Introduction to Discrete Analysis

Lecturer : Prof. Tim Gowers

==========
(4th October 2018, Thursday)

Chapter 1. The discrete Fourier transform

Let N be some fixed positive integer. Write ω for $e^{2\pi i/N}$. and $\mathbb{Z}_N$ for $\mathbb{Z}/N\mathbb{Z}$. Let $f : \mathbb{Z}_N \to \mathbb{C}$. Given $r \in \mathbb{Z}_N$ define $\widehat{f}(r)$ to be $\frac{1}{N}\sum_{x\in\mathbb{Z}_N} f(x)\omega^{-rx}$. From now on, we shall use the notation $\mathbb{E}_{x\in\mathbb{Z}_N}$ for $\frac{1}{N}\sum_{x\in\mathbb{Z}_N}$(so that $\widehat{f}(r) = \mathbb{E}_x f(x)e^{-2\pi irx/N}$). If we write ω_r for the function $x \mapsto \omega^{rx}$ and $\langle f, g\rangle$ for $\mathbb{E}_x f(x)\overline{g(x)}$, then $\widehat{f}(r) = \langle f, \omega_r\rangle$.

Let us write $\left\|f\right\|_p$ for $(\mathbb{E}_x|f(x)|^p)^{1/p}$ and call the resulting space $L_p(\mathbb{Z}_N)$.

Important convention : we use averages for the "original functions" in "physical space" and sums for their Fourier transforms in "frequency spaces".

Lemma 1) (Parseval's identity) If $f, g : \mathbb{Z}_N \to \mathbb{C}$, then $\langle\widehat{f},\widehat{g}\rangle = \langle f, g\rangle$. (be aware that inner product are not defined the same in the physical space and the frequency space)

proof)

$$\begin{aligned}\langle\widehat{f},\widehat{g}\rangle &= \sum_r \widehat{f}(r)\overline{\widehat{g}(r)} = \sum_r (\mathbb{E}_x f(x)\omega^{-rx}\overline{(\mathbb{E}_y g(y)\omega^{-ry})})\\ &= \mathbb{E}_x\mathbb{E}_y f(x)\overline{g(y)}\sum_r \omega^{-r(x-y)}\\ &= \mathbb{E}_x\mathbb{E}_y f(x)\overline{g(y)}\Delta_{xy}\\ &= \mathbb{E}_x f(x)\mathbb{E}_y\overline{g(y)}\Delta_{xy} = \mathbb{E}_x f(x)\overline{g(x)} = \langle f, g\rangle\end{aligned}$$

where $\Delta_{xy} = N$ if $x = y$ and 0 if $x \neq y$.(in analogy with δ-distribution).

(End of proof) □

Beware that inner product in the Fourier space and the physical space are defined in a different way : $\langle\widehat{f},\widehat{g}\rangle = \sum_r \widehat{f}(r)\overline{\widehat{g}(r)}$ while $\langle f, g\rangle = \mathbb{E}_x f(x)\overline{g}(x)$

Lemma 2) (Convolution identity) $\widehat{f * g}(r) = \widehat{f}(r)\widehat{g}(r)$

proof) The convolution $f * g(x)$ is defined to be $\mathbb{E}_{y+z=x} f(y)g(z) = \mathbb{E}_y f(y)g(x-y)$. Then,

$$\begin{aligned}\widehat{f * g}(r) &= \mathbb{E}_x f * g(x)\omega^{-rx} = \mathbb{E}_x\mathbb{E}_{y+z=x} f(y)g(z)\omega^{-rx}\\ &= \mathbb{E}_x\mathbb{E}_{y+z=x} f(y)g(z)\omega^{-ry}\omega^{-rz} = \mathbb{E}_y f(y)\omega^{-ry}\mathbb{E}_z g(z)\omega^{-rz} = \widehat{f}(r)\widehat{g}(r)\end{aligned}$$

(End of proof) □

Lemma 3) (Inversion formula) $f(x) = \sum_r \widehat{f}(r)\omega^{rx}$.

proof)

$$\sum_r \widehat{f}(r)\omega^{rx} = \sum_r \mathbb{E}_y f(y)\omega^{r(x-y)} = \mathbb{E}_y f(y) \sum_r \omega^{r(x-y)} = \mathbb{E}_y f(y)\Delta_{xy} = f(x)$$

(End of proof) □

Further Observations :

- If f is real-valued, then $\widehat{f}(-r) = \mathbb{E}_x f(x)\omega^{rx} = \overline{\mathbb{E}_x f(x)\omega^{-rx}} = \overline{\widehat{f}(r)}$.
- If $A \subset \mathbb{Z}_N$, write A(instead of 1_A or χ_A) for the characteristic function of A. Then $\widehat{A}(0) = \mathbb{E}_x A(x) = \frac{|A|}{N}$, the **density** of A. Also, $\|\widehat{A}\|_2^2 = \langle \widehat{A}, \widehat{A}\rangle = \langle A, A\rangle = \mathbb{E}_x A(x)^2 = |A|/N$.
- Let $f : \mathbb{Z}_N \to \mathbb{C}$. Given $\mu \in \mathbb{Z}_N$ with $(\mu, N) = 1$, define $f_\mu(x)$ to be $f(\mu^{-1}x)$. Then $\widehat{f_\mu}(r) = \mathbb{E}_x f_\mu(x)\omega^{-rx} = \mathbb{E}_x f(x/\mu)\omega^{-rx} = \mathbb{E}_x f(x)\omega^{-r\mu x} = \widehat{f}(\mu r)$.

Roth's Theorem

Theorem 4) (Szemerédi's theorem for $k = 3$) For every $\delta > 0$, there exists N such that if $A \subset \{1, \cdots, N\}$ is a set of size at least δN, then A must contain an arithmetic progression of length 3.

Basic strategy : show that if A has density $\geq \delta$ and no 3AP, then there is a long arithmetic progression $P \subset \{1, \cdots, n\}$ such that

$$|A \cap P| \geq (\delta + c(\delta))|P|$$

==========

(9th October, Tuesday)

Theorem 4) (Roth's theorem, or Szemerédi's theorem for $k = 3$) For every $\delta > 0$, there exists N such that if $A \subset \{1, \cdots, N\}$ is a set of size at least δN, then A must contain an arithmetic progression of length 3.

Lemma 5) Let $A, B, C \subset \mathbb{Z}_N$ have densities α, β, γ. If $\max_{r\neq 0} |\widehat{A}(r)| \leq \frac{\alpha(\beta\gamma)^{1/2}}{2}$ and $\frac{\alpha\beta\gamma}{2} > \frac{1}{N}$ then there exist $x, d \in \mathbb{Z}_N$ with $d \neq 0$ s.t. $(x, x+d, x+2d) \in A \times B \times C$.

proof) Assume N is odd.

Consider the function $A(x)B(x+d)C(x+2d)$, which is non-zero if and only if there is a arithmetic progression of length 3. So $\mathbb{E}_{x,d}A(x)B(x+d)C(x+2d)$ counts the density of x, d satisfying such condition.

$$\begin{aligned}
&\mathbb{E}_{x,d}A(x)B(x+d)C(x+2d) = \mathbb{E}_{x+z=2y}A(x)B(y)C(z)\\
=&\mathbb{E}_u\big(\mathbb{E}_{x+z=u}A(x)C(z)\big)\mathbb{E}_{2y=u}B(y) = \mathbb{E}_u A * C(u)B_2(u) = \langle A * C, B_2\rangle\\
=&\langle \widehat{A * C}, \widehat{B}_2\rangle = \langle \widehat{A}\widehat{C}, \widehat{B}_2\rangle = \sum_r \widehat{A}(r)\widehat{C}(r)\widehat{B}(-2r)\\
=&\alpha\beta\gamma + \sum_{r\neq 0} \widehat{A}(r)\widehat{C}(r)\widehat{B}(-2r)
\end{aligned}$$

(Recall, $B_\mu(x)$ is defined to be $B(\mu^{-1}x)$ whenever $(\mu, N) = 1$. We can generalize this to write $B_\mu(x) = \mathbb{E}_{\mu u = x}B(u)$)

The first term $\alpha\beta\gamma$ gives a rough estimate of the value just in terms of the densities and the second term gives an estimate for how random the sets A, B, C behave.

$$\begin{aligned}
\big|\sum_{r\neq 0} \widehat{A}(r)\widehat{B}(-2r)\widehat{C}(r)\big| &\leq \frac{\alpha(\beta\gamma)^{1/2}}{2} \sum_{r\neq 0} \big|\widehat{B}(-2r)\big|\big|\widehat{C}(r)\big|\\
&\leq \frac{\alpha(\beta\gamma)^{1/2}}{2}\big(\sum_r \big|\widehat{B}(-2r)\big|^2\big)^{1/2}\big(\sum_r \big|\widehat{C}(r)\big|^2\big)^{1/2} \quad \text{(Cauchy-Schwarz)}\\
&= \frac{\alpha(\beta\gamma)^{1/2}}{2}\big\|\widehat{B}\big\|_2\big\|\widehat{C}\big\|_2 = \frac{\alpha(\beta\gamma)^{1/2}}{2}\big\|B\big\|_2\big\|C\big\|_2 \quad \text{(Parseval)}\\
&= \frac{\alpha\beta\gamma}{2}
\end{aligned}$$

The contribution to $\mathbb{E}_{x,d}A(x)B(x+d)C(x+2d)$ from $d=0$ is at most $\frac{1}{N}$, so if $\frac{\alpha\beta\gamma}{2} > \frac{1}{N}$, we are done.

(End of proof) □

Now let A be a subset of $\mathbb{Z}_N$ of density $\geq \delta$ and let $B = C = A \cap [\frac{N}{3}, \frac{2N}{3}]$.

- If B has density $< \frac{\delta}{5}$, then either $A \cap [1, \frac{N}{3}]$ or $A \cap [\frac{2N}{3}, N]$ has density at least $\frac{2\delta}{5}$ so in that case we find an AP(arithmetic progression), P(one of $\mathbb{Z}_N \cap [1, N/3]$ and $\mathbb{Z}_N \cap [2N/3, N]$), of length about $N/3$ such that $|A \cap P|/|P| \geq 6\delta/5$.
- Otherwise, if $\max_{r\neq 0} |\widehat{A}(r)| \leq \frac{\delta^2}{10}$. We may pick N large so that $\frac{\delta^3}{50} > \frac{1}{N}$. Then $A \times B \times C$ contains a 3AP by the **Lemma 5** and therefore A contains a 3AP.
- So if A does not contain a 3AP, then either we find P of length about $N/3$ with $|A \cap P|/|P| \geq 6\delta/5$ or we find $r \neq 0$ such that $|\widehat{A}(r)| \geq \delta^2/10$. (This is sometimes called the *dichotomy of order and chaos*)

Definition) If X is a finite set and $f : X \to \mathbb{C}$, $Y \subset X$, write $\mathrm{osc}(f\big|_Y)$ to mean $\max_{y_1,y_2\in Y} |f(y_1) - f(y_2)|$.

