Topics in Combinatorics

W.T. Gowers, F.R.S.

Lent 2004

1 Random walks on graphs

We shall consider a graph G with n vertices that is regular of degree d. (We shall allow loops and multiple edges.) We can define a random walk on G by starting at a vertex $x = x_0$, and at each step moving from the current vertex x_n to one of its neighbours—each chosen with probability $\frac{1}{d}$.

The *adjacency matrix* A of G is defined by

$$A_{xy} = \begin{cases} 1 & xy \text{ is an edge of } G \\ 0 & \text{otherwise} \end{cases}$$

(or more generally A_{xy} is defined to be the number of edges from x to y).

The *transition matrix* T is $\frac{1}{d}A$. Observe that T_{xy} is the probability that you go to y if you are at x.

The useful thing about T is that $(T^k)_{xy}$ is the probability that, after k steps, you are at y if you start at x. (Easy inductive proof.)

T is a symmetric matrix, so there is an orthonormal basis of eigenvectors $v_0, v_1, \dots, v_{n-1}$ with eigenvalues $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{n-1}$.

Lemma 1. *With T defined as above, $\lambda_0 = 1$, and $\lambda_1 < 1$ if G is connected. Furthermore, $\lambda_{n-1} \geq -1$, with equality iff G is bipartite.*

Proof. Let $\phi : V(G) \to \mathbb{R}$ be the constant function taking the value 1 everywhere. Then

$$T\phi(x) = \sum_{y \in V(G)} T_{xy}\phi(y) = \frac{1}{d} \sum_{y \in N(x)} \phi(y) = 1 = \phi(x)$$

(as G is d-regular). So ϕ is an eigenvector with eigenvalue 1.

Now we show that $|\lambda_i| \leq 1$ for every i. Write $\|f\|_1$ for $\sum_x |f(x)|$. Then, for any $\phi : V(G) \to \mathbb{R}$,

$$\begin{aligned}
\|T\phi\|_1 &= \sum_{x \in V(G)} \left| \frac{1}{d} \sum_{y \in N(x)} \phi(y) \right| \\
&\leq \frac{1}{d} \sum_{x \in V(G)} \sum_{y \in N(x)} |\phi(y)| \\
&= \frac{1}{d} \sum_{y \in V(G)} |\phi(y)| \sum_{x \in N(y)} 1 \\
&= \sum_{y \in V(G)} |\phi(y)| \\
&= \|\phi\|_1.
\end{aligned}$$

Hence if $T\phi = \lambda\phi$ then $|\lambda| \leq 1$.

Now, let us suppose that $\|T\phi\|_1 = \|\phi\|_1$, so that equality holds in the inequality above. Then $\left| \sum_{y \in N(x)} \phi(y) \right| = \sum_{y \in N(x)} |\phi(y)|$ for every x, so no x has neighbours y and z with $\phi(y) > 0$ and $\phi(z) < 0$.

Suppose that ϕ is an eigenvector and $\phi(x) = 0$. Then

$$T\phi(x) = \frac{1}{d} \sum_{y \in N(x)} \phi(y) = 0,$$

so $\phi(y) = 0$ for all $y \in N(x)$. If G is connected then ϕ vanishes everywhere, a contradiction. So if ϕ is an eigenvector with eigenvalue ± 1 then $\phi(x)$ is never 0. If the eigenvalue is 1 then the sign of $\phi(y)$ is the same as that of $\phi(x)$ for all $y \in N(x)$. Hence ϕ has constant sign, so $\langle \phi, \mathbf{1} \rangle = \sum_x \phi(x).1 \neq 0$. So if ϕ is one of the v_i then $\phi = v_0$.

If the eigenvalue is -1, then $\phi(y)$ has opposite sign to $\phi(x)$ for every $y \in N(x)$. So $\operatorname{sign} \phi$ defines a bipartition of G.

□

Lemma 2. *Let G be a graph with n vertices, regular of degree d. Suppose that $\lambda = \max\{|\lambda_1|, |\lambda_{n-1}|\} < 1$. Then the random walk on G converges to the uniform distribution.*

Proof. Let v_0, v_1, …, v_n be an orthonormal basis of eigenvectors of the transition matrix T (ordered as before). Then we can write the initial distribution p_0 of the random walk as $\mu_0 v_0 + \mu_1 v_1 + \cdots + \mu_{n-1} v_{n-1}$ for some scalars μ_0, μ_1, …, μ_{n-1}.

We know that $\mu_0 v_0$ is the uniform distribution by looking at the expression $\sum_x \sum_{i=0}^{n-1} \mu_i v_i(x) = \sum_x p_0(x) = 1$. But $\langle v_i, v_0 \rangle = 0$ for all $i > 0$,

i.e. $\sum_x v_i(x) = 0$ for all $i > 0$, so this is also $\sum_x \mu_0 v_0(x)$, so $\mu_0 v_0$ is a constant function summing to 1.

The distribution p_k after k steps is then

$$T^k p_0 = \mu_0 v_0 + \mu_1 \lambda_1^k v_1 + \cdots + \mu_{n-1} \lambda_{n-1}^k v_{n-1}.$$

So

$$\|T^k p_0 - \mu_0 v_0\|_2^2 = \sum_{i=1}^{n-1} \mu_i^2 \lambda_i^{2k} \leq \lambda^{2k} \sum_{i=1}^{n-1} \mu_i^2 \leq \lambda^{2k}$$

since $\sum_{i=0}^{n-1} \mu_i^2 = \|p_0\|_2^2 \leq 1$. Hence $\|T^k p_0 - \mu_0 v_0\|_1 \leq \lambda^k \sqrt{n}$.

But $\lambda^k \to 0$, so p_k converges to the uniform distribution in ℓ_1 (or in ℓ_2 or in any other norm). □

Definition. The *total variation distance* between two probability distributions μ_1 and μ_2 on a set X (which we will take to be finite) is given by $\max_{A \subset X}\{|\mu_1(A) - \mu_2(A)|\}$. It is easy to see that the best A to choose is either $\{x : \mu_1(\{x\}) \geq \mu_2(\{x\})\}$ or $\{x : \mu_1(\{x\}) < \mu_2(\{x\})\}$. But since $\sum \mu_1(\{x\}) = \sum \mu_2(\{x\})$, these give the same answer, so both must give $\frac{1}{2}\|\mu_1 - \mu_2\|_1$.

Example. The *discrete cube* Q_n has vertex set $\mathbb{P}[n]$, where $[n] = \{1, 2, \ldots, n\}$. Two sets A, $B \subset [n]$ are joined by an edge if $|A \Delta B| = 1$, i.e. if A and B differ by exactly one element. Equivalently, it is the set of 01-sequences of length n, with $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ joined to $(\varepsilon_1', \varepsilon_2', \ldots, \varepsilon_n')$ iff $\varepsilon_i = \varepsilon_i'$ for all but one value of i.

We shall write down a complete system of eigenvectors.

Definition. Let $B \subset [n]$. The *Walsh function* $W_B : \mathbb{P}[n] \to \mathbb{R}$ is defined by $W_B(A) = (-1)^{|A \cap B|}$, e.g. $W_\emptyset(A) = 1$ for every $A \subset [n]$ and $W_{[n]}(A) = (-1)^{|A|}$.

Claim 1. *If $B \neq B'$ then $\langle W_B, W_{B'} \rangle = 0$.*

Proof. For $i \in B \Delta B'$, we have

$$\sum_{A \subset [n]} W_B(A) W_{B'}(A) = \sum_{A \subset [n] - \{i\}} \big(W_B(A) W_{B'}(A) + W_B\left(A \cup \{i\}\right) W_B\left(A \cup \{i\}\right) \big).$$

If $i \in B - B'$ then $W_B(A) = -W_B\left(A \cup \{i\}\right)$ and $W_{B'}(A) = W_{B'}\left(A \cup \{i\}\right)$, so the big bracket is zero for every A. Similarly if $i \in B' - B$. □

Claim 2. *Each W_B is an eigenvector of T with eigenvalue $1 - \frac{2|B|}{n}$.*

Proof. We have

$$\begin{aligned} TW_B(A) &= \frac{1}{n}\sum_{\substack{C\subset[n]\\ |A\Delta C|=1}} W_B(C) \\ &= \frac{1}{n}\sum_{i=1}^{n} W_B\left(A\Delta\{i\}\right) \\ &= \frac{1}{n}\sum_{i\in B} W_B\left(A\Delta\{i\}\right) + \frac{1}{n}\sum_{i\notin B} W_B\left(A\Delta\{i\}\right) \\ &= \frac{1}{n}W_B(A)\big(-|B|+n-|B|\big) \\ &= \left(1-\frac{2|B|}{n}\right)W_B(A). \end{aligned}$$

□

This proves that $\lambda_1 = 1-\frac{2}{n}$ and $\lambda_{n-1} = -1$. We can deal with the fact that the graph is bipartite by staying still with probability $\frac{1}{2}$. Then the transition matrix T changes to $\frac{1}{2}(T+I)$, so then the eigenvalue λ changes to $\frac{\lambda+1}{2}$.

We can do better as follows. If we start at a single vertex—wlog $\emptyset$—then the initial distribution is $\delta_\emptyset$. Since $\|W_B\|_2 = 2^{\frac{n}{2}}$, we have

$$\delta_\emptyset = 2^{-n}\sum_{B\subset[n]}\langle\delta_\emptyset, W_B\rangle W_B = 2^{-n}\sum_{B\subset[n]} W_B.$$

If we use the transition matrix $T = \frac{1}{2}(I+\frac{1}{n}A)$ then W_B has eigenvalue $1-\frac{|B|}{n}$. Then $T^k\delta_\emptyset = 2^{-n}\sum_{B\subset[n]}\left(1-\frac{|B|}{n}\right)W_B$. So

$$\begin{aligned} \|T^k\delta_\emptyset - 2^{-n}W_\emptyset\|_2^2 &= 2^{-2n}\sum_{\substack{B\subset[n]\\ B\neq\emptyset}}\left(1-\frac{|B|}{n}\right)^{2k}\|W_B\|_2^2 \\ &= 2^{-n}\sum_{r=1}^{n}\left(1-\frac{r}{n}\right)^{2k}\binom{n}{r} \\ &\leq 2^{-n}\sum_{r=1}^{n} e^{\frac{-2kr}{n}}n^r \\ &= 2^{-n}\sum_{r=1}^{n}\left(e^{\log n-\frac{2k}{n}}\right)^r. \end{aligned}$$

If $k = Cn \log n$ with $C \geq 1$ then this is at most $2^{-n} \sum_{r=1}^{n} e^{-Cr \log n} \leq 2^{-n} 2n^{-c}$, so $\|T^k \delta_\emptyset - 2^{-n} W_\emptyset\|_2 \leq 2^{\frac{1}{2}} 2^{-\frac{n}{2}} n^{-\frac{C}{2}}$, and so $\|T^k \delta_\emptyset - 2^{-n} W_\emptyset\|_1 \leq 2^{\frac{1}{2}} n^{-\frac{C}{2}}$. Hence the mixing time is $\lesssim n \log n$ (in fact $\approx$).

