^3 |P|^2.$$

In other words, the length of $|P| \ll_\delta 1$. But length of N is also bounded by $|P|$, and so $N \ll_\delta 1$.

Proving Roth's Theorem

Let $A \subset [N]$ with $|A| = \delta N$.

Step 2. So if we take ϵ sufficiently small with respect to δ, we obtain a lower bound.

$$\sum_{x-2y+z=0} 1_{A\cap P}(x)1_{A\cap P}(y)1_{A\cap P}(z) \gg \delta^3|P|^2.$$

Step 3. We're almost done, except it's possible that this contains only trivial 3APs – e.g. (x, x, x). In which case, impose a crude bound

$$|P| \geq \sum_{x-2y+z=0} 1_{A\cap P}(x)1_{A\cap P}(y)1_{A\cap P}(z) \gg \delta^3|P|^2.$$

In other words, the length of $|P| \ll_\delta 1$. But length of N is also bounded by $|P|$, and so $N \ll_\delta 1$.

non-trivial

Step 4. Summarising: Either A contains 3APs, or N is bounded in terms of δ.

#

Arithmetic Regularity

Recap

Let $A \subset [N]$, with density $\delta := |A|/N$. How do we count solutions in A

$$\underbrace{\sum_{c_1x_1+\ldots c_sx_s=b}} \underbrace{1_A(x_1)\ldots 1_A(x_s)}\,?$$

- If A is Fourier Uniform, we have an asymptotic formula for the number of solutions.
- If A is arbitrary dense subset, we can restrict to a progression $P \subset A$ and obtain a local asymptotic there.

Recap

Let $A \subset [N]$, with density $\delta := |A|/N$. How do we count solutions in A

$$\sum_{c_1x_1+\dots c_sx_s=b} 1_A(x_1)\dots 1_A(x_s)\,?$$

- If A is Fourier Uniform, we have an asymptotic formula for the number of solutions.
- If A is arbitrary dense subset, we can restrict to a progression $P \subset A$ and obtain a local asymptotic there.

Arithmetic regularity gives us a framework for **globally** counting solutions on A.

Arithmetic Regularity Lemma, Take # 1

The Fourier transform of $f\colon \mathbb{F}_p^n \to \mathbb{C}$ is a function $\hat{f}\colon \mathbb{F}_p^n \to \mathbb{C}$ defined by

$$\hat{f}(r) := \frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} f(x) e(-r \cdot x)$$

where $r \cdot x = \sum r_i x_i$.

Arithmetic Regularity Lemma, Take # 1

The Fourier transform of $f\colon \mathbb{F}_p^n \to \mathbb{C}$ is a function $\hat{f}\colon \mathbb{F}_p^n \to \mathbb{C}$ defined by

$$\hat{f}(r) := \frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} f(x) e(-r \cdot x)$$

where $r \cdot x = \sum r_i x_i$.

ϵ-Uniformity

We say that $A \subseteq \mathbb{F}_p^n$ is ϵ-**uniform** if $|\hat{1}_A(r)| \leq \epsilon$ for all $r \in \mathbb{F}_p^n \setminus \{0\}$.

The Fourier transform of $f\colon \mathbb{F}_p^n \to \mathbb{C}$ is a function $\hat{f}\colon \mathbb{F}_p^n \to \mathbb{C}$ defined by

$$\hat{f}(r) := \frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} f(x) e(-r \cdot x)$$

where $r \cdot x = \sum r_i x_i$.

ϵ-Uniformity

We say that $A \subseteq \mathbb{F}_p^n$ is ϵ-**uniform** if $|\hat{1}_A(r)| \le \epsilon$ for all $r \in \mathbb{F}_p^n \setminus \{0\}$.
For an affine space $W \subset \mathbb{F}_p^n$, we say that A is ϵ-**uniform on** W if $A \cap W$ is ϵ-uniform when viewed as a subset of W.

Arithmetic Regularity Lemma, Take #1

Green's Arithmetic Regularity [Gre05]

For every $\epsilon > 0$ and prime p, there exists integer M s.t. for every $A \subseteq \mathbb{F}_p^n$, there is some subspace $W \subset \mathbb{F}_p^n$ with codimension at most M s.t. A is ϵ-uniform on all but at most ϵ-fraction of cosets of W.

Green's Arithmetic Regularity [Gre05]

For every $\epsilon > 0$ and prime p, there exists integer M s.t. for every $A \subseteq \mathbb{F}_p^n$, there is some subspace $W \subset \mathbb{F}_p^n$ with codimension at most M s.t. A is ϵ-uniform on all but at most ϵ-fraction of cosets of W.