Lemma 6) Let $r \in \widehat{\mathbb{Z}}_N$(or just $\mathbb{Z}_N$) and let $\epsilon > 0$. Then there is a partition of $\{1, 2, \cdots, N\}$ into arithmetic progressions P_i of length at least $c(\epsilon)\sqrt{N}$ such that $\mathrm{osc}(\omega_r\big|_{P_i}) \leq \epsilon$ for each i.

proof) Let $t = \lfloor \sqrt{N} \rfloor$. Of the numbers $1, \omega^r, \cdots, \omega^{tr}$ there must be two that differ by at most $2\pi/t$(by sort-of-pigeon hole principle.) If $|\omega^{ar} - \omega^{br}| \leq 2\pi/t$ with $a < b$, then $|1 - \omega^{dr}| \leq 2\pi/t$ where $d = b - a$. Now, by the triangular inequality, if $u < v$ then

$$|\omega^{urd} - \omega^{vrd}| \leq |\omega^{urd} - \omega^{(u+1)rd}| + \cdots + |\omega^{(v-1)rd} - \omega^{vrd}| \leq \frac{2\pi}{t}(v-u)$$

So if P is a progression with common difference d and length l, then $\mathrm{osc}(\omega_r\big|_P) \leq \frac{2\pi l}{t}$. So divide up $\{1, \cdots, N\}$ into residue classes root d and partition each residue class into parts of length between $\frac{\epsilon t}{4\pi}$ and $\frac{\epsilon t}{2\pi}$(this is possible, since $d \leq t \leq \sqrt{N}$.) We are done, with $c(\epsilon) = \frac{\epsilon}{16} (< \frac{\epsilon}{4\pi})$.

(End of proof) □

Now let us use the information that $r \neq 0$ and $|\widehat{A}(r)| \geq \delta^2/10$.
Define the **balanced function** f of A by $f(x) = A(x) - |A|/N$ for each x. Note that $\widehat{f}(0) = 0$ and $\widehat{f}(r) = \widehat{A}(r)$ for all $r \neq 0$.
Now let $P_1, \cdots, P_m$ be given by **Lemma 6** with $\epsilon = \delta^2/20$. Then

$$\begin{aligned}
\frac{\delta^2}{10} \leq |\widehat{f}(r)| &= \frac{1}{N}\Big|\sum_x f(x)\omega^{-rx}\Big| \leq \frac{1}{N}\sum_i \Big|\sum_{x\in P_i} f(x)\omega^{-rx}\Big| \\
&\leq \frac{1}{N}\sum_{i=1}^{m}\Big[\Big|\sum_{x\in P_i} f(x)\omega^{-rx_i}\Big| + \Big|\sum_{x\in P_i} f(x)(\omega^{-rx} - \omega^{-rx_i})\Big|\Big] \quad \text{where } x_i \text{ arbitrary} \in P_i \\
&\leq \frac{1}{N}\sum_{i=1}^{N}\Big|\sum_{x\in P_i} f(x)\Big| + \frac{\delta^2}{20}
\end{aligned}$$

So $\sum_{i=1}^{n}\big|\sum_{x_i\in P_i} f(x)\big| \geq \delta^2 N/20$. Also, $\sum_{i=1}^{m}\sum_{x\in P_i} f(x) = 0$. So

$$\sum_{i=1}^{m}\Big(\Big|\sum_{x\in P_i} f(x)\Big| + \sum_{x\in P_i} f(x)\Big) \geq \frac{\delta^2}{20}\sum_{i=1}^{m}|P_i|$$

Therefore, $\exists i$ s.t. $|\sum_{x\in P_i} f(x)| + \sum_{x\in P_i} f(x) \geq \frac{\delta^2}{20}|P_i|$ by pigeon hole principle and hence $\sum_{x\in P_i} f(x) \geq \frac{\delta^2}{40}|P_i|$. This implies $|A \cap P_i| = \big|\sum_{x\in P_i} f(x) + \frac{|A|}{N}\big| \geq (\delta + \delta^2/40)|P_i|$

==========

(11th October, 2018)

(Example sheets handed out - do not try to use particular theorem, but to apply the idea of using Fourier analysis in combinatorics.

Why does Fourier analysis have anything to do with combinatorics? Because we can formulate various problems in terms of Fourier transformation, e.g. by convolution and Parseval's law.)

•What we have showed : Let $A \subset \mathbb{Z}_N$, $|A| \geq \delta N$. Then,

(1) A contains a 3AP *or*

(2) N is even *or*

(3) $\exists P \subset \{1, \cdots, N\}$, $|P| \geq N/3$ s.t. $|A \cap P| \geq \frac{6\delta}{5}|P|$ *or*

(4) $\exists P \subset \{1, \cdots, N\}$, $|P| \geq \frac{\delta^2}{320}\sqrt{N}$ s.t $|A \cap P| \geq (\delta + \delta^2/40)|P|$

If (2) holds, write $N = N_1 + N_2$ with N_1, N_2 odd, $N_1, N_2 \simeq N/2$. Then A has density at least δ in one of $\{1, \cdots, N_1\}$ or $\{N_1 + 1, \cdots, N_1 + N_2\}$.

If (4) holds (NB (3)⇒ (4)) then we pass to P and start again(that is, replace $\mathbb{Z}_N$ with P and A with $A \cap P$). After $40/\delta$ iterations, the density at least doubles. So the total number of iterations we can have is at most $\leq \frac{40}{\delta} + \frac{40}{2\delta} + \frac{40}{4\delta} + \cdots \leq \frac{80}{\delta}$.

If $\frac{\delta^2}{320}\sqrt{N} \geq N^{1/3}$ at each iteration, and $\delta^3/25 \geq N^{-1}$ (which follow from the first condition) then after $\frac{80}{\delta}$ iterations, we have is $N \geq N^{(1/3)^{80/\delta}}$. So the argument works provided

$$\begin{aligned} & N^{(1/3)^{80/\delta}} \geq \left(\frac{320}{\delta^2}\right)^6 \Leftrightarrow \left(\frac{1}{3}\right)^{80/\delta} \log N \geq 6(\log 320 + 2\log\frac{1}{\delta}) \\ \Leftrightarrow \quad & \frac{-80}{\delta}\log 3 + \log\log N \geq \log 6 + \log(\log 320 + 2\log\frac{1}{\delta}) \\ \Leftarrow \quad & \log\log N \geq \frac{160}{\delta} \\ \Leftarrow \quad & \delta \geq \frac{160}{\log\log N} \end{aligned}$$

(just the follow the idea of calculation„, do not try to learn a particular thing from this calculation)
(In fact 320 above should be corrected to 640 because of (2))

Bogolyubov's Method

Definition) Let $K \subset \widehat{\mathbb{Z}}_N$ and let $\delta > 0$. The **Bohr set** $B(K, \delta)$ has two definitions.

(1) $B(K, \delta) = \{x \in \mathbb{Z}_N : rx \in [-\delta N, \delta N] \quad \forall r \in K\}$ (arc-length definition)

(2) $B(K, \delta) = \{x \in \mathbb{Z}_N : |1 - \omega^{rx}| \leq \delta \quad \forall r \in K|\}$ (chord-length definition)

It turns out that Bohr set has a lot of structure.

Definition) Let G be an Abelian group and let A, B be subsets of G. Then

$$\begin{aligned} A + B &= \{a + b : a \in A, b \in B\} \\ A - B &= \{a - b : a \in A, b \in B\} \\ rA &= \{a_1 + \cdots + a_r : a_1, \cdots, a_r \in A\} \end{aligned}$$

Lemma 7) *(Bogolyubov)* Let $A \subset \mathbb{Z}_N$ be a set of density α. Then $2A - 2A$ contains a Bohr set $B(K, 1/4)$(arc) with $|K| \leq \alpha^{-2}$.

proof) Observe that $x \in 2A - 2A$ if and only if $A * A * (-A) * (-A)(x) \neq 0$(think for a second). But

$$\begin{aligned} A * A * (-A) * (-A)(x) &= \sum_r \widehat{A * A * (-A) * (-A)}(r)\omega^{rx} \quad \text{(inversion formula)} \\ &= \sum_r \left|\widehat{A}(r)\right|^4 \omega^{rx} \end{aligned}$$

Let $K = \{r : |\widehat{A}(r)| \geq \alpha^{3/2}\}$. Then

$$\alpha = \|\widehat{A}\|_2^2 = \sum_r |\widehat{A^2}(r)|^2 \geq \alpha^3 |K|$$

so $|K| \leq \alpha^{-2}$. Now suppose that $x \in B(K, 1/4)$. Then

$$\sum_r |\widehat{A}(r)|^4 \omega^{rx} = \alpha^4 + \sum_{r \in K, r \neq 0} |\widehat{A}(r)|^4 \omega^{rx} + \sum_{r \notin K} |\widehat{A}(r)|^4 \omega^{rx}$$

The real part of the second term is non-negative, since $rx \in [-N/4, N/4]$ when $r \in K$. Also the final term can be bounded by

$$\Big| \sum_{r \notin K} |\widehat{A}(r)|^4 \omega^{rx} \Big| \leq \sum_{r \notin K} |\widehat{A}(r)|^4 < \alpha^3 \sum_{r \notin K} |\widehat{A}(r)|^2 \leq \alpha^4$$

It follows that $\mathrm{Re}\Big(\sum_r |\widehat{A}(r)|^4 \omega^{rx} \Big) > 0$, so $x \in 2A - 2A$.

(End of proof) □

Lemma 8) Let $K \subset \mathbb{Z}_N$ and let $\delta > 0$. Then

(i) $B(K, \delta)$ has density at least $\delta^{|K|}$.

(ii) $B(K, \delta)$ contains a mod-N arithmetic progression of length $\geq \delta N^{1/|K|}$.

proof)

(i) Let $K = \{r_1, \cdots, r_k\}$. Consider the N number of k-tuples $(r_1 x, \cdots, r_k x) \in \mathbb{Z}_N^k$(for $x \in \mathbb{Z}_N$). If we intersect this set of $k-tuples$ with a random "box" $Q = [t_1, t_1 + \delta N] \times \cdots \times [t_k, t_k + \delta N]$. Then the expected number of k-tuples in the box is $\delta^k N$(since each intersection has probability δ^k.)

But if $(r_1 x, \cdots, r_k x)$ and $(r_1 y, \cdots, r_k y)$ belong to this box, then $x - y \in B(K, \delta)$, so the expected number of in $(r_1 y, \cdots, r_k y)$ is smaller or equal to the expected number of y in $B(K, \delta)$. i.e. if we let q be a random point in $Q \cap \mathbb{Z}_N^k$, then

$$\mathbb{E}(\sum_y 1_{y \in B(K,\delta)}) \geq \mathbb{E}(\sum_y 1_{q \in Q} 1_{q+y \in Q}) = \sum_y \mathbb{P}(y \in Q) = N|\delta|^K$$

(there must be an easy argument... in fact, in the lecture, it was concluded before so the expected number of, so there must be more intuitive way of understanding this)

(ii) If we take $\eta > N^{-1/k}$, then by (i) we get that $|B(K, \eta)| > 1$ so $\exists x \in B(K, \eta)$ s.t. $x \neq 0$. But then $dx \in B(K, d\eta)$ for every d. So if $d\eta \leq \delta$ then $B(K, d\eta) \subset B(K, \delta)$ so $dx \in B(K, \delta)$, that gives us an AP of length at least δ/η. So we get one of length at least $\delta N^{1/k}$.

(End of proof) □

==========

(16th October, Tuesday)

(about the example sheet - Question 2, has to fix $\alpha^2\beta^2 N^4$ to $\alpha^2\beta^2 N^3$: a natural way to think about this is "normalizing". We decide one element from other three, so the factor of N should be 3.)