Example. For convenience, let N be odd and consider the graph on $\mathbb{Z}_N$, the integers modulo N, with x joined to $x \pm 1$. Let $\omega = e^{\frac{2\pi i}{N}}$ and for each r let $f_r(x) = \omega^{rx}$. Then

$$\begin{aligned}(Tf_r)(x) &= \frac{1}{2} f_r(x+1) + \frac{1}{2} f_r(x-1) \\ &= \frac{1}{2}(\omega^{r(x+1)} + \omega^{r(x-1)}) \\ &= \frac{1}{2}(\omega^r + \omega^{-r})\omega^{rx} = \cos \frac{2\pi r}{N} . f_r(x).\end{aligned}$$

So f_r is an eigenvector with eigenvalue $\cos \frac{2\pi r}{N}$. So is f_{-r}, since cos is even. It follows that $x \mapsto \cos \frac{2\pi r x}{N}$ and $x \mapsto \sin \frac{2\pi r x}{N}$ are eigenvectors with eigenvalue $\cos \frac{2\pi r}{N}$. So $\lambda_1 = \cos \frac{2\pi}{N}$ and $\lambda_{N-1} = \cos 2\pi \frac{(N+1)/2}{N} = \cos\left(\pi + \frac{\pi}{N}\right)$ so the spectral gap is $\sim N^{-2}$. So the mixing time is $\lesssim N^2 \log N$.

Lemma 3. *Let G be a d-regular graph. Then*

$$\lambda_1 = \sup \left\{ \frac{\langle Tf, f \rangle}{n \operatorname{var} f} : f \not\equiv 0, \sum_x f(x) = 0 \right\}.$$

Proof. Let v_0, v_1, $\ldots$, v_{n-1} be an orthonormal basis of eigenvectors with eigenvalues $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{n-1}$. Then, as $\sum_{x \in V(G)} f(x) = 0$, we can expand f as $f = \mu_1 v_1 + \cdots + \mu_{n-1} v_{n-1}$. Then $\langle f, Tf \rangle = \sum_{i=1}^{n-1} \lambda_i \mu_i^2$ and $n \operatorname{var} f = \sum_{i=1}^{n-1} \mu_i^2$ $(= \|f\|_2^2)$. But $\sum_{i=1}^{n-1} \lambda_i \mu_i^2 \leq \lambda_1 \sum_{i=1}^{n-1} \mu_i^2$ with equality if $\mu_1 = 1$ and the other μ_i are zero. □

Now

$$\begin{aligned}\langle f, Tf \rangle &= \sum_x f(x) Tf(x) \\ &= \frac{1}{d} \sum_x f(x) \sum_{y \in N(x)} f(y) \\ &= \frac{1}{d} \sum_{\substack{(x,y) \\ x \sim y}} f(x) f(y)\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2d} \sum_{\substack{(x,y) \\ x \sim y}} \left(f(x)^2 + f(y)^2 - (f(x) - f(y))^2 \right) \\
&= n \operatorname{var} f - \frac{1}{2d} \sum_{\substack{(x,y) \\ x \sim y}} (f(x) - f(y))^2 .
\end{aligned}
$$

Also, $\operatorname{var} f = \frac{1}{2n^2} \sum_{(x,y)} (f(x) - f(y))^2$ (exercise).

If we denote by $G\text{-var}\, f$ the quantity $\frac{1}{2dn} \sum_{(x,y):x \sim y} (f(x) - f(y))^2$, this shows that the spectral gap is at least δ if $G\text{-var}\, f \geq \delta \operatorname{var} f$ for every f such that $\sum_x f(x) = 0$ and hence for all f.

So we have shown:

Lemma 4. *Let G be a d-regular graph. Then the spectral gap, $1 - \lambda_1$, is given by*

$$1 - \lambda_1 = \inf_f \frac{G\text{-var}\, f}{\operatorname{var} f},$$

where the infimum is taken over all non-constant $f : V(G) \to \mathbb{R}$.

Proposition 5 (Discrete Poincaré Inequality (Diaconis, Stroock)). *Let G be a d-regular graph with n vertices. Suppose that there is a system $\mathcal{P}$ of directed paths with the following properties:*

- *for every pair (x, y) there is a path $P_{xy} \in \mathcal{P}$ from x to y of length at most m; and*
- *no (directed) edge appears in more than t of the paths.*

Then the spectral gap is at least $\frac{n}{mtd}$.

Proof. Let $f : V(G) \to \mathbb{R}$ be any function. Then (writing e_- for the start-vertex and e_+ for the end-vertex of an edge e)

$$
\begin{aligned}
\operatorname{var} f &= \frac{1}{2n^2} \sum_{(x,y)} (f(x) - f(y))^2 \\
&= \frac{1}{2n^2} \sum_{(x,y)} \left(\sum_{e \in P_{xy}} (f(e_+) - f(e_-)) \right)^2 \\
&\leq \frac{1}{2n^2} \sum_{(x,y)} |P_{xy}| \sum_{e \in P_{xy}} (f(e_+) - f(e_-))^2 \qquad \text{(Cauchy-Schwarz)} \\
&\leq \frac{m}{2n^2} \sum_{e} \sum_{P_{xy} \ni e} (f(e_+) - f(e_-))^2
\end{aligned}
$$

$$\begin{aligned} &\leq \frac{mtd}{n} \cdot \frac{1}{2dn} \sum_e (f(e_+) - f(e_-))^2 \\ &= \frac{mtd}{n} G\text{-var } f. \end{aligned}$$

□

Example. Let G be the discrete n-dimensional cube. Given two vertices A and B (which are subsets of $[n]$), define a path $A = A_0, A_1, \ldots, A_n = B$ as follows:

- if $k \in B$ then $A_k = A_{k-1} \cup \{k\}$;
- if $k \notin B$ then $A_k = A_{k-1} - \{k\}$.

Now let CD be an edge of the graph G and suppose that C and D differ only in the kth place. If CD belongs to the path P_{AB} then

$$A \cap \{k+1, k+2, \ldots, n\} = D \cap \{k+1, k+2, \ldots, n\}$$

and

$$B \cap \{1, 2, \ldots, k\} = D \cap \{1, 2, \ldots, k\}.$$

Hence the number of pairs (A, B) with $CD \in P_{AB}$ is at most $2^k 2^{n-k} = 2^n$, i.e. we can take $t = 2^n$.

So the spectral gap is at least $\frac{2^n}{n2^n n} = \frac{1}{n^2}$. Note that this is a good estimate but is not best possible.

Alternative way to see that $t = 2^n$: note that, given CD, the pair (A, B) is determined by $A \cap \{1, 2, \ldots, k\}$ and $B \cap \{k+1, k+2, \ldots, n\}$. So we can define a surjection from $\mathbb{P}[n]$ to $\{(A, B) : CD \in P_{AB}\}$ by

$$E \mapsto \Big(\big(E \cap [1,k]\big) \cup \big(D \cap [k+1,n]\big), \big(D \cap [1,k]\big) \cup \big(E \cap [k+1,n]\big) \Big)$$

(where for integers a and b, $[a, b]$ denotes the set $\{a, a+1, \ldots, b\}$). So $t = |V(G)|$.

Approximating the permanent

Let A be an $n \times n$ matrix. The *permanent* of A is $\sum_{\pi \in S_n} \prod_{i=1}^n A_{i\pi(i)}$, i.e. "det A without the signs".

We shall restrict attention to 01-matrices. There is a natural correspondence between these and bipartite graphs, with an edge from i to j if $A_{ij} = 1$.

The product $\prod_{i=1}^n A_{i\pi(i)} = 1$ iff all the edges $i\pi(i)$ belong to the graph G, so the permanent of A is the number of perfect matchings in G.

Two ideas lie behind the proof to follow:

1. If you can efficiently generate a random matching in the graph (from the uniform distribution) then you can do approximate counting with high probability.

2. To do (1), devise a random walk in the space of all matchings and show that it mixes rapidly.

We shall restrict attention to *dense* bipartite graphs—i.e. ones where the minimum degree is at least $\frac{n}{2}$. (It is still $\#P$-hard to calculate the number of matchings exactly.)

Broder's Algorithm

Let G be a dense bipartite graph with n vertices on each side, let $\mathcal{M}$ be the set of perfect matchings in G and let $\mathcal{M}'$ be the set of *near-perfect matchings*—i.e. those matchings with exactly one vertex unmatched on each side. Let $\mathcal{N} = \mathcal{M} \cup \mathcal{M}'$.

We define a random walk on $\mathcal{N}$ as follows: given $M \in \mathcal{N}$, choose an edge $e \in G$, $e = uv$, uniformly at random, and do the following:

(i) if $M \in \mathcal{M}$ and $e \in M$ then remove e;

(ii) if $M \in \mathcal{M}'$ and u, v are unmatched in M then add e;

(iii) if $M \in \mathcal{M}'$, u is unmatched in M, and v is matched to w then replace the edge wv by e;

(iv) similarly if u is matched and v is not;

(v) otherwise, do nothing.

We can define a corresponding graph on $\mathcal{N}$ (with several loops at each vertex) which is regular of degree $|E(G)|$.

Defining a system of paths

First we show how to get from $M' \in \mathcal{M}'$ to some $M \in \mathcal{M}$. Let u and v be the unmatched vertices in M'. If $uv \in E(G)$ then let $M = M' \cup \{uv\}$. Otherwise, u must have a neighbour y and v a neighbour x such that $xy \in M'$. Pick such a pair and move to a matching M in two steps as follows: first, replaced xy by uy and then add in the edge xv. From this matching M, we can recover M' if we are told what u and v were, so at most n^2 of the $M' \in \mathcal{M}'$ end up at M. This also shows that $|\mathcal{M}'| \leq n^2|\mathcal{M}|$.

Now we shall see how to get from a matching M_1 to a matching M_2. First, put an arbitrary total ordering on the set of all cycles in G, and for each cycle pick a "start" vertex. The symmetric difference of M_1 and M_2 is a disjoint union of cycles. For each of these cycles in turn, we do the following. First, remove the edge of M_1 incident to the start-vertex. Next, keep replacing edges of M_1 by edges of M_2 wherever possible. When all edges belong to M_2, add in the M_2-edge incident to the start-vertex.

Bounding the bottleneck parameter

How many pairs (M_1, M_2) use any given transition MN in the path just defined? We cannot hope to say, but we can try to bound the ratio of this to $|\mathcal{N}|$. We shall do that by showing that the extra information needed to recover (M_1, M_2) if you know MN is in 1-1 correspondence with a subset of $\mathcal{N}$.

Suppose that the transition MN lies in the path from M_1 to M_2, where M and N are near-matchings, and M_1 and M_2 are perfect matchings. If we know MN and $M_1 \Delta M_2$ (the union of the cycles included in the path from M_1 to M_2), then we can certainly determine M_1 and M_2. We can determine $M_1 \Delta M_2$ from $M_1 \Delta M_2 \Delta (M \cup N)$. However, this will not be a near-matching. But it can be made into a near-matching by the removal of one edge, and we can reconstruct the union of cycles from it with this edge removed. Similar arguments work for the other kinds of transition so the number of paths using MN is at most $|\mathcal{N}|$ for any transition MN.

Since any matching in $\mathcal{M}$ is connected to at most n^2 near-matchings by the paths discussed earlier, we can get from any point in $\mathcal{N}$ to any other with each edge used at most $2n^4|\mathcal{N}|$ times. The paths have length at most $2n$ and the degree of each vertex in the graph on $\mathcal{N}$ is at most n^2. So the spectral gap is at most $\frac{|\mathcal{N}|}{2n.n^2.2n^4|\mathcal{N}|} = \frac{1}{4n^7}$. So the walk mixes in polynomial time.