As an application of this, one obtains:

Roth's Theorem in $\mathbb{F}_3^n$ with popular common difference

For all $\epsilon > 0$, there exists $n_0 = n_0(\epsilon)$ s.t. for all $n \geq n_0$ and every $A \subseteq \mathbb{F}_3^n$ with $|A| = \alpha 3^n$, there exists $y \neq 0$ s.t.

$$|\{x \in \mathbb{F}_3^n : x, x+y, x+2y \in A\}| \geq (\alpha^3 - \epsilon)3^n.$$

Green-Tao Arithmetic Regularity

Question. What if it's not obvious how to neatly partition your object into easily describable pieces?

Green-Tao Arithmetic Regularity

Question. What if it's not obvious how to neatly partition your object into easily describable pieces?

Answer. A perspective, popularised by Tao, is to view regularity as a decomposition result rather than a partition result. More precisely, a regularity lemma is something that tells you:

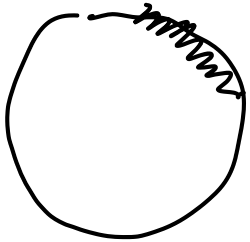
$$A \approx \text{structured part} + \text{uniform error} + \text{small error}.$$

Green-Tao Arithmetic Regularity

Here's a first approximation of the result:

Green-Tao Arithmetic Regularity [GT10]

$$1_A = f_{str} + f_{sml} + f_{unf}$$



where

- f_{str} = Bohr translate
- f_{sml} = small exceptional set
- f_{unf} = Fourier uniform set

error

How does this help us globally count solutions? The slogan:

$$\sum_{c_1x_1+\cdots+c_sx_s=b} 1_A(x_1)\dots 1_A(x_s) \approx \sum_{c_1x_1+\cdots+c_sx_s=b} f_{str}(x_1)\dots f_{str}(x_s).$$

More precisely:

Green-Tao Arithmetic Regularity [GT10]

Let $\epsilon > 0$, and $\mathcal{F}\colon \mathbb{Z}_{\geq 0} \to [0, \infty)$ a growth function. There exists an integer $M = M(\epsilon, \mathcal{F})$ such that for all A:

$$1_A = f_{str} + f_{sml} + f_{unf}$$

satisfying

- $f_{str} = F(\theta n)$ for an M-Lipschitz $F\colon \mathbb{T}^M \to [0,1]$, $\theta \in \mathbb{T}^M$.
- $f_{sml}\colon [N] \to [-1,1]$ s.t. $||f_{sml}||_1$ is small, where $||f_{sml}||_1 \leq \epsilon ||1_{[N]}||_1$
- $f_{unf}\colon [N] \to [-1,1]$ s.t. $||\hat{f}_{unf}||_\infty \leq ||\hat{1}_N||_\infty / \mathcal{F}(M)$

Stable Regularity

Question: What can graph theory tell us about structure vs. randomness?

Graph-theoretic Distortions

In Week 1, I mentioned that part of my original excitement about viewing $X(k)$ as a graph is that this perspective might highlight certain patterns in the distribution of rational points that were previously obscured.

Graph-theoretic Distortions

In Week 1, I mentioned that part of my original excitement about viewing $X(k)$ as a graph is that this perspective might highlight certain patterns in the distribution of rational points that were previously obscured.

If we're no longer viewing $X(k)$ as graph, how might graph theory still give us clues about structure vs. randomness within $X(k)$?

Szemerédi Regularity Lemma (SRL)

For every $\epsilon > 0$, there exists $N(\epsilon)$ s.t. every finite graph G may be partitioned into m classes $V_1 \cup \cdots \cup V_m$ where $m \leq N$ and

- All of the pairs V_i, V_j satisfy $||V_i| - |V_j|| \leq 1$.
- All but at most ϵm^2 of the pairs (V_i, V_j) are ϵ-regular.

Szemerédi Regularity Lemma (SRL)

For every $\epsilon > 0$, there exists $N(\epsilon)$ s.t. every finite graph G may be partitioned into m classes $V_1 \cup \cdots \cup V_m$ where $m \leq N$ and

- All of the pairs V_i, V_j satisfy $||V_i| - |V_j|| \leq 1$.
- All but at most ϵm^2 of the pairs (V_i, V_j) are ϵ-regular.