Recall : $B(K, \delta) = \{x \in \mathbb{Z}_N : rx \in [-\delta N, \delta N] \text{ for all } r \in K\}$

Definition) Let A, B be subsets of Abelian groups and let $\phi : A \to B$. Then ϕ is a **Freiman homomorphism of order** k if

$$a_1 + \cdots + a_k = a_{k+1} + \cdots + a_{2k} \quad \Rightarrow \quad \phi(a_1) + \cdots \phi(a_k) = \phi(a_{k+1}) + \cdots + \phi(a_{2k})$$

If $k = 2$, we call this a **Freiman homomorphism**. In that case, the condition is equivalent to

$$a - b = c - d \quad \Rightarrow \quad \phi(a) - \phi(b) = \phi(c) - \phi(d)$$

If ϕ has an inverse which is also a Freiman homomorphism of order k then ϕ is a **Freiman isomorphism of order** k.

Consider $\begin{matrix}\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot\end{matrix}$ and $\cdots, \cdots, \cdots$. They have same additive structure, in some sense, and the Freiman isomorphism describes this. Freiman isomorphism is intended to describe a hidden lattice structure of a additive system.

Lemma 9) Assume $0 \notin K$ and N prime. If $\delta < 1/4$, then $B(K,\delta)$ (arc) is Freiman isomorphic to the intersection $[-\delta N, \delta N]^{|K|} \cap \Lambda \subset \mathbb{R}^{|K|}$ where Λ is some lattice in $\mathbb{R}^{|K|}$.

proof) Let $K = \{r_1, \cdots, r_k\}$ and let $\Lambda = N\mathbb{Z}^k + \{(r_1 x, \cdots, r_k x) : x \in \mathbb{Z}\}$. Write $\underline{r}$ for $(r_1, \cdots, r_k)$. *Claim* that $B(K,\delta) \cong \Lambda \cap [-\delta N, \delta N]^k$.

Define a map $\phi : B(K,\delta) \to \Lambda \cap [-\delta N, \delta N]^k$ by

$$x \mapsto (\langle r_1 x\rangle, \cdots, \langle r_k x\rangle)$$

where $\langle u\rangle$ means the least-modulus residue of u mod N. If $x + y = z + w$, then $\underline{r}x + \underline{x}y = \underline{r}z + \underline{r}w$ in $\mathbb{Z}_N^k$. But for each i, $\langle r_i x\rangle + \langle r_i y\rangle - \langle r_i z\rangle - \langle r_i w\rangle \in [-4\delta N, 4\delta N]$. Since $\delta < 1/4$, that implies that $\langle r_i x\rangle + \langle r_i y\rangle - \langle r_z\rangle - \langle r_i w\rangle = 0$. So $\langle \underline{r}x\rangle + \langle \underline{r}y\rangle - \langle \underline{r}\rangle - \langle \underline{r}w\rangle = \underline{0}$.

That already implies that ϕ is an injection.

If $\underline{r}x + \underline{a}N \in [-\delta N, \delta N]^k$(this is a way of writing a typical point in a lattice), then $r_i x \in [-\delta N, \delta N]$ mod N for each i, so $x \in B(K,\delta)$ and $\phi(x) = \underline{r}x + \underline{a}N$. So ϕ is a surjection.

If $\underline{r}x + \underline{a}N + \underline{r}y + \underline{b}N = \underline{r}z + \underline{c}N + \underline{r}w + \underline{d}N$, then $r_i(x+y) = r_i(z+w)$ mod N, s $x + y = z + w$ mod N, so the inverse of ϕ is also a Freiman homomorphism.

(End of proof) $\square$

(we will not prove in this course, but later in next term, we will prove a generalized version of AP contained in an intersection of K and a box - this requires more knowledge in geometry of numbers, so cannot be dealt here)

Lemma 10) Let Λ be a lattice and let C be a symmetric convex body, both in $\mathbb{R}^k$. Then $|\Lambda \cap C| \le 5^k |\Lambda \cap \frac{C}{2}|$ (5 is not an optimal bound, but convenient to prove - how do we reduce this?)

proof) Let $x_1, \cdots, x_n$ be a maximal subset of $\Lambda \cap C$ such that for all $i \neq j$, $x_j \notin x_i + \frac{C}{2}$. Then by maximality, the sets $x_i + C/2$ cover all of $\Lambda \cap C$. So we see that

$$|A \cap C| = \big|\cup_i (x_i + C/2) \cap A\big| \le n\big|A \cap C/2\big|$$

Also, the sets $x_i + C/4$ are disjoint subsets of $\mathbb{R}^k$, and they are all contained in $C + \frac{C}{4} = 5C/4$. That is $\cup_i (x_i + C/4) \subset 5C/4$ and therefore

$$\big|\cup_i (x_i + C/4)\big| = n\big|C/4\big| \le \big|5C/4\big|$$

Hence $n \le \frac{\mathrm{vol}(5C/4)}{\mathrm{vol}(C/4)} = 5^k$, and we have the conclusion along with the previous inequality.

Corollary 11) If N is primes, $0 \notin K$, $|K| = k$, $\delta < 1/4$, then $|B(K,\delta)| \le 5^k |B(k, \delta/2)|$

2. Sum Sets and their Structure

Suppose A is a subset of integers, and has k such that $|A + A| \le k|A|$. What can we say about the set A? What can we say about the bound of $|rA - sA|$? (will prove $|rA - sA| \le k^{r+s}|A|$)

Lemma 1) (George Petridis - a former student of T.W.Gowers) Let A_0 and B be finite subsets of an Abelian group such that $|A_0 + B| \le K_0|A_0|$. Then there exists a subset $A \subset A_0$ and $K \le K_0$ such that $|A + B + C| \le K|A + C|$ for every finite subset C of the group.

-the style of proof is so different from other combinatorics proofs. It uses induction on C.(Lesson to take : do not dispose a method for proof even before thinking about it)

proof) Choose $A \subset A_0$ that minimizes the ratio $|A+B|/|A|$ and let the minimal ratio be K.

$\star$ Claim : This particular choice of A and K works - We prove this by induction on C.

proof) If $C = \phi$, we are all happy.

Now assume it for C and let $x \notin C$. Then $A + (C \cup \{x\}) = (A+C) \cup \Big[(A+x)\backslash(A'+x)\Big]$ where $A' = \{a \in A : a + x \in A + C\}$. This is a disjoint union, so

$$|A + (C \cup \{x\})| = |A + C| + |A| - |A'|$$

Also, $A + B + (C \cup \{x\}) = (A + B + C) \cup \big((A + B + x)\backslash(A' + B + x)\big)$ since if $a + x \in A + C$, then $a + B + x \subset A + B + C$.

$$\begin{aligned} |A + B + (C \cup \{x\})| &\leq |A + B + C| + |A + B| - |A' + B| \\ &\leq K|A + C| + K|A| - K|A'| \\ &= |A + (C \cup \{x\})| \end{aligned}$$

by induction and minimality property of A

(End of proof) □

==========

(18th October, Thursday)

(Example class next Wednesday afternoon, 24th October)

Corollary 2) If A, B are finite subsets of an Abelian group and $|A + B| \leq K|A|$ then there exists $A' \subset A$, $A' \neq \phi$ such that $|A' + rB| \leq K^r|A'|$ for every positive integer r.

proof) Choose A' as we chose A in **Lemma 1**, that is, minimizing $|A' + B|/|A'|$. Then $|A' + rB| = |A' + B + (r-1)B| \leq K|A' + (r-1)B|$ and $|K' + B| \leq K|A'|$. So we are done by induction.

(End of proof) □

Corollary 3) If $|A + A| \leq K|A|$ or $|A - A| \leq K|A|$, then $|rA| \leq K^r|A|$.

proof) Just set $B = A$ or $-A$ in **Corollary 2**, and note that $|A' \pm rA| \geq |rA|$.

(End of proof) □

Lemma 4) *(Ruzsa triangle inequality)* Let A, B, C be finite subsets of an Abelian group. Then

$$A||B - C| \leq |A - B||A - C|$$

proof) Define a map $\phi : A \times (B - C) \to (A - B) \times (A - C)$ as follows : For each $x \in B - C$, choose functions $b(x) \in B$ and $c(x) \in C$ such that $b(x) - c(x) = x$. Now given (a, x) with $a \in A$, $x \in B - C$, set $\phi(a, x) = (a - b(x), a - c(x))$.

Note that given $(a - b(x), a - c(x)) \in (A - B) \times (A - C)$, we may find $x = b(x) - c(x) = (a - c(x)) - (a - b(x))$. And then, having worked with x, we know $b(x)$, and $a = a - b(x) + b(x)$, so a is determined too, so ϕ is a injection, and the inequality follows.

(End of proof) □

Why does this lemma deserves its name 'triangular inequality?' We can rewrite it as

$$\frac{|B - C|}{|B|^{1/2}|C|^{1/2}} \leq \frac{|A - B|}{|A|^{1/2}|B|^{1/2}} \cdot \frac{|A - C|}{|A|^{1/2}|C|^{1/2}}$$

So if we define the **Rusza distance** $d(A, B)$ to be $\frac{|A-B|}{|A|^{1/2}|B|^{1/2}}$, then the inequality says $d(B, C) \leq d(A, B)d(A, C)$. (be aware that, $\log(d(A, B))$ still does not define a metric... this defines a metric if and only if A is a coset of a subgroup. But this inequality is still useful, since this helps to measure a distance between two sets using other two relations.)

Corollary 5) If $|A + B| \leq K|A|$, then $|rB - sB| \leq K^{r+s}|A|$ for all positive integers r, s.

proof) Pick A' as before. Then by **Corollary 2** with same A and B, has

$$|A' + rB| \leq K^r|A'| \quad \text{and} \quad |-A' - sB| = |A' + sB| \leq K^s|A'|$$

Therefore by Ruzsa triangle inequality,

$$|A'||rB - sB| \leq K^{r+s}|A'|^2 \quad \Rightarrow \quad |rB - sB| \leq K^{r+s}|A|$$

(End of proof) □

Corollary 6) *(Plünnecke's theorem)* If $|A + A| \leq K|A|$ or $|A - A| \leq K|A|$, then $|rA - sA| \leq K^{r+s}|A|$.

proof) Apply **Corollary 5** with $B = -A$ or $B = A$.

(End of proof) □

If you follow the lines of proof, you'll find them trivial but the proof themselves are highly non-trivial.

Lemma 7) *(Ruzsa's embedding Lemma)* Let $A \subset \mathbb{Z}$ be finite and suppose that $|kA - kA| \leq C|A|$. Then there exists a prime $p > 4C|A|$ and a subset $A' \subset A$ of size at least $|A|/k$ such that A' is Freiman isomorphic of order k to a subset of $\mathbb{Z}_p$.
(In fact, the proof tells us that such subset $A' \subset A$ exists for any prime $p > 2C|A|$, with $|A'| \geq \lfloor |A|/k \rfloor$)

This gives a pretty dense subgroup of a cyclic group. This is beneficial later one

proof) Consider the following composition of maps, with prime $q > 2k \max\{|a| : a \in A\}$

$$\phi : \mathbb{Z} \xrightarrow{\text{reduce mod } q} \mathbb{Z}_q \xrightarrow{\times \text{ by some(random) } r \neq 0} \mathbb{Z}_q \xrightarrow{\text{least non-neg. residue}} \mathbb{Z} \xrightarrow{\text{reduce mod } p} \mathbb{Z}_p$$

where q is a prime bigger than $\text{diam}(A)$(the largest element of $A-A$) and p is a prime $\in (2C|A|, 4C|A|]$. Call ϕ for this composition. The first, second and fourth parts are group homomorphisms, and thus Freiman homomorphism of all orders. Also, the third map is a Freiman homomoprhism of order k if you restrict to a subinterval of $[0, q-1]$ of length $\leq q/k$. To see this, write $\langle u \rangle$ for least non-negative residue, let I be an interval of length $\leq q/k$ (and therefore $< q/k$) and $u_1, \cdots, u_{2k} \in I$. If $u_1 + \cdots + u_k - u_{k+1} - \cdots - u_{2k} = 0$, then

$$\langle u_1 \rangle + \cdots + \langle u_k \rangle - \langle u_{k+1} \rangle - \cdots - \langle u_{2k} \rangle \equiv 0 \mod q$$

and also $\langle u_1 \rangle + \cdots + \langle u_k \rangle - \langle u_{k+1} \rangle - \cdots - \langle u_{2k} \rangle \in (-q, q)$. So it is zero.