Random sampling $\rightsquigarrow$ approximate counting

Suppose we know how to sample uniformly at random from $\mathcal{N}(G)$. We can use this to estimate $|M(G)|/|\mathcal{N}(G \cup \{e\})|$ as follows.

Sample repeatedly from $\mathcal{N}(G \cup \{e\})$ and count how many times you get an element of $\mathcal{M}(G)$. For this to work, it is important that $|\mathcal{N}(G \cup \{e\})|$ is not too much larger than $|\mathcal{M}(G)|$. Any matching in $G \cup \{e\}$ is either a matching in G or the union of $\{e\}$ and a near-matching in G. So

$$|\mathcal{M}(G \cup \{e\})| \leq |\mathcal{N}(G)| \leq (n^2+1)|\mathcal{M}(G)|$$

and so

$$|\mathcal{N}(G \cup \{e\})| \leq (n^2+1)^2|\mathcal{M}(G)|.$$

By the same argument we can estimate $|\mathcal{N}(G\cup\{e\})|/|\mathcal{N}(G\cup\{e,e'\})|$ and so on. Since we know $|\mathcal{N}(K_{n,n})|$ exactly, the size of $\mathcal{M}(G)$ can be calculated from at most n^2 ratios.

We have to make sure that the sum of the error probabilities is small, and that the product of the errors in the ratio estimates is small. This can easily be achieved in polynomial time.

Coupling

This is another technique for proving rapid mixing that works for some very symmetric walks.

Example. The cube. Think of it as 01-sequences of length n. Define two random walks x_0, x_1, x_2, ... and y_0, y_1, y_2, ... as follows. Let x_0 be a fixed vertex and y_0 a random vertex. At time t, choose uniformly at random from $\{1, 2, \dots, n\}$ and randomly decide whether to take $(x_t)_i$ and $(y_t)_i$ to be 0 or 1 (but making the same decision for both), and leave the other coordinates the same as they were for x_{t-1} and y_{t-1}. What do we know about these walks?

(i) Individually, the walks (x_t) and (y_t) are the usual walks with holding probability $\frac{1}{2}$;

(ii) y_t is uniformly distributed for all t;

(iii) if each i has been chosen then $x_t = y_t$.

The probability that some i has not been chosen by time t is at most $n(1-1/n)^t \le e^{\log n - t/n}$, which is small for $t \gg n\log n$. So the mixing time is roughly $n\log n$.

2 Quasirandom graphs

If G is a graph and $x \in V(G)$ then we shall write N_x for the set of neighbours of x.

Theorem 6. *Let G be an $\frac{N}{2}$-regular bipartite graph with vertex sets X and Y of size N. Then the following statements are equivalent, in the sense that $c_i \to 0 \implies c_j \to 0$.*

(i) $\sum_{x,x'\in X} |N_x \cap N_{x'}|^2 \le \frac{N^4}{16} + c_1 N^4$.

(ii) The number of labelled 4-cycles (x, y, x', y') in G with x, $x' \in X$ and y, $y' \in Y$ is at most $\frac{N^4}{16} + c_1 N^4$.

(iii) $\big||N_x \cap N_{x'}| - \frac{N}{4}\big| \le c_2 N$ *for all but at most* $c_2 N^2$ *pairs* $(x, x') \in X^2$.

(iv) For every $A \subset X$ *and* $B \subset Y$*, the number of edges from* A *to* B *differs from* $\frac{1}{2}|A||B|$ *by at most* $c_3 N^2$.

Proof. **(i) $\iff$ (ii).** This follows from the fact that $|N_x \cap N_{x'}|^2$ is the number of pairs $(y, y') \in Y^2$ such that (x, y, x', y') is a labelled 4-cycle of the required kind.

(i) $\implies$ (iii). Assume (i) and consider the sum

$$\sum_{x,x' \in X} \left(|N_x \cap N_{x'}| - \frac{N}{4} \right)^2 = \sum_{x,x' \in X} |N_x \cap N_{x'}|^2 - \frac{N}{2} \sum_{x,x' \in X} |N_x \cap N_{x'}| + \frac{N^4}{16}.$$

Now

$$\sum_{x,x' \in X} |N_x \cap N_{x'}| = \sum_{y \in Y} |N_y|^2 \ge N^{-1} \left(\sum_{y \in Y} |N_y| \right)^2 = \frac{N^3}{4}.$$

Hence the original sum is at most $\frac{N^4}{16} + c_1 N^4 - \frac{N^4}{8} + \frac{N^4}{16} = c_1 N^4$. So $\big||N_x \cap N_{x'}| - \frac{N}{4}\big| > c_1^{\frac{1}{3}} N$ for at most $c_1^{\frac{1}{3}} N^2$ pairs (x, x').

(iii) $\implies$ (i). If (iii) holds then

$$\begin{aligned} \sum_{x,x' \in X} |N_x \cap N_{x'}|^2 &\le N^2 \left(\frac{N}{4} + c_2 N \right)^2 + c_2 N^2 . N^2 \\ &= \frac{N^4}{16} + \left(\frac{1}{2} c_2 + c_2^2 + c_2 \right) N^4. \end{aligned}$$

(i) $\implies$ (iv). Let

$$G(x, y) = \begin{cases} 1 & xy \in E(G) \\ 0 & \text{otherwise} \end{cases}$$

and let

$$f(x, y) = G(x, y) - \frac{1}{2} = \begin{cases} \frac{1}{2} & xy \in E(G) \\ -\frac{1}{2} & \text{otherwise} \end{cases}.$$

Then the left-hand side of (i) is

$$\begin{aligned} & \sum_{x,x' \in X} \sum_{y,y' \in Y} G(x, y) G(x, y') G(x', y) G(x', y') \\ = & \sum_{x,x' \in X} \sum_{y,y' \in Y} \left(\frac{1}{2} + f(x, y) \right) \left(\frac{1}{2} + f(x, y') \right) \left(\frac{1}{2} + f(x', y) \right) \left(\frac{1}{2} + f(x', y') \right) \\ = & \frac{N^4}{16} + \sum_{x,x' \in X} \sum_{y,y' \in Y} f(x, y) f(x, y') f(x', y) f(x', y'). \end{aligned}$$

All the other 14 terms are zero. For example,

$$\begin{aligned}\sum_{x,x'\in X}\sum_{y,y'\in Y}\frac{1}{2}f(x,y')f(x',y)f(x',y') &= \frac{1}{2}\sum_{x'\in X}\sum_{y,y'\in Y}f(x',y)f(x',y')\sum_{x\in X}f(x,y')\\ &= 0\end{aligned}$$

as $\sum_{x\in X}f(x,y')=0$ for all $y'\in Y$.

So now let us assume that (iv) is false, and obtain a lower bound for $\sum_{x,x'\in X}\sum_{y,y'\in Y}f(x,y)f(x,y')f(x',y)f(x',y')$. It is

$$\begin{aligned}\sum_{x,x'\in X}\Bigl(\sum_{y\in Y}f(x,y)f(x',y)\Bigr)^2 &\geq \sum_{x,x'\in A}\Bigl(\sum_{y\in Y}f(x,y)f(x',y)\Bigr)^2\\ &= \sum_{x,x'\in A}\sum_{y,y'\in Y}f(x,y)f(x,y')f(x',y)f(x',y')\\ &= \sum_{y,y'\in Y}\Bigl(\sum_{x\in A}f(x,y)f(x,y')\Bigr)^2\\ &\geq \sum_{y,y'\in B}\Bigl(\sum_{x\in A}f(x,y)f(x,y')\Bigr)^2\\ &\geq |B|^{-2}\Bigl(\sum_{y,y'\in B}\sum_{x\in A}f(x,y)f(x,y')\Bigr)^2\\ &= |B|^{-2}\left(\sum_{x\in A}\Bigl(\sum_{y\in B}f(x,y)\Bigr)^2\right)^2\\ &\geq |B|^{-2}\left(|A|^{-1}\Bigl(\sum_{x\in A}\sum_{y\in B}f(x,y)\Bigr)^2\right)^2\\ &= |A|^{-2}|B|^{-2}\Bigl(\sum_{x\in A}\sum_{y\in B}f(x,y)\Bigr)^4\\ &= |A|^{-2}|B|^{-2}\left(|E(A,B)|-\frac{1}{2}|A||B|\right)^4\\ &\geq |A|^{-2}|B|^{-2}(c_3N^2)^4\\ &\geq c_3^4N^4.\end{aligned}$$

Hence

$$\sum_{x,x'\in X}|N_x\cap N_{x'}|^2\geq\frac{N^4}{16}+c_3^4N^4.$$

(iv) $\Longrightarrow$ (i). Choose an edge xy randomly from G and consider the number of edges from $N_y \subset X$ to $N_x \subset Y$, assuming $\sum_{x,x' \in X} |N_x \cap N_{x'}|^2 > \frac{N^4}{16} + c_1 N^4$. On average it is

$$\begin{aligned} \frac{2}{N^2} \sum_{xy \in E(G)} \sum_{\substack{y' \in N_x \\ x' \in N_y}} \mathbf{1}_{x'y' \in E(G)} &= \frac{2}{N^2} \sum_{x,x' \in X} |N_x \cap N_{x'}|^2 \\ &> \frac{N^2}{8} + 2c_1 N^2 \\ &= \frac{1}{2}|N_x||N_y| + 2c_1 N^2. \end{aligned}$$

Hence there exist x and y such that this is true. □

Lemma 7. *Let G be a bipartite graph with vertex sets X and Y of size N, with N even. Suppose that $\left|\deg(x) - \frac{N}{2}\right| \leq cN$ for all but at most cN vertices $x \in X \cup Y$. Then there is an $\frac{N}{2}$-regular graph H with $|E(H) \Delta E(G)| \leq 21cN^2$.*

Proof. We can remove fewer than $2cN^2 + cN^2 = 3cN^2$ edges to obtain a graph with all degrees at most $\frac{N}{2}$ and having at least

$$\frac{N^2}{2} - 2cN^2 - cN^2 - 3cN^2 = \frac{N^2}{2} - 6cN^2$$

edges. Now add edges for as long as possible while keeping the maximum degree at most $\frac{N}{2}$.

If we can no longer do this, then every vertex in X of degree less than $\frac{N}{2}$ is joined to every vertex in Y of degree less than $\frac{N}{2}$. Pick $x \in X$ and $y \in Y$, each of degree less than $\frac{N}{2}$. Choose $z \in X$ not joined to y. Since $\deg z = \frac{N}{2}$ (as it is not joined to y) and $\deg x < \frac{N}{2}$, we can find w joined to z but not x. Now replace the edge zw by the edges xw and zy. The degrees of w and z stay the same, while those of x and y go up by 1. We have only changed 3 edges and increased the total number of edges by 1. So at most $18cN^2$ further changes are necessary. □

Theorem 8. *Let G be an $\frac{N}{2}$-regular graph with N vertices. Then the following statements are equivalent:*

(i) $\sum_{x,x' \in X} |N_x \cap N_{x'}|^2 \leq \frac{N^4}{16} + c_1 N^4$.

(ii) The number of labelled 4-cycles is at most $\frac{N^4}{16} + c_1 N^4$.

(iii) $\left||N_x \cap N_{x'}| - \frac{N}{4}\right| \leq c_2 N$ *for all but at most* $c_2 N^2$ *pairs* (x, x').

(iv) $\left|E(A,B) - \frac{1}{2}|A||B|\right| \leq c_3 N^2$ *for all* $A, B \subset X$.