Discussion. Informally, SRL says any large dense graph G can be approximated by well-behaved pieces. But we can see this approximation isn't perfect:

Szemerédi Regularity (Review)

Szemerédi Regularity Lemma (SRL)

For every $\epsilon > 0$, there exists $N(\epsilon)$ s.t. every finite graph G may be partitioned into m classes $V_1 \cup \cdots \cup V_m$ where $m \leq N$ and

- All of the pairs V_i, V_j satisfy $||V_i| - |V_j|| \leq 1$.
- All but at most ϵm^2 of the pairs (V_i, V_j) are ϵ-regular.

Discussion. Informally, SRL says any large dense graph G can be approximated by well-behaved pieces. But we can see this approximation isn't perfect:

- The number of pieces $N(\epsilon)$ in the partition given by the lemma is potentially huge, a tower of exponentials ($2^{2^{\cdots^{2}}}$)
- There exists ϵ-irregular pairs.

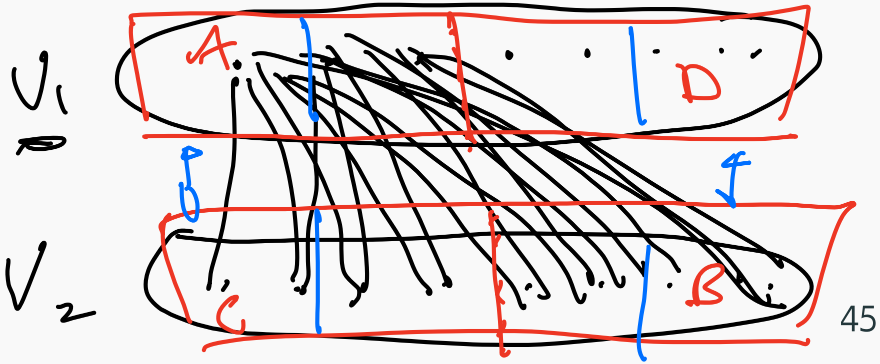
It was asked by Szemerédi if the existence of ϵ-irregular pairs were necessary. It was soon noticed that that the answer was yes.

Half-Graph

A half-graph is a bipartite graph with vertex classes $A = \{a_1, \dots, a_n\}$ and $B = \{b_1, \dots, b_n\}$ where $a_i b_j$ is an edge iff $i \leq j$.

Draw.

$i \le j$



Stable Regularity

Surprisingly, it turns out the only barriers to ϵ-regularity are these half-graphs.

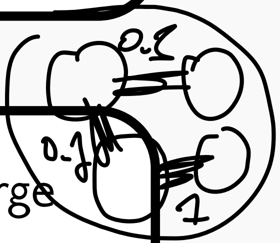
Definition

For $r \in \mathbb{N}$, call a graph G r-**edge stable** if it contains no half-graph of length r.

Surprisingly, it turns out the only barriers to ϵ-regularity are these half-graphs.

Definition

For $r \in \mathbb{N}$, call a graph G r-**edge stable** if it contains no half-graph of length r.



Stable Regularity [MS14]

For each $\epsilon > 0$ and $r \in \mathbb{N}$, there exists $N = N(\epsilon, r)$ s.t. for any sufficiently large finite r-edge stable graph G, for some ℓ with $\ell \leq N$, the graph G can be partitioned into dijsoint pieces $A_1, \ldots, A_\ell$ and:

- The partition is equitable, i.e. the sizes of the pieces differ by at most 1.
- <u>**All**</u> pairs are ϵ-regular, and moreover have density either $> 1 - \epsilon$ or $< \epsilon$.
- $N < (\frac{4}{\epsilon})^{2^{r+3}-7}$.

Stable Arithmetic Regularity

In fact, one can identify analogues of half-graphs in the arithmetic setting.

In fact, one can identify analogues of half-graphs in the arithmetic setting.

Definition

A subset A of a finite abelian group G is k-stable if there does not exist sequences $a_1, \dots, a_k, b_1, \dots, b_k \in G$ s.t. $a_i + b_j \in A$ iff $\boxed{i \leq j}$.

In fact, one can identify analogues of half-graphs in the arithmetic setting.

Definition

A subset A of a finite abelian group G is k-stable if there does not exist sequences $a_1, \dots, a_k, b_1, \dots, b_k \in G$ s.t. $a_i + b_j \in A$ iff $i \leq j$.

Stable Arithmetic Regularity [TW19]

For all $\epsilon \in (0,1)$, $k \geq 2$, and primes p, there exists $n_0 = n_0(k, \epsilon, p)$ s.t. the following holds for all $n \geq n_0$. Suppose that $G := \mathbb{F}_p^n$ and $A \subseteq G$ is k-stable. **Then**, there is a subspace $H \leq G$ of codimension $m \leq M$ s.t. for **any** $g \in G$, either $|(A - g) \cap H| \leq \epsilon|H|$ or $|H \setminus (A - g)| \leq \epsilon|H|$.