By the pigeon hole principle, for any r we can find I of length $\leq q/k$ such that $A' = \{a \in A : ra \in I \bmod q\}$ has size at least $|A|/k$ (In fact, this is not true : Let $q = ku + s$ and $|A| = kw + v$, where $s, v \in [0, k-1]$. $I_1, \cdots, I_k$ be disjoint intervals of each length $u \leq q/k$. Then $\left| \cup_j \{a \in A : ra \in I_j\} \right| \geq |A| - s = kw + (v - s)$. Then by pigeon hole principle, one can choose I_j with $\left| \{a \in A : ra \in I_j\} \right| \geq w = \lfloor |A|/k \rfloor$. So $|A|/k$ should be replaced by $\lfloor |A|/k \rfloor$. How do I fix this problem?) Then $\phi|_{A'}$ is a Freiman homomorphism of order k.

It remains to prove that ϕ is an isomorphism to its image. That is, we must show that if

$$a_1 + \cdots + a_k - a_{k+1} - \cdots - a_{2k} \neq 0 \ (a_i \in A)$$
$$\text{then} \quad \langle a_1 \rangle + \cdots + \langle a_k \rangle - \langle a_{k+1} \rangle - \cdots - \langle a_{2k} \rangle \not\equiv 0 \mod p$$

But if the a_i are chosen such that the ra_i all belong to the same interval of length $\leq q/k$ then

$$\left| \langle ra_1 \rangle + \cdots + \langle ra_k \rangle - \langle ra_{k+1} \rangle - \cdots - \langle ra_{2k} \rangle \right| < q$$

and

$$\langle ra_1 \rangle + \cdots + \langle ra_k \rangle - \langle ra_{k+1} \rangle - \cdots - \langle ra_{2k} \rangle \equiv r(a_1 + \cdots + a_k - a_{k+1} - \cdots - a_{2k}) \mod q$$

So all that can go wrong is if $r(a_1 + \cdots + a_k - a_{k+1} - \cdots - a_{2k})$ is xp for some $x \neq 0$ with $|x| < q/p$(if $|x| \geq q/p$, then its contradicts the first inequality). The number of values to avoid is at most $2q/p = |(-p/q, p/q)|$, so for each $a_1 + \cdots + a_k - a_{k+1} - \cdots - a_{2k}$ the probability of going wrong if $r \pmod q$ is chosen randomly is at most $2/p$. And since $|kA - kA| \leq C|A|$, the probability of going wrong is at most $\frac{2}{p}C|A|$, if a_j are chosen at random. Since $p > 2C|A|$, there exists some r s.t. we get an Freiman isompophism of order k.

(End of proof) □

==========

(23rd October, Wednesday)

(Tomorrow example class MR3, 2pm)
(Question 10 in examples sheets - need condition that B has property P)

Freiman's theorem (a version of)

We shall now start with a set $A \subset \mathbb{Z}$ with $|A+A| \leq C|A|$ and put together several of the previous results to say a lot about the structure of A.
(The theorem proved here is not the most general version. For a stronger version, see the attached note of Brad Hannigan-Daley)

- By Plünnecke's theorem, $|8A - 8A| \leq C^{16}|A|$
- By Ruzsa's embedding lemma, A has a subset A' of size at least $|A|/8$ that is 8-isomorphic to a subset $A'' \subset \mathbb{Z}_p$ with $p \leq 4C^{16}|A|$. The density of A'' in $\mathbb{Z}_p$ is $\alpha \geq \frac{1}{32C^{16}}$.
- By Bogolyubov's lemma, $2A'' - 2A''$ contains a Bohr set $B(K, 1/4)$ with $|K| \leq \alpha^{-2}$, which is 2-isomorphic to a set B that is the intersection of a symmetric convex with a lattice of dimension at most α^{-2}. But $2A'' - 2A'' \stackrel{2}{=} 2A' - 2A'$ (think for a minute - if $A' \stackrel{8}{=} A''$ then we should have $2A' - 2A' \stackrel{2}{=} 2A'' - 2A''$), and therefore $2A' - 2A'$ has a subset B that is isomorphic to B'.
- Now let $X \subset A$ be maximal such that the sets $x + B$ with $x \in X$ are disjoint. Then $A \subset X + B - B$, by maximality. Also,

  $$|X||B| = |X + B| \leq |3A - 2A| \leq C^5|A| \quad \Rightarrow |X| \leq \quad C^5|A|/|B|$$

  But by **Lemma 8** about Bohr sets, $|B| \geq 4^{\alpha^{-2}}|A|$, so $|X| \leq 4^{\alpha^{-2}}C^5 \leq 4^{1024C^{32}}C^5$. So A is contained in the union of at most $4^{1024C^{32}}C^5$ translates of $B - B$
- If $B = \Lambda \cap K_0$ (where K_0 is a symmetric convex body) then $B - B \subset \Lambda \cap 2K_0$, and also $|B - B| \leq C^2|A| \leq 5^{\alpha^{-2}}|B|$. So this

We need a additional statement about the geometry of numbers to prove the real Freiman's theorem.

****

(not lectured)

Minkowski's Second Theorem : Let K be a symmetric convex body and Λ a lattice in $\mathbb{R}^k$. Let $\lambda_1 \leq \cdots \leq \lambda_k$ be the successive minima of C with respect to Λ. Then $\lambda_1 \cdots \lambda_k \leq 2^k \det(\Lambda)/\mathrm{vol}(K)$.

Lemma) (Volume Packing Lemma) Let Λ, Λ' be lattices with $\lambda \subset \Lambda' \subset \mathbb{R}^n$. Then

$$\det(\Lambda) = \det(\Lambda')[\Lambda : \Lambda]$$

Proposition) Let N be a large prime, $r_1, \cdots, r_k$ be distinct residue classes mod N with $k \geq 2$, and let $\delta \in (0, 1)$. Then the Bohr set $\mathcal{B}(r_1, \cdots, r_k; \delta)$ in $\mathbb{Z}/N$ contains a proper generalized arithmetic progression of dimension k and size at least $(\delta/k)^k N$.

It follows the Freiman's theorem.

Freiman's Theorem : Let A be a finite subset of $\mathbb{Z}$ with $|A + A| \leq C|A|$. Then there exist constants d' and S depending only on C such that A is contained in a generalized arithmetic progression of dimension at most d' and size at most S.

****

The Balog-Szemerédi-Gowers theorem

Definition) Let A be a subset of an Abelian group. An **additive quadruple** in A is a quadruple $(a,b,c,d) \in A^4$ such that $a+b=c+d$. (Equivalently, it's a quadruple such that $a-b=c-d$)

If $|A|=n$, then the number of additive quadruples in A is at most n^3 (this is obtained if A is a subgroup or a coset of a subgroup.) We may expect that as the structure of A gets closer to the structure of a group, then larger the number of quadruples there would be. In fact, the converse statement - *if* A^4 contains at least cn^3 additive quadruples, *then* A has a subset A' of size at least $c'n$ with $|A-A'| \leq C|A'|$ where c' and C depend (nicely) on c only - also holds.(We are going to prove the last statement.) Hence, the size of number of quadruples can be used as a measure of how close the structure of the subset A is close to structure of a group.(though this statement cannot be justified quantitatively.)

==========

(25th October, Thursday)

Lemma 8) Let $A_1, A_2, \cdots, A_m$ be subsets of $[n]=\{1,2,\cdots,n\}$ of average density at least δ. Then for any $\eta > 0$ we can find a set $B \subset [m]$ of size at least $\delta m/\sqrt{2}$ such that the proportion of pairs $(i,j) \in B^2$ with $|A_i \cap A_j| \geq \eta\delta^2 n/2$ is at least $1-\eta$.

proof) Choose $y \in [n]$ uniformly at random, and set $B=\{i : y \in A_i\}$. The probability that we pick both i and j is $|A_i \cap A_j|/n$, which we can write as $\langle A_i, A_j \rangle$. So the expected number of pairs we pick is

$$\mathbb{E}|B|^2 = \sum_{i,j} \langle A_i, A_j \rangle = \Big\| \sum_i A_i \Big\|_2^2 \geq \Big\| \sum_i A_i \Big\|_1^2 = (\delta m)^2 = \delta^2 m^2$$

The probability that we pick i,j with $|A_i \cap A_j| \leq \eta\delta^2 n/2$ is at most $\eta\delta^2/2$. Call such a pair (i,j) **bad**. Then

$$\mathbb{E}|B|^2 - \eta^{-1}\mathbb{E}(\#\text{ bad pairs picked}) \geq \delta^2 m^2 - \eta^{-1}\frac{\eta\delta^2}{2}m^2 = \frac{\delta^2 m^2}{2}$$

Therefore, there exists B such that $|B| \geq \delta m/\sqrt{2}$ and the proportion of pairs in B^2 that are bad is at most η.

(End of proof) □

The following is an equivalent statement.

Corollary 9) Let G be a bipartite graph with finite vertex sets X, Y and density at least δ. Then there is a subset $B \subset X$ of size at least $\delta|X|/\sqrt{2}$ such that for at least $(1-\eta)|B|^2$ pairs $(x_1,x_2) \in B^2$, there are at least $\frac{\eta\delta^2}{2}|Y|$ paths of length 2 from x_1 to x_2.
(Note : density of a bipartite graph $= |E(G)/(|X||Y|)|$)

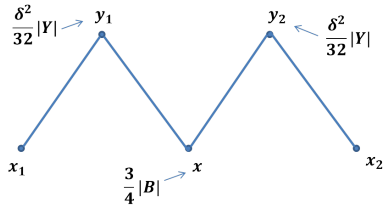
proof) Put $X=[m]$ and $Y=[n]$ and $A_i = \{y \in Y : (i,y) \in E, i \in X, y \in Y\} \subset Y$ be as in the setting of **Lemma 8**. If we choose a subset B of X, then there is a path of length 2 from $x_1 \in B$ to $x_2 \in B$ passing through B if and only if $|A_{x_1} \cap A_{x_2}|$ is non-empty, therefore we can just apply **Lemma 8** to conclude.