(v) $\left|E(A,A) - \frac{1}{2}|A|^2\right| \leq c_4 N^2$ *for every* $A \subset X$.

(vi) The second-largest eigenvalue of the adjacency matrix is at most $c_5 N$.

(vii) Let H *be any fixed graph with* k *vertices and let* $\phi : V(H) \to V(G)$ *be a random function. Let* E *be the event that* $\phi(x)\phi(y) \in E(G)$ *precisely when* $xy \in E(H)$. *Then* $\left|\mathbb{P}(E) - 2^{-\binom{k}{2}}\right| \leq c_6$.

Proof. (i) $\iff$ (ii) $\iff$ (iii) $\iff$ (iv) by Theorem 6. (Make G into a bipartite graph by doubling up the vertex set.) Trivially (iv) $\implies$ (v) and (vii) $\implies$ (ii).

(v) $\implies$ (iv). Let A, B be such that $\left|e(A,B) - \frac{1}{2}|A||B|\right| > c_3 N^2$. Note that

$$2e(A,B) = e(A \cup B, A \cup B) - e(A - B, A - B) - e(B - A, B - A) + e(A \cap B, A \cap B).$$

Also,

$$|A||B| = \frac{1}{2}|A \cup B|^2 - \frac{1}{2}|A - B|^2 - \frac{1}{2}|B - A|^2 + \frac{1}{2}|A \cap B|^2.$$

Hence there is some set C such that $\left|e(C,C) - \frac{1}{2}|C|^2\right| \geq \frac{c_3 N^2}{2}$.

(i) $\iff$ (vi). Suppose we write

$$G(x,y) = \sum_{i=0}^{N-1} \lambda_i u_i(x) u_i(y)$$

where u_0,u_1, …, u_{N-1} is an orthonormal basis of eigenvectors (of the adjacency matrix G) and λ_0, λ_1, …, λ_{N-1} are the eigenvalues with $\lambda_0 = \frac{N}{2}$. Then

$$\begin{aligned}
\sum_{x,y \in V(G)} G(x,y) &= \sum_{x,y \in V(G)} G(x,y)^2 \\
&= \sum_{x,y \in V(G)} \left(\sum_{i=0}^{N-1} \lambda_i u_i(x) u_i(y) \right)^2 \\
&= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \lambda_i \lambda_j \sum_{x,y \in V(G)} u_i(x) u_i(y) u_j(x) u_j(y) \\
&= \sum_{i=0}^{N-1} \lambda_i^2
\end{aligned}$$

(as $\sum_{x\in V(G)} u_i(x)u_j(x) = \delta_{ij}$). Also,

$$\begin{aligned}
&\sum_{x,y,z,w\in V(G)} G(x,y)G(y,z)G(z,w)G(w,x)\\
=\ &\sum_{0\le i,j,k,l\le N-1} \lambda_i\lambda_j\lambda_k\lambda_l \sum_{x,y,z,w\in G(v)} u_i(x)u_i(y)u_j(y)u_j(z)u_k(z)u_k(w)u_l(w)u_l(x)\\
=\ &\sum_{0\le i,j,k,l\le N-1} \lambda_i\lambda_j\lambda_k\lambda_l\langle u_i,u_l\rangle\langle u_i,u_j\rangle\langle u_j,u_k\rangle\langle u_k,u_l\rangle\\
=\ &\sum_{i=0}^{N-1}\lambda_i^4,
\end{aligned}$$

i.e.

$$\sum_{x,x'\in V(G)} |N_x\cap N_{x'}|^2 = \sum_{i=0}^{N-1}\lambda_i^4.$$

If $|\lambda_i| \le c_5N$ for every $i>0$ then

$$\sum_{i=1}^{N-1}\lambda_i^4 \le (c_5N)^2\sum_{i=0}^{N-1}\lambda_i^2 = c_5^2N^2\frac{N^2}{2} = \frac{1}{2}c_5^2N^4,$$

and so

$$\sum_{x,x'\in V(G)} |N_x\cap N_{x'}|^2 \le \frac{N^4}{16}+\frac{1}{2}c_5^2N^4.$$

Conversely, if there is some $i>0$ for which $|\lambda_i| > c_5N$, then

$$\sum_{x,x'\in V(G)} |N_x\cap N_{x'}|^2 = \sum_{i=0}^{N-1}\lambda_i^4 > \frac{N^4}{16}+c_5^4N^4.$$

(iv) $\Longleftrightarrow$ (vii). We shall, in fact, prove the following stronger result:

> Let G be a k-partite graph with vertex sets $X_1, X_2, \ldots, X_k$, all of size $N = 2^kM$, such that all the bipartite graphs spanned by (X_i, X_j), $i\ne j$, are $\frac{N}{2}$-regular and have the property that, whenever $A\subset X_i$ and $B\subset X_j$, then $\left|e(A,B)-\frac{1}{2}|A||B|\right| \le c_3N^2$. Let H be a graph with vertex set $\{1,2,\ldots,k\}$ and let $x_i\in X_i$ be chosen randomly. Then
>
> $$\left|\mathbb{P}\left[x_ix_j\in E(G) \iff ij\in E(H)\right] - 2^{-\binom{k}{2}}\right| \le c_6$$
>
> for some c_6 that (depends on k and) tends to 0 as $c_3\to 0$.

Let $\mathcal{E}$ be the event $[x_ix_j \in E(G) \iff ij \in E(H)]$. Pick x_i randomly. Then a necessary condition for $\mathcal{E}$ is that $x_i \in N_{x_1} \cap X_i$ whenever $1i \in E(H)$, and $x_i \in X_i - N_{x_1}$ whenever $1i \notin E(H)$. Let

$$A_i = \begin{cases} X_i \cap N_{x_1} & 1i \in E(H) \\ X_i - N_{x_1} & 1i \notin E(H) \end{cases} \qquad (i = 2, 3, \ldots, k).$$

Then the probability that $x_i \in A_i$, $i = 2, 3, \ldots, k$, is $2^{-(k-1)}$.

It remains to discuss $\mathbb{P}(\mathcal{E}|x_i \in A_i, i = 2, 3, \ldots, k)$. First, let us show that the graph spanned by A_i and A_j is approximately regular. If there are $c\frac{N}{2}$ vertices in A_i of degree greater than $\frac{N}{4} + c\frac{N}{2}$ in A_j then let B be the set of these bad vertices. We have

$$e(B, A_j) > \left(\frac{N}{4} + c\frac{N}{2}\right)|B| = \frac{1}{2}|B||A_j| + c\frac{N}{2}|B| \geq \frac{1}{2}|B||A_j| + \frac{c^2}{4}N^2,$$

contradicting our assumption if $c \geq 2\sqrt{c_3}$. We may proceed similarly if $c\frac{N}{2}$ have degree less than $\frac{N}{4} - c\frac{N}{2}$.

So, we can make the graph spanned by A_i and A_j become $\frac{N}{4}$-regular after changing at most $84\sqrt{c_3}\left(\frac{N}{2}\right)^2$ edges. In the modified graph, if $A \subset A_i$ and $B \subset A_j$ then

$$\left|e(A, B) - \frac{1}{2}|A||B|\right| \leq (c_3 + 21\sqrt{c_3})\, N^2 \leq 100\sqrt{c_3}\left(\frac{N}{2}\right)^2$$

(if $\sqrt{c_3} \geq c_3$).

By induction, the probability that x_ix_j is an edge of the modified graph precisely when ij is an edge of H given that $x_i \in A_i$ for $i = 2, 3, \ldots, k$ differs from $2^{-\binom{(k-1)}{2}}$ by at most c, where $c \to 0$ as $100\sqrt{c_3} \to 0$. The probability that some x_ix_j is a modified edge is at most $84\sqrt{c_3}\binom{k-1}{2} \to 0$ so $\mathbb{P}(\mathcal{E}) - 2^{-(k-1)}2^{-\binom{k-1}{2}} \to 0$ as $c_3 \to 0$. □

Lemma 9. *Let V and W be two n-dimensional real inner product spaces, and let $\alpha : V \to W$ be a linear map. Then α can be written $\sum_{i=0}^{n-1} \lambda_i v_i \otimes w_i$, where $(v_i)_{i=0}^{n-1}$ and $(w_i)_{i=0}^{n-1}$ are orthonormal bases and $v \otimes w$ denotes the rank-1 map $u \mapsto \langle u, v\rangle w$.*

Proof. Pick $v \in V$ and $w \in W$ with $\|v\| = \|w\| = 1$ so as to maximize $\langle \alpha v, w\rangle$. Clearly $w = \frac{\alpha v}{\|\alpha v\|}$ and then $\langle \alpha v, w\rangle = \|\alpha v\|$, so we are looking for v that maximizes $\frac{\|\alpha v\|}{\|v\|}$.

If we have this v and $\langle u, v\rangle = 0$, then $\|v + \delta u\|^2 = \|v\|^2 + o(\delta)$ and $\|\alpha(v + \delta u)\|^2 = \|\alpha v\|^2 + 2\delta\langle \alpha u, \alpha v\rangle + o(\delta)$, so if $\langle \alpha v, \alpha u\rangle \neq 0$ we can choose a

δ such that $\frac{\|\alpha(v+\delta u)\|}{\|v+\delta u\|} > \frac{\|\alpha v\|}{\|v\|}$, a contradiction. Therefore whenever $\langle u, v\rangle = 0$ then also $\langle \alpha u, \alpha v\rangle = 0$.

Set $v_0 = v$, $w_0 = \frac{\alpha v}{\|\alpha v\|}$, $V_1 = \langle v_0\rangle^\perp$ and $W_1 = \langle w_0\rangle^\perp$. Then $\alpha : V_1 \to W_1$, so, by induction, we can write the restriction of α to V_1 as $\sum_{i=1}^{n-1} \lambda_i v_i \otimes w_i$, where $(v_i)_{i=1}^{n-1}$ and $(w_i)_{i=1}^{n-1}$ are orthonormal and all the v_i are orthogonal to v_0 and all the w_i are orthogonal to w_0. □

It is not hard to show that if we apply this decomposition to the (bipartite) adjacency matrix of a bipartite graph G with $n+n$ vertices then the number of labelled 4-cycles (x, y, x', y') with $x, x' \in X$ is $\sum_{i=0}^{n-1} \lambda_i^4$. So, as with graphs, if G is $\frac{n}{2}$-regular, we can say that G is quasirandom iff for all $i > 0$, $|\lambda_i| \le cn$ for small c. Since $\lambda_0 = \frac{n}{2}$ and $v_0 = w_0 = \frac{1}{\sqrt{n}}(1, 1, \dots, 1)$, this says that $\sum_{xy\in E(G)} v(x)w(y) \le cn\|v\|_2\|w\|_2$ whenever $\sum_{x\in X} v(x) = \sum_{y\in Y} w(y) = 0$.

Alternatively, $G(x, y) - \frac{1}{2}$ has small correlation with any rank-1 matrix.

3 Szemerédi's Regularity Lemma

Let G be a graph, and let A, $B \subset V(G)$. The *density* of the pair (A, B) is defined to be

$$d(A, B) = \frac{|E(A, B)|}{|A||B|} = \frac{|\{(x, y) \in A \times B : xy \in E(G)\}|}{|A \times B|}.$$

The pair is said to be *ε-regular* if, whenever $A' \subset A$ and $B' \subset B$ with $|A'| \ge \varepsilon|A|$ and $|B'| \ge \varepsilon|B|$, then $|d(A', B') - d(A, B)| \le \varepsilon$. A partition X_1, X_2, …, X_k of $V(G)$ is said to be *ε-regular* if

$$\sum \left\{ \frac{|X_i||X_j|}{n^2} : (X_i, X_j) \text{ is not } \varepsilon\text{-regular} \right\} \le \varepsilon.$$

Equivalently, the partition is ε-regular if for $(x, y) \in V(G)^2$, the probability that (x, y) belongs to an irregular pair is less than ε.