Test Problems

Regularity Problems

Our study group has led us to ask questions about the local interactions of rational points, and how this relates to their global distribution.

Problem #1: Geometric Regularity Lemma

Let $X(k)$ be the set of rational points of a Fano variety X/k (maybe to make our lives easier, assume that X is a del Pezzo surface). Prove an analogue of the regularity lemma in this setting.

Problem #2: Extensions of Geometric Regularity

Once we solve Problem #1, there are a couple of important extensions of the regularity lemma that we should also figure out the analogues of. In particular:

- Prove an analogue of the counting lemma. What does this tell us about the distribution of rational points?
- Prove an analogue of the stable regularity lemma. What is the analogue of the half-graph in our setting? What can this tell us about exceptional thin sets?

There are some lingering questions regarding local distribution [Hua17] that appear to be unanswered, and are of independent interest:

Problem # 3: Repulsion

Based on our understanding of local distribution (or otherwise), prove a repulsion principle for rational points for a suitable class of Fano varieties, analogous to the one proved by McKinnon [McK11]. $v(P, Q) \geq c\, d(P,Q)$

Problem # 4: Role of Exceptional Thin Sets

Fix a rational point P. Explain how the local distribution around P changes based on whether P lives in the exceptional thin set or not.

There are some lingering questions regarding local distribution [Hua17] that are of independent interest:

Problem #5

Under what conditions might we deduce Manin-Peyre's Conjecture from the Local Distribution Conjecture of rational points?

Fourier Analysis works for counting 3APs (because it's a linear equation with at least 3 variables). However, higher arithmetic progressions correspond to systems of equations, for which Fourier Analysis works less well. To this end, Gowers developed "Higher Order Fourier Analysis."

Problem #6: Higher-Order Fourier Analysis

Extend the methods of Higher Order Fourier Analysis to the study of rational points. What kind of leverage does this give us? E.g. can we improve upon various results that were (conditionally) proved via the Circle Method?

Remark. The full power of the Green-Tao Arithmetic Regularity Lemma is done not just for Fourier uniformity, but for higher-order equivalents in terms of so-called Gowers norms. There is also a higher-order analogue of stability in this setting [TW21].

References i

Ben Green.
A Szemerédi-type regularity lemma in abelian groups.
Geometric and Functional Analysis, 15(2):340–376, 2005.

Ben Green and Terence Tao.
***An Irregular Mind*, volume 21, chapter An arithmetic regularity lemma, associated counting lemma, and applications.**
Springer, Berlin, 2010.

Zhizhong Huang.
Distribution asymptotique fine des points de hauetuer born'ee sur les variétés algébriques.
PhD thesis, L'université Greoble Alpes, 2017.

References ii

David McKinnon.
Vojta's conjecture implies the Batryev-Manin conjecture for $K3$ surfaces.
Bulletin of the London Mathematical Society, 43(6):1111–1118, 2011.

Maryanthe Malliaris and Saharon Shelah.
Regularity lemmas for stable graphs.
Trans. Amer. Math Soc, 366:1551–1585, 2014.

Emmanuel Peyre.
Hauteurs et mesures de Tamagawa sur les variétés de Fano.
Duke Math. J., 79(1):101–218, 1995.

Sean Prendiville.
Fourier methods in combinatorial number theory.
https://www.dropbox.com/s/kuyaq62srdsc375/Fourier%20methods%20online.pdf?raw=1, 2020.

Yuri Tschinkel.
Fujita's program and rational points.
In Károly Böröczky, János Kollár, and Tamás Szamuely, editors, *Higher Dimensional Varieties and Rational Points*, volume 12 of *Bolyai Society Mathematical Studies*, pages 283–310. Springer Berlin, Heidelberg, 2003.

- Caroline Terry and Julia Wolf.
  Stable arithmetic regularity in the finite field model stable arithmetic regularity in the finite field model.
  Bulletin of the London Mathematical Society, 51(1):70–88, 2019.

- Caroline Terry and Julia Wolf.
  Higher-order generalizations of stability and arithmetic regularity higher-order generalizations of stability and arithmetic regularity.
  arXiv:2111.01739, 2021.

- Yufei Zhao.
  Graph theory and additive combinatorics.
  `https://yufeizhao.com/gtacbook/`.