(End of proof) □

Corollary 10) Let G be a bipartite graph as above with density δ. Then there is a subset $B' \subset X$ of size at least $\delta|X|/2\sqrt{2}$ such that for every $x_1, x_2 \in B'$ there are at least $\frac{\delta^5}{2048\sqrt{2}}|X||Y|^2$ paths of length 4 from x_1 to x_2.

proof) Choose B as in **Corollary 9** with $\eta = 1/16$ (with size at least $\delta|X|/\sqrt{2}$). Define a graph Γ with vertex set B, joining x_1 to x_2 if there are at least $\frac{\delta^2}{32}|Y|$ paths of length 2 from x_1 to x_2 in G. By **Corollary 9**, the average degree in Γ is at least $(1-\frac{1}{16})|B| = \frac{15}{16}|B|$. Therefore, (by pigeon hole principle - if there are $\leq |B|/2$ vertices with degree $|B|$ and $\geq |B|/2$ vertices with degree $7|B|/8$, then the average degree would be $\leq 15|B|/16$) there are at least $|B|/2$ vertices in B of degree at least $\frac{7}{8}|B|$. Let B' be the set of all such vertices. If $x_1, x_2 \in B'$, then there are at least $1-(1-\frac{7}{8})-(1-\frac{7}{8}) = \frac{3}{4}|B|$ vertices in B joined to both x_1 and x_2 in Γ. Therefore, there at least

$$\Big(\frac{\delta^2|Y|}{32}\Big) \times \Big(\frac{\delta^2|Y|}{32}\Big) \times \Big(\frac{3}{4}|B|\Big) \geq \frac{\delta^4}{1024}|Y|^2\frac{3}{4}\frac{\delta|X|}{\sqrt{2}} \quad \text{paths of length 4 from } x_1 \text{ to } x_2$$



[Note : setting $\eta = 1/8$ would(?) give a bound of $\frac{\delta^4}{256}|Y|^2\frac{1}{2}\frac{\delta|X|}{\sqrt{2}}$]

(End of proof) □

Lemma 11) *(The BSG Lemma)* Let A be a subset of size n of an Abelian group H. Suppose that there are at least cn^3 additive quadruples in A. Then A has subset A' of size at least $c'n$ with $|A' - A'| \leq C|A'|$, where c' and C depend on c only.

proof) For each d, let $f(d)$ be the number of ways of writing $d = a_1 - a_2$ with $a_1, a_2 \in A$. Then $\sum_d f(d)^2 \geq cn^3$, because $f(d)^2$ is the number of quadruples (a_1, a_2, a_3, a_4) with $d = a_1 - a_2 = a_3 - a_4$ for each d. Call d **popular** if $f(d) \geq cn/2$. Then

$$cn^3 \leq \sum_d f(d)^2 = \sum_{d \text{ popular}} f(d)^2 + \sum_{d \text{ unpopular}} f(d)^2$$
$$\leq (\#\text{ popular d}) \times n^2 + \frac{cn}{2}n^2 \quad (\text{since } \sum_d f(d) = n^2)$$

Therefore, the number of popular d is at least $cn/2$.

Now define a graph with vertex set A, by joining a_1 to a_2 if $a_1 - a_2$ (or $a_2 - a_1$) is popular. Each popular difference contributes at least $cn/2$ edges so the average degree of the graph is at least $c^2/4$.

By duplicating the vertex set, create a corresponding bipartite graph G. By **Corollary 10** with $\delta = c^2/4$, we can find a subset $B \subset A$ of size at least $\frac{c^2}{8\sqrt{2}}|A|$ such that for any $a_1, a_2 \in B$ there are at least $\frac{\delta^5}{2048\sqrt{2}}|A|^3$ paths of length 4 from a_1 to a_2. Each such paths of length 4 gives us at least $(\frac{c}{2}|A|)^4$ ways of writing $a_2 - a_1$ as $b_1 - b_2 + b_3 - b_4 + b_5 - b_6 + b_7 - b_8$ with all $b_i \in A$. So

$$|B - B| \times \frac{\delta^5}{2048\sqrt{2}}|A|^3(\frac{c}{2}|A|)^4 \leq |A|^8$$

so $|B - B| \leq C'|A| \leq C|B|$

(End of proof) □

==========

(30th October, Tuesday)

(When revising this course, try to remember the structure of the proofs rather than the exact numbers. The exact number is not very important)

3. Quasirandom Graphs

The box norm is an extremely useful tool in studying Quasirandomness.

The box norm

Let X and Y be finites sets and $f : X \times Y \to \mathbb{C}$. We define the **box norm** $\|f\|_\square$ of f by the formula

$$\|f\|_\square^4 = \mathbb{E}_{x_1,y_1,x_2,y_2} f(x_1, y_1)\overline{f(x_1, y_2)f(x_2, y_1)}f(x_2, y_2)$$

This is a non-negative real number, because we can rewrite this as

$$\|f\|_\square^4 = \mathbb{E}_{x_1,x_2}\left(\left|\mathbb{E}_y f(x_1, y)\overline{f(x_2, y)}\right|^2\right)$$

It is not entirely clear if whether this really is a norm.

If $f_1, f_2, f_3, f_4 : X \times Y \to \mathbb{C}$, then their **box inner product** $[f_1, f_2, f_3, f_4]$ is

$$[f_1, f_2, f_3, f_4] = \mathbb{E}_{x_1,y_1,x_2,y_2} f_1(x_1, y_1)\overline{f_2(x_1, y_2)f_3(x_2, y_1)}f_4(x_2, y_2)$$

We shall *temporarily* use the notation $[f_1, f_2, f_3, f_4] = \boxed{\begin{matrix} f_1 & f_2 \\ f_3 & f_4 \end{matrix}}$ in order to expose the symmetry.

Lemma 1) For any four functions $f_{00}, f_{01}, f_{10}, f_{11}$ we have

$$\left|\boxed{\begin{matrix} f_{00} & f_{01} \\ f_{10} & f_{11} \end{matrix}}\right| \leq \|f_{00}\|_\square \|f_{01}\|_\square \|f_{10}\|_\square \|f_1\|_\square$$

This is called **Box Cauchy-Schwarz inequality**.

proof)

$$\begin{aligned}
&\mathbb{E}_{x_0,y_0,x_1,y_1} f_{00}(x_0, y_0)\overline{f_{01}(x_0, y_1)f_{10}(x_1, y_0)}f_{11}(x_1, y_1) \\
=&\mathbb{E}_{x_0,x_1}\Big(\mathbb{E}_{y_0} f_{00}(x_0, y_0)\overline{f_{10}(x_1, y_0)}\Big)\overline{\Big(\mathbb{E}_{y_1} f_{01}(x_0, y_1)\overline{f_{11}(x_1, y_1)}\Big)} \\
\leq&\Big(\mathbb{E}_{x_0,x_1}\Big|\mathbb{E}_{y_0} f_{00}(x_0, y_0)\overline{f_{10}(x_1, y_0)}\Big|^2\Big)^{1/2}\Big(\mathbb{E}_{x_0,x_1}\Big|\mathbb{E}_{y_1} f_{01}(x_0, y_1)\overline{f_{11}(x_1, y_1)}\Big|^2\Big)^{1/2} \\
=&\boxed{\begin{matrix} f_{00} & f_{00} \\ f_{10} & f_{10} \end{matrix}}^{1/2}\boxed{\begin{matrix} f_{01} & f_{01} \\ f_{11} & f_{11} \end{matrix}}^{1/2}
\end{aligned}$$

By Symmetry (interchanging the roles of x and y) we also have, for any functions $g_{00}, g_{01}, g_{10}, g_{11}$,

$$\left\|\boxed{\begin{matrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{matrix}}\right\| \leq \boxed{\begin{matrix} g_{00} & g_{01} \\ g_{00} & g_{01} \end{matrix}}^{1/2}\boxed{\begin{matrix} g_{10} & g_{11} \\ g_{10} & g_{11} \end{matrix}}^{1/2}$$

Combining these two inequalities in the obvious way gives the result.

(End of proof) □

(the whole point of using this weird notation was to visualize symmetry and to simplify the argument. Can give up with this notation and write a lengthy proof)

Corollary 2) $\|\cdot\|_\square$ is a norm.

proof) The only property that is not completely straightforward is the triangular inequality. Let f_0, $f_1 : X \times Y \to \mathbb{C}$. Then

$$\begin{aligned}
\|f_0 + f_1\|_\square^4 &= [f_0 + f_1, f_0 + f_1, f_0 + f_1, f_0 + f_1] \\
&= \sum_{\epsilon\in\{0,1\}^4} [f_{\epsilon_1}, f_{\epsilon_2}, f_{\epsilon_3}, f_{\epsilon_4}] \\
&\leq \sum_{\epsilon\in\{0,1\}^4} \|f_{\epsilon_1}\|_\square \|f_{\epsilon_2}\|_\square \|f_{\epsilon_3}\|_\square \|f_{\epsilon_4}\|_\square \\
&= (\|f_0\|_\square + \|f_1\|_\square)^4
\end{aligned}$$

(End of proof) □

Remark : Suppose that $f(x, y) = g(x)$ for every x, y. Then

$$\|f\|_\square^4 = \mathbb{E}_{x_1,x_2} g(x_1)\overline{g(x_1)g(x_2)}g(x_2) = \|g\|_2^4$$

so $\|f\|_\square = \|g\|_2$.

Corollary 3) *(The box-norm inequality)* If $f : X \times Y \to \mathbb{C}$, $u : X \to \mathbb{C}$, $v : Y \to \mathbb{C}$ then

$$|\mathbb{E}_{x,y} f(x, y)u(x)v(y)| \leq \|f\|_\square \|u\|_2 \|v\|_2$$

proof) Apply the box Cauchy-Schwarz inequality to $f_1 = f$, $f_2(x,y) = u(x)$, $f_3(x,y) = v(y)$, $f_4(x,y) = 1$ and use the above remark.

(End of proof) □

Lemma 4) Let $f : X \times Y \to \mathbb{C}$. Then $\|f\|_\square \geq |\mathbb{E}_{x,y} f(x,y)|$.

proof) Note that we can write $\|f\|_\square^4 = \mathbb{E}_{x_1,x_2} \left|\mathbb{E}_y f(x_1,y)\overline{f(x_2,y)}\right|^2$, so

$$\begin{aligned}\|f\|_\square^4 &= \mathbb{E}_{x_1,x_2} \left|\mathbb{E}_y f(x_1,y)\overline{f(x_2,y)}\right|^2 \geq \left|\mathbb{E}_{x_1,x_2}\mathbb{E}_y f(x_1,y)\overline{f(x_2,y)}\right|^2 \\ &= \left|\mathbb{E}_y \left|\mathbb{E}_x f(x,y)\right|^2\right|^2 \geq \left|\mathbb{E}_y \mathbb{E}_x f(x,y)\right|^4\end{aligned}$$

(End of proof) □

Lemma 5) Let $F : X \times Y \to \mathbb{R}$ be such that $\mathbb{E}_{x,y} F(x,y) = \delta > 0$ and $\|F\|_\square^4 \leq \delta^4(1+c)$. Let $f(x,y) = F(x,y) - \delta$. Then $\|f\|_\square \leq |\delta|(c^{1/4} + (c/2)^{1/2})$.