Theorem 10 (Szemerédi's Regularity Lemma). *Let $\varepsilon > 0$. Then there exists $K = K(\varepsilon)$ such that for every graph G, there is an ε-regular partition of $V(G)$ into at most K sets.*

Definition. Given a partition $X_1 \cup X_2 \cup \cdots \cup X_m$ of $V(G)$, define the *mean-square density* to be

$$\sum_{i,j=1}^{m} \frac{|X_i||X_j|}{n^2} d(X_i, X_j)^2.$$

Since $\sum_{i,j=1}^{m} \frac{|X_i||X_j|}{n^2} = 1$ and $0 \le d(X_i, X_j) \le 1$, we see that the mean-square density lies between 0 and 1.

Lemma 11. *Let X_1, X_2, ..., X_m be a partition, and let Y_1, Y_2, ..., Y_M be a partition that refines X_1, X_2, ..., X_m. Then the mean-square density with respect to Y_1, Y_2, ..., Y_M is at least as big as that with respect to X_1, X_2, ..., X_m.*

Proof. Suppose each X_i is split into $X_{i1} \cup X_{i2} \cup \cdots \cup X_{ir_i}$. Then

$$\begin{aligned} d(X_i, X_j)^2 &= \left(\sum_{s,t} \frac{|X_{is}||X_{jt}| d(X_{is}, X_{jt})}{|X_i||X_j|} \right)^2 \\ &\le \left(\sum_{s,t} \frac{|X_{is}||X_{jt}|}{|X_i||X_j|} \right) \left(\sum_{s,t} \frac{|X_{is}||X_{jt}|}{|X_i||X_j|} d(X_{is}, X_{jt})^2 \right). \end{aligned}$$

Hence

$$\frac{|X_i||X_j|}{n^2} d(X_i, X_j)^2 \le \sum_{s,t} \frac{|X_{is}||X_{jt}|}{n^2} d(X_{is}, X_{jt})^2.$$

□

Remark. The same proof shows (or the lemma itself implies) a similar statement for bipartite graphs: if G has vertex sets X and Y partitioned into $X_1 \cup X_2 \cup \cdots \cup X_r$ and $Y_1 \cup Y_2 \cup \cdots \cup Y_s$, and $Z_1 \cup Z_2 \cup \cdots \cup Z_R$ and $W_1 \cup W_2 \cup \cdots \cup W_S$ refine $X_1 \cup X_2 \cup \cdots \cup X_r$ and $Y_1 \cup Y_2 \cup \cdots \cup Y_s$ respectively, then

$$\sum_{i=1}^{r} \sum_{j=1}^{s} \frac{|X_i||X_j|}{|X||Y|} d(X_i, X_j)^2 \le \sum_{i=1}^{R} \sum_{j=1}^{S} \frac{|Z_i||W_j|}{|X||Y|} d(Z_i, W_j)^2.$$

Lemma 12. *Let (X, Y) be some pair of sets of vertices in a graph G, and suppose that $d(X, Y) = \alpha$ and (X, Y) is not ε-regular. Then there are partitions $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$ such that*

$$\sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 \ge \alpha^2 + \varepsilon^4.$$

Proof. By the non-ε-regularity, we can find $X_1 \subset X$ and $Y_1 \subset Y$ such that $|d(X_1, Y_1) - \alpha| \geq \varepsilon$ and $|X_1| \geq \varepsilon|X|$, $|Y_1| \geq \varepsilon|Y|$. Let $u(X_i, Y_i) = d(X_i, Y_i) - \alpha$. Then

$$\begin{aligned}
\varepsilon^4 &\leq \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} u(X_i, Y_j)^2 \\
&= \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 - 2\alpha \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j) + \alpha^2 \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} \\
&= \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 - \alpha^2.
\end{aligned}$$

□

Lemma 13. *Let G be a graph with N vertices, and let $X_1 \cup X_2 \cup \cdots \cup X_m$ be a partition of the vertices that is not ε-regular. Then there is a refinement $X_{11} \cup \cdots \cup X_{1r_1} \cup X_{21} \cup \cdots \cup X_{2r_2} \cup \cdots \cup X_{m1} \cup \cdots \cup X_{mr_m}$ such that each r_i is at most 2^{2m} and the mean-square density is bigger by at least ε^5.*

Proof. Let $I = \{(i,j) : (X_i, X_j) \text{ is not } \varepsilon\text{-regular}\}$. Let α^2 be the mean-square density of G with respect to $X_1 \cup X_2 \cup \cdots \cup X_m$.

For each $(i,j) \in I$, Lemma 12 gives us partitions $X_i = A_1^{ij} \cup A_2^{ij}$ and $X_j = B_1^{ij} \cup B_2^{ij}$ such that

$$\sum_{p,q=1}^{2} \frac{|A_p^{ij}||B_q^{ij}|}{|X_i||X_j|} d(A_p^{ij}, B_q^{ij})^2 \geq d(X_i, X_j)^2 + \varepsilon^4.$$

Now, for each i, let $X_{i1} \cup X_{i2} \cup \cdots \cup X_{ir_i}$ be a partition of X_i into at most 2^{2m} sets such that every A_1^{ij}, A_2^{ij}, B_1^{ij} and B_2^{ij} is a union of some of the X_{ih}. Then, by Lemma 11 (and the remark after it),

$$\sum_{p=1}^{r_i} \sum_{q=1}^{r_j} \frac{|X_{ip}||X_{jq}|}{|X_i||X_j|} d(X_{ip}, X_{jq})^2 \geq d(X_i, X_j)^2 + \varepsilon^4$$

for all $(i,j) \in I$. Multiplying both side by $\frac{|X_i||X_j|}{N^2}$ and summing over all (i,j), splitting into a sum over I and a sum over I^c, we have

$$\begin{aligned}
\sum_{i,j=1}^{m} \sum_{p=1}^{r_i} \sum_{q=1}^{r_j} \frac{|X_{ip}||X_{jq}|}{N^2} d(X_{ip}, X_{jq})^2 &\geq \sum_{i,j=1}^{m} \frac{|X_i||X_j|}{N^2} d(X_i, X_j)^2 + \varepsilon^4 \sum_{(i,j)\in I} \frac{|X_i||X_j|}{N^2} \\
&\geq \alpha^2 + \varepsilon^5.
\end{aligned}$$

□

Proof (of Theorem 10). Start with the trivial partition of $V(G)$ into one set. If it is ε-regular then we are done. If not, refine it into at most four sets in such a way that mean-square density increases by ε^5.

Continue this process. If at stage k we have a partition into m sets then at stage $k+1$ we have one into at most $m.2^{2m} \leq 2^{2^m}$. Hence the process must stop at an ε-regular partition after at most ε^{-5} steps. The number of cells in the final partition is therefore at most $2^{2^{2^{\cdot^{\cdot^{\cdot^2}}}}\Big\}2\varepsilon^{-5}}$. □

Lemma 14. *Let G be a graph, and let X, Y, $Z \subset V(G)$. Suppose that (X,Y), (Y,Z) and (X,Z) are ε-regular and that $d(X,Y) = \alpha$, $d(Y,Z) = \beta$ and $d(X,Z) = \gamma$. Then the number of $(x,y,z) \in X \times Y \times Z$ forming triangles in G is at least $(1-2\varepsilon)(\alpha-\varepsilon)(\beta-\varepsilon)(\gamma-\varepsilon)|X||Y||Z|$, provided $\alpha, \beta, \gamma \geq 2\varepsilon$.*

Proof. For each $x \in X$, write $d_Y(x)$ for $|N_x \cap Y|$ and $d_Z(x)$ for $|N_x \cap Z|$. Then the number of $x \in X$ such that $d_Y(x) < (\alpha-\varepsilon)|Y|$ is at most $\varepsilon|X|$ by ε-regularity of (X,Y). Similarly, at most $\varepsilon|X|$ of the $x \in X$ have $d_Z(x) < (\beta-\varepsilon)|Z|$. If $d_Y(x) \geq (\alpha-\varepsilon)|Y|$ and $d_Z(x) \geq (\beta-\varepsilon)|Z|$ then, by ε-regularity of (Y,Z), the number of edges between $N_x \cap Y$ and $N_x \cap Z$ is at least $(\gamma-\varepsilon)(\alpha-\varepsilon)(\beta-\varepsilon)|Y||Z|$. Summing over $x \in X$ gives the result. □

Theorem 15. *For all $\varepsilon > 0$, there exists some $\delta > 0$ with the following property: given any graph G with n vertices and at most δn^3 triangles, it is possible to remove at most εn^2 edges from G to make it triangle-free.*

Proof. Let $X_1 \cup X_2 \cup \cdots \cup X_M$ be an $\frac{\varepsilon}{4}$-regular partition with $M \leq M(\varepsilon)$. From G, remove the edge xy if

(i) $(x,y) \in X_i \times X_j$ with (X_i, X_j) not an $\frac{\varepsilon}{4}$-regular pair; or

(ii) $(x,y) \in X_i \times X_j$ with $d(X_i, X_j) < \frac{\varepsilon}{2}$; or

(iii) $x \in X_i$ with $|X_i| \leq \frac{\varepsilon n}{4M}$.

The number of edges removed by (i) is at most $\sum_{(i,j)\in I} |X_i||X_j| \leq \frac{\varepsilon n^2}{4}$, since the partition is $\frac{\varepsilon}{4}$-regular. The number removed by (ii) is at most $\sum_{i,j} \frac{\varepsilon}{2}|X_i||X_j| = \frac{\varepsilon n^2}{2}$. The number removed by (iii) is at most $Mn\frac{\varepsilon n}{4M} = \frac{\varepsilon n^2}{4}$.

Now suppose that, after these edges have been removed, there is still a triangle $(x,y,z) \in X_i \times X_j \times X_k$, say. Then the pairs (X_i, X_j), (X_j, X_k) and (X_i, X_k) are all $\frac{\varepsilon}{4}$-regular with density at least $\frac{\varepsilon}{2}$ (or we would have removed these edges). Also $|X_i|, |X_j|, |X_k| > \frac{\varepsilon n}{4M}$. By Lemma 14, G contains at least $\left(1-\frac{\varepsilon}{2}\right)\left(\frac{\varepsilon}{4}\right)^3\left(\frac{\varepsilon N}{4M}\right)^3$ labelled triangles. So let $\delta = \frac{\varepsilon^6}{2^{20}M^3}$ and the result is proved. □

Theorem 16. *Let $\delta > 0$. Then there exists N_0 such that if $N \geq N_0$ and $A \subset [N]^2$ with $|A| \geq \delta N^2$ then A contains a triple of the form (x, y), $(x+d, y)$, $(x, y+d)$ with $d > 0$.*

Proof. The set $A + A = \{\mathbf{x} + \mathbf{y} : \mathbf{x}, \mathbf{y} \in A\}$ is contained in $[2N]^2$ so there must exists $\mathbf{z}$ that can be written as $\mathbf{x} + \mathbf{y}$ in at least $\frac{(\delta N^2)^2}{(2N)^2} = \frac{\delta^2 N^2}{4}$ ways. Pick such a $\mathbf{z}$ and write $A' = A \cap (\mathbf{z} - A)$ and $\delta' = \frac{\delta^2}{4}$. Then A' has size at least $\delta' N^2$ and has the property that if A' contains a triple of the form (x, y), $(x + d, y)$, $(x, y + d)$ for $d < 0$ then so does $\mathbf{z} - A$, and hence A contains such a triple with $d > 0$. So, replacing A by A' and δ by δ', we may assume that A contains no such triple for $d \neq 0$.