This inequality can be thought to be an estimation of box norm-variance of F.

proof) For each x, let $g(x) = \mathbb{E}_y f(x,y)$ and let $h(x,y) = f(x,y) - g(x)$. Then $\mathbb{E}_x g(x) = \mathbb{E}_{x,y} f(x,y) = \mathbb{E}_{x,y} F(x,y) - \delta = 0$ and for every x, $\mathbb{E}_y h(x,y) = g(x) - g(x) = 0$. Now

$$\begin{aligned}\|F\|_\square^4 &= \mathbb{E}_{x_1,x_2} \left|\mathbb{E}_y F(x_1,y)\overline{F(x_2,y)}\right|^2 \\ &= \mathbb{E}_{x_1,x_2} \left|\mathbb{E}_y (\delta + g(x_1) + h(x_1,y))(\delta + \overline{g(x_2)} + \overline{h(x_2,y)})\right|^2 \\ &= \mathbb{E}_{x_1,x_2} \left|\mathbb{E}_y (\delta + g(x_1))(\delta + \overline{g(x_2)}) + \mathbb{E}_y h(x_1,y)\overline{h(x_2,y)}\right|^2\end{aligned}$$

==========

(1st November, Thursday)

proof continued) For $F(x,y) = \delta + g(x) + h(x,y)$,

$$\|F\|_\square^4 = \mathbb{E}_{x_1,x_2} \left|(\delta + g(x_1))(\delta + \overline{g(x_2)}) + \mathbb{E}_y h(x_1,y)\overline{h(x_2,y)}\right|^2$$

When we expand the mod squared, we get three terms as follows. The first is

$$\begin{aligned}&\mathbb{E}_{x_1,x_2}(\delta + g(x_1))\overline{(\delta + g(x_1))(\delta + g(x_2))}(\delta + g(x_2) \\ =&|\delta|^4 + |\delta|^2\mathbb{E}|g(x_1)|^2 + |\delta|^2\mathbb{E}|g(x_2)|^2 + \mathbb{E}|g(x_1)|^2|g(x_2)|^2 = (|\delta|^2 + \|g\|_2^2)\end{aligned}$$

The second is

$$2\mathrm{Re}(\mathbb{E}_{x_1,x_2}(\delta + g(x_1))\overline{(\delta + g(x_1))}\mathbb{E}_y \overline{h(x_1,y)}h(x_2,y)) = \mathbb{E}_y 2\mathrm{Re}\left|\mathbb{E}_x(\delta + g(x))\overline{h(x,y)}\right|^2 \geq 0$$

The third is $\|h\|_\square^4$.

Putting this together, has

$$\|F\|_\square^4 \geq (\delta^2 + \|g\|_2^2)^2 + \|h\|_\square^4$$

Therefore, if $\|F\|_\square^4 \leq |\delta|^4(1+c)$, we have $\|h\|_\square^4 \leq c|\delta|^4$ and

$$|\delta|^2 + \|g\|_2^2 \leq |\delta|^2(1+c)^{1/2} \leq |\delta|^2(1 + \frac{c}{2})$$

so $\|g\|_2 \leq \sqrt{\frac{c}{2}}\delta$.

Therefore, $\|f\|_\square = \|F - \delta\|_\square \leq |\delta|(c^{1/4} + (c/2)^{1/2})$.

(End of proof) □

We are now ready to do some quasirandom graphs. The followings are **quasirandomness conditions**

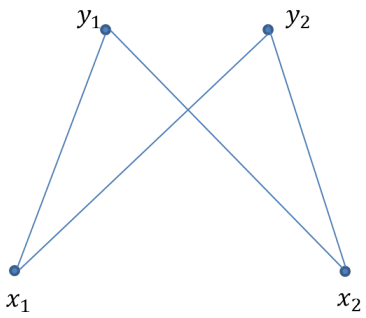
Lemma 6) *(Quasirandomness Conditions)* Let G be a bipartite graph of density δ with finite vertex sets X, Y. Then the following statements are "equivalent", in the sense that if one holds for sufficiently small c_i, then the others hold for small c_j.

(i) $\|G\|_{\square}^4 \leq \delta^4(1+c_1)$. ($\|G\|_{\square}^4$ is the **4-cycle density** of G.)

(ii) For each vertex $x, x_1, x_2 \in X$, let $\delta(x) = \mathbb{E}_y G(x,y)$ =(density of neighbourhood of x in Y), and let $\delta(x_1,x_2) = \mathbb{E}_y G(x_1,y)G(x_2,y)$ =(density of neighbourhood intersection). Similarly define $\delta(y)$, $\delta(y_1,y_2)$ for $y, y_1, y_2 \in Y$. Then

$$\mathbb{E}_{x_1,x_2}\left[\left(\delta(x_1,x_2)-\delta^2\right)^2\right] \leq c_2\delta^4$$

(be aware that this object is not the variance of $\delta(x_1,x_2)$, because δ^2 is not its mean)



(iii) For every $A \subset X$ of density α and $B \subset Y$ of density β,

$$\left|\mathbb{E}_{x,y}G(x,y)A(x)B(y) - \delta\alpha\beta\right| \leq c_3\delta$$

proof)

(i)⇔(ii) Observe first that $\|G\|_{\square}^4 = \mathbb{E}_{x_1,x_2}\delta(x_1,x_2)^2$. Also,

$$\mathbb{E}(\delta(x_1,x_2)-\delta^2)^2 = \mathbb{E}_{x_1,x_2}\delta(x_1,x_2)^2 - 2\delta^2\mathbb{E}_{x_1,x_2}\delta(x_1,x_2) + \delta^4$$
$$\mathbb{E}(\delta(x_1,x_2)) = \mathbb{E}_y\mathbb{E}_{x_1,x_2}G(x_1,y)G(x_2,y) = \mathbb{E}_y\delta(y)^2 \geq (\mathbb{E}_y\delta(y))^2 = \delta^2$$

(⇒) So if $\|G\|_{\square}^4 \leq \delta^4(1+c)$, then $\mathbb{E}(\delta(x_1,x_2)-\delta^2)^2 \leq \delta^4(1+c) - 2\delta^4 + \delta^4 = \delta^4 c$.

(⇐) Conversely, if $\mathbb{E}(\delta(x_1,x_2)-\delta^2)^2 \leq c\delta^4$, then $\mathbb{E}(\delta(x_1,x_2)-\delta^2) \leq c^{1/2}\delta^2$, and hence

$$\mathbb{E}(\delta(x_1,x_2)) \leq (1+c^{1/2})\delta^2$$

Then

$$\|G\|_{\square}^4 = \mathbb{E}_{x_1,x_2}\delta(x_1,x_2)^2 \leq c\delta^4 + 2\delta^2(1+c^{1/2})\delta^2 - \delta^4 = \delta^4(1+c+2c^{1/2})$$

(i)⇒(iii) (Note : this implication is the most useful one) Suppose (i) holds. Then

$$\left|\mathbb{E}_{x,y}G(x,y)A(x)B(y) - \delta\alpha\beta\right| = \left|\mathbb{E}_{x,y}(G-\delta)(x,y)A(x)B(y)\right|$$

But by **Lemma 5**, $\|G-\delta\|_{\square} \leq \delta(c_1^{1/4} + (c_1/2)^{1/2})$ so by the box-norm inequality(**Corollary 3**), this is at most

$$\delta(c_1^{1/4} + (c_1/2)^{1/2})\alpha^{1/2}\beta^{1/2} \quad \left(\leq \delta(c_1^{1/4} + (c_1/2)^{1/2})\right)$$

(iii)⇒(i) *(with assumption of G being regular)* If G is regular (on both sides) (that is, the vertices connected to a particular vertex has regular density δ), then the proof is very short.

If (i) is false, then

$$\mathbb{E}_{x_2,y_2}\mathbb{E}_{x_1,y_1}G(x_1,y_1)G(x_1,y_2)G(x_2,y_1)G(x_2,y_2) > \delta^4(1+c_1)$$

since $\mathbb{P}[G(x_2,y_2)=1]=\delta$, it follows (by pigeon hole principle) that there exist x_2,y_2 such that

$$\mathbb{E}_{x_1,y_1}G(x_1,y_1)G(x_1,y_2)G(x_2,y_1) > \delta^3(1+c_1)$$

Let

$$A=\{x : G(x,y_2)=1\}, \quad B=\{y : G(x_2,y)=1\}$$

Then A and B have density δ (by regularity) and

$$\left|\mathbb{E}_{x,y}G(x,y)A(x)B(y)-\delta^3\right| > \delta^3 c_1$$

==========

(6th November, Tuesday)

proof continued) *(without assumption of G being regular)* If we have

$$\mathbb{E}_{x_1,x_2,y_1,y_2}G(x_1,y_1)G(x_1,y_2)G(x_2,y_1)G(x_2,y_2) \geq \delta^4 + c\delta^4$$

then

either $\mathbb{E}_{x_1,x_2,y_1,y_2}G(x_1,y_1)G(x_1,y_2)G(x_2,y_1)G(x_2,y_2) \geq \delta\mathbb{E}_{x_1,x_2,y_1,y_2}G(x_1,y_2)G(x_2,y_1)G(x_2,y_2) + \frac{c}{2}\delta^4$
or $\mathbb{E}_{x_1,x_2,y_1,y_2}G(x_1,y_2)G(x_2,y_1)G(x_2,y_2) \geq \delta^3 + \frac{c}{2}\delta^3 = \delta\mathbb{E}_{x_1,x_2,y_1,y_2}G(x_1,y_2)G(x_2,y_1) + \frac{c}{2}\delta^3$

In the first case, we can rewrite the inequality as

$$\mathbb{E}_{x_1,x_2,y_1,y_2}\big(G(x_1,y_1)-\delta\big)G(x_1,y_2)G(x_2,y_1)G(x_2,y_2) \geq \frac{c}{2}\delta^4$$

Since $G(x_2,y_2)=1$ with probability δ, it follows that there exist x_2,y_2 such that

$$\mathbb{E}_{x_2,y_2}\big(G(x_1,y_1)-\delta\big)G(x_1,y_2)G(x_2,y_1) \geq \frac{c\delta^3}{2}$$

Then, as before, take $A=N(y_2)=\{x : G(x,y_2)=1\}$, $B=N(x_2)$ and we get

$$\left|\mathbb{E}_{x,y}G(x,y)A(x)B(y)-\delta\frac{|A|}{|X|}\frac{|B|}{|Y|}\right| \geq \frac{c\delta^3}{2}$$

which contradicts the condition (iii).

In the second case, we can rewrite the inequality as

$$\mathbb{E}_{x_1,x_2,y_1,y_2}G(x_1,y_2)G(x_2,y_1)(G(x_2,y_2)-\delta) \geq \frac{c\delta^3}{2}$$

It follows that there exist x_1,y_1 such that

$$\mathbb{E}_{x_2,y_2}G(x_1,y_2)G(x_2,y_1)G(x_2,y_2) - \delta\mathbb{E}_{x_2,y_2}G(x_1,y_2)G(x_2,y_1) \geq \frac{c\delta^3}{2}$$

This time set $A=N(y_1)$, $B=N(x_1)$ then we have contradiction to condition (iii).

(End of proof) □

Lemma 7) *(Counting random)* Let G be a k-partite graph with vertex sets $X_1,\cdots,X_k$. Write $G(X_i,X_j)$ for the induced bipartite subgraph with vertex sets (X_i,X_j). Suppose that the density of $G(X_i,X_j)$ is α_{ij} and that writing G_{ij} for the restriction G_{ij} for the restriction of G to $X_i\times X_j$, $\left\|G_{ij}-\alpha_{ij}\right\|_\square \leq c$. Let H be a graph with vertex set $[k]$ and let $x_i\in X_i$ be chosen independently at random. Then

$$\left|\mathbb{E}_{x_1,\cdots,x_k}\prod_{ij\in E(H)}G_{ij}(x_i,x_j) - \prod_{ij\in E(H)}\alpha_{ij}\right| \leq 2^{|E(H)|}c$$

(can do better than this, but it is easier to prove with bound $2^{|E(H)|}c$.)

proof) Let $G_{ij} = f_{ij} + \alpha_{ij}$. Then $\|f_{ij}\|_\square \leq c$ for all i, j. Then

$$\mathbb{E} \prod_{ij \in E(H)} G_{ij}(x_i, x_j) = \mathbb{E} \prod_{ij \in E(H)} (\alpha_{ij} + f_{ij}(x_i, x_j))$$

The main term is $\prod_{ij \in E(H)} \alpha_{ij}$ (when we expand the product into $2^{|E(H)|}$) terms). Every other term can be written in the form

$$\mathbb{E}_{x_i, x_j} f_{ij}(x_i, x_j) u(x_i) v(x_j)$$

with $\|u\|_\infty, \|v\|_\infty \leq 1$ if we fix all variables other than x_i and x_j. Therefore, by the box-norm inequality, it has size at most c, and this remains true after averaging over the other variables.

(End of proof) $\square$

Szemerédi's Regularity Lemma

Szemerédi's regularity lemma says that almost every k-partite graph is actually very close to a quasirandom graph.

Let X be a finite set and let $\mathcal{P} = \{X_1, \cdots, X_k\}$ be a partition $X = X_1 \cup \cdots \cup X_k$. Given a function $f : X \to \mathbb{R}$, define the conditional expectation of f with respect to $\mathcal{P}$ by

$$\mathbb{E}[f|\mathcal{P}] = \mathbb{E}[f|\mathcal{P}](x) = \mathbb{E}[f(y)|y \text{ is in the same cell as } x]$$

Let us temporarily write Pf for $\mathbb{E}[f|\mathcal{P}]$ (P for projection).

Remarks :

(i) Note that $\mathbb{E}[\mathbb{E}[f|\mathcal{P}]] = \mathbb{E}f$.

(ii) $\langle f, Pf \rangle = \mathbb{E}_x[f(x)\mathbb{E}[f|\mathcal{P}](x)] = \mathbb{E}[\mathbb{E}[f|\mathcal{P}]^2] = \langle Pf, Pf \rangle$. Therefore, $\langle (I-P)f, Pf \rangle = 0$.

(iii) $P(Pf) = \mathbb{E}[\mathbb{E}[f|\mathcal{P}]|\mathcal{P}] = \mathbb{E}[f|\mathcal{P}] = Pf$. Together with (ii), P is an orthogonal projection.

(iv) It follows that

$$\mathbb{E}f^2 = \|f\|_2^2 = \|Pf\|_2^2 + \|(I-P)f\|_2^2 \geq \|Pf\|_2^2 = \mathbb{E}[\mathbb{E}[f|\mathcal{P}]^2]$$

So the norm decreases in taking conditional expectation.

(v) Let $\mathcal{Q}$ be a refinement of $\mathcal{P}$. Then

$$\mathbb{E}[\mathbb{E}[f|\mathcal{Q}]|\mathcal{P}] = \mathbb{E}[f|\mathcal{P}]$$

It follows from (iv) applied to $\mathbb{E}[f|\mathcal{Q}]$ that

$$\mathbb{E}[\mathbb{E}[f|\mathcal{Q}]^2] \geq \mathbb{E}[\mathbb{E}[f|\mathcal{P}]^2]$$

==========

(8th November, Thursday)

(Second Example Class at 16th November, Friday, 2 pm)

- Let G be a bipartite graph with (finite) vertex sets X, Y. Let $A \subset X$, $B \subset Y$. Then the **density** $d(A, B) = d_G(A, B)$ is

$$\mathbb{E}_{x \in A, y \in B} G(x, y)$$

- We shall say that (A, B) is **ϵ-regular** if $\forall A' \subset A$, $B' \subset B$, we have

$$\left| \mathbb{E}_{x \in A, y \in B} G(x, y) A'(x) B'(y) - d(A, B) \mathbb{E}_{x \in A, y \in B} A'(x) B'(y) \right| \leq \epsilon$$

Note, this is one of the quasirandomness conditions on the subgraph of G induced by A' and B'.

Theorem 8) *(Szemerédi's regularity lemma)* Let G be a bipartite graph with vertex sets X, Y and let $\epsilon > 0$. Then there exist partitions $X = X_1 \cup \cdots \cup X_r$, $Y = Y_1 \cup \cdots \cup Y_s$ with $r, s \leq k(\epsilon)$ such that if you choose $(x, y) \in X \times Y$ at random, then the probability that it belongs to an ϵ-regular pair (X_i, Y_j) is at least $1 - \epsilon$.

Can actually make stronger statements - such as making the partition sets of almost the same size.

Definition) Given partitions $X_1 \cup \cdots \cup X_r$ of X and $Y_1 \cup \cdots \cup Y_s$ of Y, let $\mathcal{P}$ be the partition of $X \times Y$ into the sets $X_i \times Y_j$. The **mean-square density of** G w.r.t. the partition $X_1 \cup \cdots \cup X_r$ and $Y_1 \cup \cdots \cup Y_s$ is $\mathbb{E}[\mathbb{E}[G|\mathcal{P}]^2]$. *i.e.* It is the expectation of $d(X_i, Y_j)^2$ where (X_i, Y_j) is the pair containing a random edge (x, y).

The following lemma holds the key fact deriving the result.

Lemma 9) Let $A \subset X$, $B \subset Y$ and suppose that (A, B) is not ϵ-regular. Then there are partitions $A = A_0 \cup A_1$, $B = B_0 \cup B_1$ such that the mean-square density of $G\big|_{A \times B}$ is at least $d(A, B)^2 + \epsilon^2$.

proof) By hypothesis we can find $A_0 \subset A$, $B_0 \subset B$ such that

$$\Big|\mathbb{E}_{x \in A, y \in B}\Big(G(x, y) - d(A, B)\Big)A_0(x)B_0(y)\Big| \geq \epsilon$$

Set $A_1 = A \backslash A_0$, $B_1 = B \backslash B_0$. Then

$$\begin{aligned}\mathbb{E}[\mathbb{E}[G|\mathcal{P}]^2] =& \big(\mathbb{E}[\mathbb{E}[G|\mathcal{P}]]\big)^2 + \mathrm{Var}(\mathbb{E}[G|\mathcal{P}]) \\ =& d(A, B)^2 + \mathbb{E}\big(\mathbb{E}[G|\mathcal{P}] - d(A, B)\big)^2\end{aligned}$$

But

$$\begin{aligned}\mathbb{E}\big(\mathbb{E}[G|\mathcal{P}] - d(A, B)\big)^2 &\geq \mathbb{P}[x \in A_0, y \in B_0]\Big(d(A_0, B_0) - d(A, B)\Big)^2 \\ &= \frac{|A_0||B_0|}{|A||B|}\Big(\frac{1}{|A_0||B_0|}\sum_{x \in A_0, y \in B_0}\big(G(x, y) - d(A, B)\big)\Big)^2 \\ &\geq \Big(\mathbb{E}_{x \in A, y \in B}\big(G(x, y) - d(A, B)\big)A_0(x)B_0(y)\Big)^2 \geq \epsilon^2\end{aligned}$$

(End of proof) $\square$

Lemma 10) Suppose that the partition $X_1 \cup \cdots \cup X_r$, $Y_1 \cup \cdots \cup Y_s$ do not satisfy the conclusion of **Theorem 8**. Suppose also that $r, s \leq m$. Then there are refinements of the partition into at most $m \cdot 2^m$ sets each, such that the mean square density goes up by at least ϵ^3.

proof) For each pair (X_i, Y_j) that is not ϵ-regular, find a partition $X_i = X_{ij}^0 \cup X_{ij}^1$, $Y_j = Y_{ij}^0 \cup Y_{ij}^1$ w.r.t. the mean-square density of $G\big|_{X_i \times Y_j}$ is at least $d(X_i, Y_j)^2 + \epsilon^2$. For each i take a common refinement of the partitions $X_{ij}^0 \cup X_{ij}^1$ into at most 2^m sets, and do similarly for Y_j.

Pick a random edge (x, y) and suppose that it belongs to (X_i, Y_j).

- If (X_i, Y_j) is not ϵ-regular (observe that this happens with probability $\geq \epsilon$ by the assumption), then the mean square density of the cell of the refined partition is at least $d(X_i, Y_j)^2 + \epsilon^2$, because inside (X_i, Y_j) the refined partition refines the partitions $X_{ij}^0 \cup X_{ij}^1$, $Y_{ij}^0 \cup Y_{ij}^1$.
- Otherwise, it is at least $d(X_i, Y_j)^2$.

So the result follows.

(End of proof) $\square$

Finally we prove the Szemerédi's reglarity leamma.

proof of Theorem 8) Start with the trivial partitions of X and Y. If those don't work, apply **Lemma 10** repeatedly. Since mean-square density is ≤ 1, there can be at most ϵ^{-3} iterations before partitions are found that work. This gives an upper bound on k obtained by iterating the function $m \mapsto m2^m$ by ϵ^{-3} times. Since $m \cdot 2^m \leq 4^m$, this gives a bound of

$$4^{4^{\cdot^{\cdot^{\cdot^4}}}}$$

with ϵ^{-3} number of 4s.

(End of proof) □

(Note, this is less or equal to $2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}$ with $2+\epsilon^{-3}$ number of 2s.)

Theorem 11) *(The triangle-removed lemma)* For every $\epsilon > 0$ there exists $\delta > 0$ such that if G is any graph with n vertices and at most δn^3 triangles. Then there is a subgraph $H \subset G$ such that $G \backslash H$ contains at most ϵn^2 edges and H is triangle free.

==========

(13th November, Tuesday)

(due to Alan Baker's memorial, cannot give example class next Friday 2pm - but will be some time next week)

A quick remark : The graphs version (as opposed to bipartite graphs version) of *Szemerédi's regularity lemma* is essentially the same but with just one partition of $V(G)$.

Theorem 8') Let G be a graph and $\epsilon > 0$. Then there exists partition $G = X_1 \cup \cdots \cup X_r$ with $r < k(\epsilon)$, such that if you choose $(x, y) \in G \times G$ at random, then the probability that it belongs to an ϵ-regular pair (X_i, X_j) is at least $1 - \epsilon$.

To prove it, form a bipartite graph with two copies of $V(G)$ and run the previous argument, but at each iteration do a further refinement to keep the two partitions "the same", so instead of having $m \mapsto m \cdot 2^m$, would have $m \mapsto m \cdot 4^m$.

We use **Theorem 8'** in the proof of **Theorem 11**.

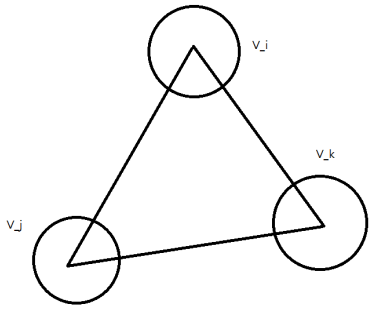
proof of Theorem 11) Apply the regularity lemma with parameter θ. It gives us a partition of the vertex set into sets $V_1, \cdots, V_K$. Throw away :

- all edges with a vertex in some V_i of density $< \epsilon/6K$
- all edges joining V_i to V_j if $d(V_i, V_j) < \epsilon/3$
- all edges joining V_i to V_j if (V_i, V_j) is not a θ-regular pair.

As long as $\theta \leq \epsilon/3$, the total density of removed edges is at most

$$2K\frac{\epsilon}{6K} + \frac{\epsilon}{3} + \theta \leq \epsilon$$

Now suppose that after these deletions, there is still a triangle. Let its vertices belong to V_i, V_j, V_k.