Now, we construct a tripartite graph as follows. The vertex sets are $X = [N]$, $Y = [N]$ and $Z = [2N]$. Think of X as the set of vertical lines through A, Y as the set of horizontal lines, and Z as the set of diagonal lines with $x + y$ constant. Join $x \in X$ to $y \in Y$ iff $(x, y) \in A$, join $x \in X$ to $z \in Z$ iff $(x, z - x) \in A$ and join $y \in Y$ to $z \in Z$ iff $(z - y, y) \in A$ (i.e. join two lines iff their intersection is in A).

Suppose G contains a triangle (x, y, z). Then (x, y), $(x, y + (z - x - y))$ and $(x + (z - x - y), y)$ are all in A, which is a contradiction unless $x + y = z$ (which is the degenerate case where the three lines meet in a point). So G has at most $N^2 = \frac{1}{64N}(4N)^3$ triangles. So for sufficiently large N, we can remove at most $\frac{\delta N^2}{2}$ edges from G to make it triangle-free. However, the degenerate triangles are edge-disjoint (since any edge in G determines a unique point of A) and there are at least δN^2 of them. So we have a contradiction. □

Corollary 17. *Let $\delta > 0$. Then if N is sufficiently large and $A \subset [N]$ is any set of size at least δN, A contains an arithmetic progression of length 3.*

Proof. Define a subset $B \subset [2N]^2$ to be $\{(x, y) : y - x \in A\}$. Then $|B| \geq \delta N^2$, so by Theorem 16 we can find x, y, and $d \neq 0$ such that $y - (x + d)$, $y - x$ and $y + d - x$ all belong to A. □

Theorem 18 (Erdős-Stone Theorem). *Let H be a graph with chromatic number k. Then a graph G with n vertices that contains no copy of H has at most $\left(1 - \frac{1}{k-1}\right)\binom{n}{2}(1 + o(1))$ edges (and this is best possible).*

Proof (sketch). The example of a complete $(k-1)$-partite graph with vertex subsets of roughly equal size shows that the estimate cannot be substantially improved.

Conversely, suppose that G contains no copy of H. Let $X_1, X_2, \ldots, X_M$ be an ε-regular partition of $V(G)$ for some small $\varepsilon > 0$. Remove edges xy if they belong to sparse or irregular pairs $X_i \times X_j$, or pairs with X_i or X_j small. Not too many edges are removed. Let G' be G with these edges removed.

If G' contains a K_k then, by a generalization of the lemma counting triangles, the resulting k-partite subgraph contains several copies of H.

(The resulting lemma says that if you randomly embed $V(H)$ into a k-partite graph with all pairs regular, respecting the k-partition of $V(H)$, then the probability that $\phi(x)\phi(y) \in E(G)$ whenever $xy \in E(H)$ is what it would be for random graphs of the appropriate density).

So G' contains no K_k, and so by Turán's theorem has at most $\left(1 - \frac{1}{k-1}\right)\binom{n}{2}$ edges. □

4 Crossing numbers and combinatorial geometry

Theorem 19. *A planar graph with $n \geq 3$ vertices has at most $3n - 6$ edges.*

Proof. Assume wlog G is maximal planar (and so, in particular, is connected). Then Euler's theorem says that $V - E + F = 2$. It is easy to see that every face has at least 3 edges, so the number of edge-face pairs is equal to $2E$ and at least $3F$. So $F \leq \frac{2}{3}E$, giving $V - \frac{E}{3} \geq 2$ and so $E \leq 3V - 6$. □

Lemma 20. *Let G be a graph with n vertices and $m \geq 4n$ edges. Then any drawing of G in $\mathbb{R}^2$ must have at least $\frac{m^3}{64n^2}$ crossings.*

Proof. First, note that the number of crossings is at least $m - (3n - 6)$. To see this, let G have $m > 3n - 6$ edges. Remove an edge involved in a crossing (which exists by Theorem 19). The number of crossings goes down by at least 1. We can do this $m - (3n - 6)$ times.

Now let $0 < p \leq 1$ and choose vertices independently and randomly with probability p, forming an induced subgraph H. Suppose that G has been drawn with t crossings. Then the expected number of vertices of H is pn, the expected number of edges is p^2m and the expected number of crossings is p^4t. Also, the expected number of crossings is at least $p^2m - 3pn$. So choose p such that $p^2m = 4pn$, i.e. take $p = \frac{4n}{m}$ (which by assumption is at most 1). Then $p^4t \geq pn$, giving $t \geq \frac{n}{p^3} = \frac{m^3}{64n^2}$. □

Definition. A *pseudoline system* in $\mathbb{R}^2$ is a collection Λ of curves l such that any two intersect in at most one point.

Theorem 21 (Szemerédi-Trotter Theorem). *Let $X \subset \mathbb{R}^2$ be a set of n points, and let Λ be a set of m pseudolines. Then the number of pairs (x, l) such that $x \in X$, $l \in \Lambda$ and $x \in l$, i.e. the number of* incidences*, is at most $8(m^{\frac{2}{3}}n^{\frac{2}{3}} + m + n)$.*

Proof. Suppose that there are t incidences. Define a graph G with vertex set X by joining x to y iff they are adjacent along some $l \in \Lambda$ (i.e. regard Λ as a drawing of G). Then G has at least $t - m$ edges. If $t \geq 2m$ then this is at least $\frac{t}{2}$. If $t \geq 2m$ and $\frac{t}{2} \geq 4n$ then there are at least $\frac{t^3}{512n^2}$ crossings by Lemma 20. But since Λ is a pseudoline system, there are at most $\binom{m}{2} \leq m^2$ crossings. So $t^3 \leq 512n^2m^2$, and so $t \leq 8m^{\frac{2}{3}}n^{\frac{2}{3}}$.

We have shown $t \geq \max\{2m, 8n\} \implies t \leq 8m^{\frac{2}{3}}n^{\frac{2}{3}}$. □

Corollary 22. *Let X be a set of n points, and Λ a pseudoline system with each $l \in \Lambda$ containing at least $k \geq 2$ points of X. Then $|\Lambda| \leq C \max\left\{\frac{n^2}{k^3}, \frac{n}{k}\right\}$.*

Proof. Let $|\Lambda| = m$. Then the number of incidences is at least mk. So $mk \leq 8(m^{\frac{2}{3}}n^{\frac{2}{3}} + m + n) \leq 16(m^{\frac{2}{3}}n^{\frac{2}{3}} + n)$. Hence either $m^{\frac{2}{3}}n^{\frac{2}{3}} \geq cmk$ or $n \geq cmk$, giving $m \leq C \max\left\{\frac{n^2}{k^3}, \frac{n}{k}\right\}$. □

Lemma 23. *Let $X \subset \mathbb{R}^2$ be a set of size n, let $A \subset \mathbb{R}$ be a set of size r and let f be a strictly concave or strictly convex function defined on some interval that contains A. Let $Z = \{(a, f(a)) : a \in A\} \subset \mathbb{R}^2$. Then*

$$|X + Z| \geq c \min\{n^{\frac{1}{2}}r^{\frac{3}{2}}, nr\}.$$

Proof. Let $N = |X + Z|$ and let Γ be the graph of f. Then the curves $x + \Gamma$ ($x \in X$) from a pseudoline system, by the convexity/concavity of f. It has size n. Each $x \in \Gamma$ intersects $X + Z$ in at least r points. Hence

$$n \leq C \max\left\{\frac{N^2}{r^3}, \frac{N}{r}\right\}$$

and so

$$N \geq c \min\{r^{\frac{3}{2}}n^{\frac{1}{2}}, rn\}.$$

□

Corollary 24. *Let A, B, $C \subset \mathbb{R}$ be sets of size n and let f be as above. Then $|A + B||f(A) + C| \geq cn^{\frac{5}{2}}$.*

Proof. We note that $B \times C + \{(a, f(a)) : a \in A\} \subset (A + B) \times (f(A) + C)$ and so, by Lemma 23, $|A + B||f(A) + C| \geq c \min\{nn^{\frac{3}{2}}, n^2n\} = cn^{\frac{5}{2}}$. □

Corollary 25. *Let A be a set of size n in $\mathbb{R}$, and let f be a strictly convex or strictly concave function. Then*

(i) either $|A+A| \geq cn^{\frac{5}{4}}$ or $|f(A)+f(A)| \geq cn^{\frac{5}{4}}$;

(ii) $|A+f(A)| \geq cn^{\frac{5}{4}}$;

(iii) either $|A+A| \geq cn^{\frac{5}{4}}$ or $|A.A| \geq cn^{\frac{5}{4}}$.

Proof. (i) Setting $B = A$ and $C = f(A)$ in Corollary 24, and noting that $|f(A)| \geq \frac{n}{2}$, we observe that $|A+A||f(A)+f(A)| \geq cn^{\frac{5}{2}}$ and the result follows immediately.

(ii) Set $B = f(A)$ and $C = A$.

(iii) We may assume wlog that half the elements of A are positive. Let $f(x) = \log x$. Since $|\log A + \log A| = |A.A|$, (iii) follows from (i) applied to the set $\{a \in A : a > 0\}$. □

Theorem 26. *Let X be a set of n points in $\mathbb{R}^2$. Then the number of pairs $(x,y) \in X^2$ with $d(x,y) = 1$ is at most $Cn^{\frac{4}{3}}$.*

Proof. Define a multigraph G with loops by drawing around each $x \in X$ a unit circle and, for y, $z \in X$, joining y to z if they are adjacent along one of the circles drawn. If the number of unit distances is t then the number of edges is t.

Remove all circles that contain at most two points of X. So we've removed at most $2n$ edges from G. Now any two vertices are joined by at most two edges. If they are joined by precisely two then remove one of them. This removes at most half the edges.

So if $t \geq 4n$ then the resulting simple graph has at least $\frac{t}{4}$ edges. If $\frac{t}{4} \geq 4n$ then by Lemma 20 there are at least $\frac{(t/4)^3}{64n^2} = \frac{t^3}{4096n^2}$ crossings. But two circles cross in at most two points so the number of crossings is at most $2\binom{n}{2} \leq n^2$. So $t^3 \leq 4096n^4$ and so $t \leq 16n^{\frac{4}{3}}$. □

Theorem 27. *Let G be a multigraph with n vertices, m edges and maximum edge-multiplicity at most d. If $m \geq 32nd$ then the number of crossings is at least $\frac{cm^3}{n^2 d}$.*

Proof. For $i = 1, 2, \ldots, \lfloor \log_2 d + 1 \rfloor$, let G_i be the multigraph consisting of all edges with multiplicities r such that $2^{i-1} \leq r < 2^i$. For each i, let the number of pairs xy that are joined in G_i be m_i. The number of edges in multigraphs G_i such that $m_i < 4n$ is at most $\sum_{i=1}^{\lfloor \log_2 d \rfloor + 1} 4n.2^i \leq 16nd$. So if we remove these then we still have at least $\frac{m}{2}$ edges.

Let $B = \{i : \text{at least } 4n \text{ pairs are joined in } G_i\}$. If $i \in B$ then I claim that the number of crossings in G_i is at least $\frac{2^{2(i-1)}m_i^3}{64n^2}$. The reason is that if we choose one edge at random from the set joining x to y (when they are joined) then the probability that any given crossing survives is at most $2^{-2(i-1)}$, but the number of crossings left is at least $\frac{m_i^3}{64n^2}$ by Theorem 20. So the number of crossings in G is at least $\sum_{i\in B}\frac{2^{2(i-1)}m_i^3}{n^2}$. But

$$\begin{aligned}\frac{m}{2} \le \sum_{i\in B} m_i 2^i = \sum_{i\in B} 2^{\frac{2i}{3}} m_i 2^{\frac{i}{3}} &\le \left(\sum_{i\in B} 2^{2i} m_i^3\right)^{\frac{1}{3}} \left(\sum_{i\in B} 2^{\frac{i}{2}}\right)^{\frac{2}{3}} \\ &\le C\left(\sum_{i\in B} 2^{2(i-1)} m_i^3\right)^{\frac{1}{3}} d^{\frac{1}{3}}.\end{aligned}$$

Hence

$$\sum_{i\in B} \frac{2^{2(i-1)}m_i^3}{n^2} \ge \frac{cm^3}{n^2 d}.$$

□

Theorem 28. *Let X be a set of n points in $\mathbb{R}^2$. Then X determines at least $cn^{\frac{4}{5}}$ distinct distances.*

Proof. Suppose that the number of distinct distances is t, and that these distances are $r_1, r_2, \ldots, r_t$. About each $x \in X$, draw circles of radius $r_1, r_2, \ldots, r_t$. Define a multigraph by joining $y, z \in X$ iff they are adjacent along one of the circles drawn (once per arc with this property). We may assume throughout the proof that n is large.

The number of edges is $n(n-1) \ge \frac{99}{100}n^2$. If $t < n^{\frac{4}{5}}$ then the number of edges from circles with at most two points is at most $2nt \le \frac{n^2}{100}$, so remove these. Let k be an integer to be chosen (it will be $n^{\frac{2}{5}}$). We now wish to bound from above the number of edges of multiplicity at least k.

The perpendicular bisector of such an edge (considered geometrically) is a line that contains at least k points of X. So we shall bound from above the number of pairs (e, l) such that e is an edge, l bisects e, and l contains $u \ge k$ points of X. The number of lines l with $2^{i-1} \le u < 2^i$ is at most $C\max\{\frac{n^2}{2^{3i}}, \frac{n}{2^i}\}$ by Corollary 22. Each such line is involved in at most $2t.2^i$ pairs (e, l), so the total number is at most

$$C\left(\sum_{k\le 2^i\le\sqrt{n}} t\frac{n^2}{2^{3i}}2^i + \sum_{\sqrt{n}\le 2^i\le n} \frac{tn2^i}{2^i}\right) \le C\left(t\frac{n^2}{k^2} + tn\log n\right).$$

We want this to be less than $\frac{n^2}{4}$, which is true if $t \le n^{\frac{4}{5}}$ and $t \le ck^2$.

After removing all edges of multiplicities at least k, we still have at least $\frac{n^2}{2}$ edges left, so by Theorem 27, there are at most $\frac{cn^6}{n^2k} = \frac{cn^4}{k}$ crossings. But the number of crossings is also at most n^2t^2, so $t^2 \geq \frac{cn^2}{k}$, and so $t \geq \frac{cn}{\sqrt{k}}$. So we have shown that $t \geq c \min\left\{k^2, \frac{n}{\sqrt{k}}\right\}$. If we choose $k = n^{\frac{2}{5}}$, we get the result. □

Theorem 29 (Beck's two-extremities theorem). *Let X be a set of n points in $\mathbb{R}^2$. Then either there are at least cn^2 lines determined by pairs of points in X, or some line contains at least cn points.*

Proof. The number of lines containing between between 2^i and 2^{i+1} points is at most $C \max\left\{\frac{n^2}{2^{3i}}, \frac{n}{2^i}\right\}$ by Corollary 22. Any line with this property contains at most 2^{2i+1} pairs of points of X. So the number of pairs belonging to lines that contain between A and $\frac{n}{A}$ points of X is at most $C' \sum_{A \leq 2^i \leq \frac{n}{A}} \max\left\{\frac{n^2}{2^i}, 2^i n\right\} \leq \frac{C''n^2}{A}$ (by consideration of two geometric sums). So choosing $A = 2C''$, there must be at least $\frac{n^2}{2}$ pairs belonging to lines with at least $\frac{n}{A}$ or at most A points.

If any line contains at least $\frac{n}{A}$ points then we are done.

Otherwise, at least $\frac{n^2}{2}$ pairs belong to lines with at most A^2 pairs, so there must be at least $\frac{n^2}{2A^2}$ lines. □

Corollary 30. *Let X be any set of n points in $\mathbb{R}^2$, not all collinear. Then there must be some point $x_0 \in X$ such that there are at least cn different lines containing x_0 and some other point of X.*

Proof. If cn points lie on a line, let x_0 be some point not on the line. Otherwise, there are at least cn^2 lines so some point must be on at least cn of them. □

Lemma 31. *Let A, $B \subset \mathbb{R}$ be sets of size n. Then the number of collinear triples in $A \times B$ is at most $Cn^4 \log n$.*

Proof. By Corollary 22, the number of lines with r points of $A \times B$ for some r with $2^i \leq r < 2^{i+1}$ is at most $\frac{Cn^4}{2^{3i}}$ (since $\frac{n^2}{2^i} \leq \frac{n^4}{(2^i)^3}$ as $r \leq n$). Such a line contains at most $C2^{3i}$ triples. So the number of collinear triples is at most $C \sum_{2^i \leq n} \frac{n^4}{2^{3i}} 2^{3i} = Cn^4 \log n$. □

Theorem 32. *Let $A \subset \mathbb{R}$ be a set of size n. Then $|A + A|^4 |A.A| \geq \frac{cn^6}{\log n}$.*

Proof. Let $|A.A| = p$ and $|A + A| = s$. Then $A \times A$ is contained in the union of the p sets $X_\lambda = \{(x, y) : xy = \lambda\}$ for $\lambda \in A.A$. So $\sum_{\lambda \in A.A} |X_\lambda| \geq n^2$,

and, by Cauchy-Schwarz, $\sum_{\lambda \in A.A} |X_\lambda|^2 \geq \frac{n^4}{p}$, so there are at least $p^{-1}n^4$ pairs $((a', b'), (a'', b''))$ such that $a'b' = a''b''$ (i.e. belonging to the same X_λ). For every $(a, b) \in A \times A$ and every such pair, we get a collinear triple $((a, b), (a + a', b + b''), (a + a'', b + b'))$ in $(A \cup (A + A)) \times (A \cup (A + A))$. By Lemma 31 (and the fact that $|A| \leq |A + A| \leq |A|^2$), $n^2 p^{-1} n^4 \leq C s^4 \log n$, and so $ps^4 \geq \frac{cn^6}{\log n}$. □

Example. Let $k \leq \sqrt{n}$. Let $X = \{1, 2, \ldots, k\} \times \left\{1, 2, \ldots, \frac{n}{k}\right\}$. Then any line $y = mx + b$ with $m, b \in \mathbb{N}$, $b \leq \frac{n}{2k}$ and $m \leq \frac{n}{2k^2}$ goes through k points of X, and there are $\frac{cn^2}{k^3}$ of them. On the other hand, if $k \geq \sqrt{n}$ then pick any $\frac{n}{k}$ lines and put k points on each. This shows that Corollary 22 is tight.

5 Monotone Circuit Complexity

A *circuit* is a directed, acyclic graph with different kinds of vertices called *inputs, outputs, AND gates, OR gates,* and *NOT gates* with the following properties:

(i) x is an input iff it has in-degree zero;

(ii) x is an output iff it has out-degree zero;

(iii) x is a NOT gate iff it has in-degree 1;

(iv) x may be both an output and a gate.

Let f be a function from the set I of inputs of some circuit C to the set $\{0, 1\}$. Then there is a unique extension g of f to $V(C)$ such that

(i) if x is an input then $g(x) = f(x)$;

(ii) if x is an AND gate then $g(x) = 1$ iff $g(y) = 1$ for every predecessor y of x;

(iii) if x is an OR gate then $g(x) = 1$ iff $g(y) = 1$ for some predecessor y of x;

(iv) if x is a NOT gate then $g(x) = 1$ iff $g(y) = 0$ for the unique predecessor y of x.

We say that the circuit C computes the function $\phi : \{0,1\}^I \to \{0,1\}^O$, where O is the set of output vertices, if $\phi(f) = g|_O$ for the function g just defined.

A circuit is *binary* if the in-degree of every vertex is at most 2. It is *monotone* if there are no NOT gates. It is easy to prove that if C is a monotone circuit then the function ϕ it computes is montotone, in the sense that if $f \leq g$ (i.e. $f(x) < g(x)$ for all x) then $\phi(f) \leq \phi(g)$.

Proposition 33. *There is a function $\phi : \{0,1\}^n \to \{0,1\}$ that cannot be computed by a binary circuit of size less than $\frac{2^n}{100n}$.*

Proof. The number of binary circuits with m vertices is at most $(10m^2)^m$. For these to compute all functions from $\{0,1\}^n$ to $\{0,1\}$, we need $(10m^2)^m \geq 2^{2^n}$ which implies $m(2\log_2 m + \log_2 10) \geq 2^n$, giving $m \geq \frac{2^n}{100n}$. □

Given a circuit C, let us label the input vertices $x_1, x_2, \ldots, x_n$ and the remaining vertices $x_{n+1}, x_{n+2}, \ldots, x_m$ in such a way that each x_i comes after its predecessors in the circuit. For each i, let

$$A_i = \big\{f : \{x_1, x_2, \ldots, x_n\} \to \{0,1\} \,:\, g(x_i) = 1 \text{ for } g \text{ as defined earlier}\big\}.$$

Then for $1 \leq i \leq n$, $A_i = \{f : f(x_i) = 1\} = E_i$ (where E stands for "elementary"). We can think of E_i as a subset of $\{0,1\}^n$—consisting of all points with i-coordinate 1.

If x_i is an AND (OR) gate with predecessors x_j and x_k then $A_i = A_j \cap A_k$ ($A_j \cup A_k$). If x_i is a NOT gate with predecessor x_j then $A_i = A_j^c$. Hence $\phi : \{0,1\}^n \to \{0,1\}$ can be computed by a circuit C of size m iff there is a sequence $A_1, A_2, \ldots, A_m$ of sets with the following properties:

(i) if $1 \leq i \leq n$ then $A_i = E_i = \{f : f(i) = 1\}$;

(ii) if $i > n$ then $A_i = A_j \cap A_k$, $A_i = A_j \cup A_k$ or $A_i = A_j^c$ for some $j, k < i$;

(iii) $A_m = A(\phi) = \{f : \phi(f) = 1\}$.

Razborov's method of approximations

The basic idea of the proof to follow, that a certain monotone function ϕ has high monotone circuit complexity, is this:

We shall define a lattice $\mathcal{L}$ of subsets of $\{0,1\}^n$ with operations $\sqcap$ and $\sqcup$ that approximate $\cap$ and $\cup$ well enough that whenever we take a sequence $A_1, A_2, \ldots, A_m$ as above and define a new sequence $B_1, B_2, \ldots, B_m$ by using $\sqcap$ and $\sqcup$ instead of $\cap$ and $\cup$ then B_m is close to A_m (if m is small). It follows

that every function with small monotone circuit complexity is approximated by some function in $\mathcal{L}$. We then need to show that $\{f : \phi(f) = 1\}$ is not approximated by a function in $\mathcal{L}$.