(V_i, V_j, V_k not necessarily distinct) Then the pairs (V_i, V_j), (V_j, V_k), (V_i, V_k) are all θ-regular of density $\delta' \geq \epsilon/3$ so

$$\begin{aligned}
&\left|\mathbb{E}_{x\in V_i, y\in V_j, z\in V_k} G(x,y)G(y,z)G(x,z) - \delta'^3\right| \\
\leq &\left|\mathbb{E}_{x,y,z}(G(x,y) - \delta')G(y,z)G(x,z)\right| + \delta'\left|\mathbb{E}_{x,y,z}(G(y,z) - \delta')G(x,z)\right| \\
\leq &\theta + \delta'\theta \leq 2\theta
\end{aligned}$$

(recall, the definition of being θ-regular was equivalent to having $\left|\mathbb{E}(G(x,y)-\delta)A(x)B(y)\right| \leq \theta$). So the number of triangles in G is at least $|V_i||V_j||V_k|\left(\frac{\epsilon}{3}\right)^2 - 2\theta$. Since θ was arbitrary, if we pick $\theta \leq \frac{\epsilon^3}{108}$, then

$$|V_i||V_j||V_k|\left(\frac{\epsilon}{3}\right)^2 - 2\theta \geq \left(\frac{\epsilon}{6K}\right)^3\left(\frac{\epsilon^3}{54}\right)n^3$$

So if we choose $\delta = \frac{1}{2}\left(\frac{\epsilon}{6K}\right)^3\frac{\epsilon^3}{54}$, and G was assumed to have $\leq \delta n^3$ from the start, then this result makes a contradiction, i.e. removal of edges from G as in the start of the theorem gives a subgraph without any triangle.

(End of proof) □

Theorem 12) *(The corners theorem)* For every $\delta_0 > 0$ there exists n such that every subset A of $[n]^2$ of density at least $\delta \geq \delta 0$ contains a triple $(x,y),(x+d,y),(x,y+d)$ with $d \neq 0$ (called a corner).
(this does not guarantee the containment of the triangle ordered in the reverse direction, but will be asked to prove this in the 3rd example sheet)

proof) For a tripartite graph with vertex sets $X = Y = [n]$, $Z = [2n]$. Join $x \in X$ to $y \in Y$ if $(x,y) \in A$, join $x \in X$ to $z \in Z$ if $(x, z-x) \in A$, join $y \in Y$ to $z \in Z$ if $(z-y, y) \in A$. If (x,y,z) is a triangle, then, writing $d = z - x - y$, we have $(x,y) \in A$, $(x, z-x) = (x, y+d) \in A$, $(z-y, y) = (x+d, y) \in A$.

So if A contains no corners(i.e. the triple of edges in the statement of the theorem), then the only triangles in the graph are "degenerate" triangles where $x + y = z$, so $d = 0$. So the number of triangles is exactly $\delta n^2 = o(n^3)$. But the degenerate triangles are edge disjoint, since any two of x, y, z determine the third if $x + y = z$(that is, if $(x,y) \in A$ then $(x, z-x) \in A$ and $(z-y, y) \in A$). So the number of edges you have to remove to get rid of all triangles is at least δn^2. For sufficiently large n, this contradicts the triangle removal lemma.

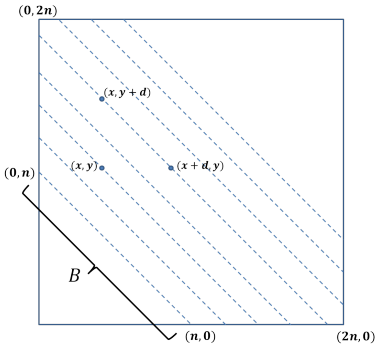
: if we choose $\epsilon = \frac{1}{2}\delta_0$, then we should have δ' such that the **Theorem 11** holds, but $\delta n^2 < \delta' n^3$ in the limit $n \to \infty$, so we should be able to find only $\frac{1}{2}\delta_0$ edges that removes all triangles, which is a contradiction.

(End of proof) □

This theorem was proved by Ruzsa, Szemerédi and Solymosi.

This lemma actually implies Roth's theorem - turn the sets in the diagonal arrangement.

second proof of Roth's theorem) Let $B \subset [n]$ be a subset of density δ. Consider the following



subset $A = \{(x,y) \in [2n]^2 : (y-x) \in B \pmod n\} \subset [2n]^2$. Then A has density at least $\frac{1}{2}\delta$ in $[2n]^2$. Then by the corners theorem, for sufficiently large n, we have $d \neq 0$ such that $(x,y),(x+d,y),(x,y+d) \in A$, which implies $y-x-d, y-x, y-x+d \in B \bmod n$. Hence B contains an arithmetic progression of length 3.

(End of proof) □

4. The Polynomial Method

Let X be a finite set and let $f : X^3 \to \mathbb{F}$ for some field $\mathbb{F}$. The **slice rank** of f is the smallest k such that there are functions $u_1, \cdots, u_k : X \to \mathbb{F}$, $v_1, \cdots, v_k : X^2 \to \mathbb{F}$, k_1, k_2 such that

$$f(x,y,z) = \sum_{i=1}^{k_1} u_i(x)v_i(y,z) + \sum_{i=k_1+1}^{k_2} u_i(y)v_i(x,z) + \sum_{i=k_2+1}^{k} u_i(z)v_i(x,y)$$

==========

(15th November, Thursday)

The frameworks we are working on is established by Croot, Lev, Pach / Ellenberg, Gijswijt/ and worked in the form we are presenting by T. Tao

Lemma 1) Let X be a finite set, let $A \subset X$ and let $f : X^3 \to \mathbb{F}$ be a function such that $f(x,y,z) \neq 0$ *iff* $x = y = z \in A$. (can think of f as a diagonal matrix) Then f has slice rank $|A|$.

proof) We can rewrite $f(x,y,z)$ as $\sum_{a\in A} \lambda\delta_a(x)\delta_a(y)\delta_a(z)$ so the slice rank is at most $|A|$.

Now suppose that

$$f(x,y,z) = \sum_{i=1}^{k_1} u_i(x)v_i(y,z) + \sum_{k_1+1}^{k_2} u_i(y)v_i(x,z) + \sum_{k_2+1}^{k} u_i(z)v_i(x,y)$$

Let V be the subspace of $\mathbb{F}^X$ that consist of functions orthogonal to all of $u_1, \cdots, u_{k_1}$. That is, $h \in V$ *iff* $\sum_x h(x)u_i(x) = 0$ for all $i \leq k$. Then V has codimension at most k_1. It follows that V contains some function h that vanishes at most k_1 times.(i.e. there are at most k_1 number of $x \in X$ s.t. $h(x) = 0$.)

: to see this, form an $m \times n$ matrix whose rows are a basis for V (so $m \geq n - k_1$). Put the matrix into reduced row-echelon form. Add up the rows to get the desired function.

Now define $g : X^2 \to \mathbb{F}$ by $g(y,z) = \sum_x f(x,y,z)h(x)$. Then $g(y,z) \neq 0$ *iff* $y = z \in A$ and $h(y) \neq 0$. This holds at least $|A| - k_1$ times, so g has rank $\geq |A| - k_1$. But also $g(y,z)$ has a formula of the form

$$\sum_{k_1+1}^{k_2} u_i(y)w_i(z) + \sum_{k_2+1}^{k} u_i(z)w_i(y) \quad \text{for some functions } w_i$$

(e.g. if $k_1 < i \leq k_2$ then $w_i(z) = \sum_x u_i(x,z)h(x)$) since h is orthogonal to each $u_1, \cdots, u_{k_1}$. Therefore, g has rank at most $k - k_1$. Therefore, $|A| \leq k$.

(End of proof) □

Let $F(n)$ be the number of sequences $\{0,1,2\}$ of length n that add up to at most $\frac{2n}{3}$.

Lemma 2) Let A be a subset of $\mathbb{F}_3^n$ such that if $x, y, z \in A$ and $x + y + z = 0$ (this is precisely an AP of length 3) then $x = y = z$. Then $|A| \leq 3F(n)$.

proof) The function $f(x,y,z) = \sum_{a\in A} \delta_a(x)\delta_a(y)\delta_a(z)$ on A^3 can be expressed by the formula

$$f(x,y,z) = \prod_{i=1}^{n}(1 - (x_i + y_i + z_i))(1 + (x_i + y_i + z_i)) = \prod_{i=1}^{n}(1 - (x_i + y_i + z_i)^2)$$

This is a polynomial in $x_1, \cdots, x_n, y_1, \cdots, y_n, z_1, \cdots, z_n$ with each x_i, y_i, z_i occurring with degree 0,1 or 2. Also, it has total degree $2n$. It is therefore a linear combination of monomials, and in each one, either x variables or y variables or z variables occur with total degree at most $\frac{2n}{3}$. Partition the monomials into three sets accordingly. It follows that we can write $\sum_{a\in A} \delta_a(x)\delta_a(y)\delta_a(z)$ as

$$\sum_{i=1}^{m_1} u_i(x)v_i(y,z) + \sum_{i=m_1+1}^{m_2} u_i(y)v_i(x,z) + \sum_{i=m_1+1}^{m} u_i(z)v_i(x,y)$$

where the u_i are monomials of total degree at most $\frac{2n}{3}$ and the v_i are polynomials (in $2n$ variables). The number of monomials in $x_1, \cdots, x_n$ such that each x_i occurs with degree $0, 1$ or 2 and the total degree is $\leq \frac{2n}{3}$ is $F(u)$. So $|A| \leq 3F(n)$ by **Lemma 1**.

(End of proof) □

Lemma 3) $F(n) \leq e^{-n/12}3^n$

proof) Let x be a random sequence in $\{0,1,2\}^{2n}$ and let $X_1, \cdots, X_n$ be independent random variables uniformly distributed in $\{-1,0,1\}$. Then

$$\begin{aligned}\mathbb{P}\Big[\sum(1-X_i) \leq \frac{2n}{3}\Big] &= \mathbb{P}\Big[\sum_{i=1}^{n} X_i \geq n/3\Big] = \mathbb{P}\Big[e^{\lambda\sum_{i=1}^n X_i} \geq e^{\lambda n/3}\Big] \quad \text{(for any } \lambda\text{)} \\ &\leq e^{-\lambda n/3}\mathbb{E}\Big[e^{\lambda\sum_{i=1}^n X_i}\Big] = e^{-\lambda n/3}\prod_{i=1}^{n}\Big(\mathbb{E}e^{\lambda X_i}\Big) \\ &= e^{-\lambda n/3}\Big(\frac{1+e^{\lambda}+e^{-\lambda}}{3}\Big)^n\end{aligned}$$

But

$$\begin{aligned}\frac{1+e^{\lambda}+e^{-\lambda}}{3} &= 1+\frac{2}{3}\Big(\frac{\lambda^2}{2!}+\frac{\lambda^4}{4!}+\cdots\Big) \\ \leq 1+\frac{\lambda^2}{3}+\Big(\frac{\lambda^2}{2}\Big)^2\frac{1}{2!}+\Big(\frac{\lambda^2}{3}\Big)^3\frac{1}{3!}+\cdots &= e^{\lambda^2/3}\end{aligned}$$

so the probability is at most $e^{-\frac{\lambda n}{3}+\frac{\lambda^2 n}{3}}$. Choosing $\lambda = 1/2$ gives an upper bound of $e^{-n/12}$

(End of proof) □

This estimate gives the following theorem.

Theorem 4) Let $A \subset \mathbb{F}_3^n$ contain no non-trivial solutions to $x+y+z=0$. Then $|A| \leq e^{-n/12}3^n$.

proof) Follows directly from **Lemma 2** and **Lemma 3**

(End of proof) □