Notation. If B, $B' \in \mathcal{L}$, define $\delta_\sqcup(B, B')$ to be $B \sqcup B' - (B \cup B')$ and $\delta_\sqcap(B, B')$ to be $B \cap B' - (B \sqcap B')$.

Let $\mathcal{L}$ be a lattice of subsets of $\{0,1\}^n$ with operations $\sqcap$ and $\sqcup$ and suppose that the elementary sets $E_i = \{\varepsilon \in \{0,1\}^n : \varepsilon_i = 1\}$ are contained in $\mathcal{L}$. Suppose that $A \sqcup B \supset A \cup B$ and $A \sqcap B \subset A \cap B$ for every A, $B \in \mathcal{L}$.

Now let us suppose that A_1, A_2, ..., A_m is a *monotone straight-line calculation* of $A_m = A$, i.e. $A_i = E_i$ if $i \leq n$, and $A_i = A_j \cup A_k$ or $A_i = A_j \cap A_k$ for some j, $k < i$ if $i > n$. Let B_1, B_2, ..., B_m be defined by $B_i = A_i$ if $i \leq n$, and $B_i = B_j \sqcup B_k$ or $B_i = B_j \sqcap B_k$ according as $A_i = A_j \cup A_k$ or $A_i = A_j \cap A_k$ for $i > n$. Write M_i, N_i for the two sets that make B_i.

Lemma 34. *Under these circumstances,*

$$B_m \subset A_m \cup \bigcup_{i=1}^{m} \delta_\sqcup(M_i, N_i)$$

and

$$B_m \supset A_m - \left(\bigcup_{i=1}^{m} \delta_\sqcap(M_i, N_i) \right).$$

Proof. The proof is by induction on m.

If $B_m = B_j \sqcup B_k$ then

$$\begin{aligned} B_m &= B_j \cup B_k \cup \delta_\sqcup(B_j, B_k) \\ &\subset A_j \cup \bigcup_{i=1}^{j} \delta_\sqcup(M_i, N_i) \cup A_k \cup \bigcup_{i=1}^{k} \delta_\sqcup(M_i, N_i) \cup \delta_\sqcup(M_m, N_m) \\ &\subset A_m \cup \bigcup_{i=1}^{m} \delta_\sqcup(M_i, N_i) \end{aligned}$$

(as $A_j \cup A_k = A_m$) and

$$\begin{aligned} B_m &\supset B_j \cup B_k \\ &\supset \left(A_j - \left(\bigcup_{i=1}^{j} \delta_\sqcap(M_i, N_i) \right) \right) \cup \left(A_k - \left(\bigcup_{i=1}^{k} \delta_\sqcap(M_i, N_i) \right) \right) \\ &\supset A_m - \left(\bigcup_{i=1}^{m} \delta_\sqcap(M_i, N_i) \right). \end{aligned}$$

Similarly, if $B_m = B_j \sqcap B_k$ then

$$\begin{aligned} B_m &= B_j \cap B_k - \delta_\sqcap(B_j, B_k) \\ &\supset \left(A_j - \left(\bigcup_{i=1}^{j} \delta_\sqcap(M_i, N_i)\right)\right) \cap \left(A_k - \left(\bigcup_{i=1}^{k} \delta_\sqcap(M_i, N_i)\right)\right) - \delta_\sqcap(B_j, B_k) \\ &\supset A_m - \left(\bigcup_{i=1}^{m} \delta_\sqcap(M_i, N_i)\right) \end{aligned}$$

and

$$\begin{aligned} B_m &\subset B_j \cap B_k \\ &\subset \left(A_j \cup \bigcup_{i=1}^{j} \delta_\sqcup(M_i, N_i)\right) \cap \left(A_k \cup \bigcup_{i=1}^{k} \delta_\sqcup(M_i, N_i)\right) \\ &\subset A_m \cup \bigcup_{i=1}^{m} \delta_\sqcup(M_i, N_i). \end{aligned}$$

□

Definition of $\mathcal{L}$

Let W, W_1, W_2, …, W_r be sets. We shall say that W_1, W_2, …, W_r *entail* W, and write W_1, W_2, …, $W_r \vdash W$ if $W_i \cap W_j \subset W$ for all $i \neq j$. (The sets W_1, W_2, …, W_r are not necessarily distinct.)

Write $[n]^{(\leq k)}$ for the set of all subsets of $\{1, 2, \ldots, n\}$ of size at most k. Say that a subset $\mathcal{A} \subset [n]^{(\leq k)}$ is *closed* if whenever W_1, W_2, …, $W_r \in \mathcal{A}$ and W_1, W_2, …, $W_r \vdash W$ then $W \in \mathcal{A}$ (where $r \geq 2$ is some fixed number to be chosen later). If $\mathcal{A} \subset [n]^{(\leq k)}$ then the *closure* of $\mathcal{A}$, written $\mathcal{A}^*$, is the intersection of all closed sets that contain $\mathcal{A}$.

Given a set $\mathcal{A} \subset [n]^{(\leq k)}$, write $\lceil \mathcal{A} \rceil$ for the set of all graphs with vertex set $[n]$ that contain a clique with vertex set A for some $A \in \mathcal{A}$.

The lattice $\mathcal{L}$ will consist of all sets of the form $\lceil \mathcal{A} \rceil$ for $\mathcal{A}$ closed. We define the operations $\sqcup$ and $\sqcap$ by $\lceil \mathcal{A} \rceil \sqcap \lceil \mathcal{B} \rceil = \lceil \mathcal{A} \cap \mathcal{B} \rceil$ and $\lceil \mathcal{A} \rceil \sqcup \lceil \mathcal{B} \rceil = \lceil (\mathcal{A} \cup \mathcal{B})^* \rceil$. It is easy to check that $\lceil \mathcal{A} \rceil \sqcap \lceil \mathcal{B} \rceil \subset \lceil \mathcal{A} \rceil \cap \lceil \mathcal{B} \rceil$ and $\lceil \mathcal{A} \rceil \sqcup \lceil \mathcal{B} \rceil \supset \lceil \mathcal{A} \rceil \cup \lceil \mathcal{B} \rceil$.

Lemma 35. *A closed set can have at most $(r-1)^k$ minimal elements.*

Proof. If $\mathcal{A}$ is closed and A_1, A_2, …, A_r, A are minimal elements of $\mathcal{A}$ then it is not possible to find $B \subsetneq A$ such that $A_i \cap A_j \subset B$ for all $i \neq j$. We shall show that any collection of sets with this property has cardinality at most $(r-1)^k$. We do this by induction on r (noting that the result is clear when $r = 2$).

Let $\mathcal{M}$ be a collection of sets with this property and let $F \in \mathcal{M}$. For each $C \subset F$, let $\mathcal{M}_C = \{A \in \mathcal{M} : A \cap F = C\}$. If $A_1, A_2, \ldots, A_{r-1}, A \in \mathcal{M}_C$ and if $B \subsetneq A$ with $A_i \cap A_j \subset B$ for all $i \neq j$, then $C \subset B$, from which it follows that $A_1, A_2, \ldots, A_{r-1}, F \vdash B$, a contradiction. It follows that the sets $A - F = A - C$ for $A \in \mathcal{M}_C$ have the same property, but with k and r replaced by $k - |C|$ and $r - 1$.

Hence by induction, $|\mathcal{M}_C| \leq (r-2)^{k-|C|}$ and so

$$|\mathcal{M}| \leq \sum_{C \subset F} (r-2)^{k-|C|} \leq \sum_{i=0}^{|F|} \binom{|F|}{i} (r-2)^{k-i} \leq \sum_{i=0}^{k} \binom{k}{i} (r-2)^{k-i} = (r-1)^k.$$

□

Example. Let $B_1, B_2, \ldots, B_k$ be disjoint sets of size $r-1$ and let $\mathcal{M} \subset [n]^{(\leq k)}$ be the collection $\{A : |A \cap B_i| = 1,\ 1 \leq i \leq k\}$. Then $|\mathcal{M}| = (r-1)^k$. If $A_1, A_2, \ldots, A_r \in \mathcal{M}$ then for all i, there exist $j \neq k$ with $A_j \cap B_i = A_k \cap B_i$. Therefore if $A_1, A_2, \ldots, A_r \vdash B$ we have $|B \cap B_i| \geq 1$ for all i.

Definition. A *g-colouring* of $[n]$ is a function $f : [n] \to [g]$. If $A \subset [n]$ then we say A is *properly coloured* (or *PC*) if every element of A has a different colour (i.e. $f|_A$ is an injection).

Lemma 36. *Let $\mathcal{A} \subset [n]^{(\leq k)}$ be a set system and let $A \in [n]^{(\leq k)}$ be such that $A_1, A_2, \ldots, A_r \vdash A$ for some $A_1, A_2, \ldots, A_r \in \mathcal{A}$. Let $f : [n] \to [g]$ be a random g-colouring of $[n]$. Then the probability that A is properly coloured and no $A' \in \mathcal{A}$ is PC is at most $\left(1 - \frac{g(g-1)\ldots(g-k+1)}{g^k}\right)^r$.*

Proof. The probability in question is clearly at most

$$\begin{aligned}
& \mathbb{P}[A \text{ is PC and no } A_i \text{ is PC for } i = 1, 2, \ldots, r] \\
\leq\ & \mathbb{P}[\text{no } A_i \text{ is PC} | A \text{ is PC}] \\
=\ & \prod_{i=1}^{r} \mathbb{P}[A_i \text{ is not PC} | A \text{ is PC}] \qquad \text{(since the } A_i - A \text{ are disjoint)} \\
=\ & \prod_{i=1}^{r} \left(1 - \mathbb{P}[A_i \text{ is PC} | A \text{ is PC}]\right).
\end{aligned}$$

But $\mathbb{P}[A_i \text{ is PC} | A \text{ is PC}] \geq \frac{g(g-1)\ldots(g-k+1)}{g^k}$ which proves the lemma. □

Lemma 37. *Let $\mathcal{A} \subset [n]^{(\leq k)}$ be a set system and let $f : [n] \to [g]$ be a random g-colouring of $[n]$. Then the probability that no $A' \in \mathcal{A}$ is PC but some $A \in \mathcal{A}^*$ is PC is at most $n^k \left(1 - \frac{g(g-1)\dots(g-k+1)}{g^k}\right)^r$.*

Proof. Enumerate the sets in $\mathcal{A}^*$ as $A_1, A_2, \dots, A_m$, starting with the sets in $\mathcal{A}$, such that each $A_i \in \mathcal{A}^* - \mathcal{A}$ is entailed by r earlier sets. Then the event in question is the disjoint union of the events 'A_i is the first set to be PC' for $i > |\mathcal{A}|$. By Lemma 36, the probability of this is at most $\left(1 - \frac{g(g-1)\dots(g-k+1)}{g^k}\right)^r$. But obviously $m \leq \sum_{i=0}^k \binom{n}{i} \leq n^k$ and so the lemma is proved. □

A g-colouring f defines a complete g-partite graph $G(f)$ in $[n]$ where you join x to y iff $f(x) \neq f(y)$. A set A is PC iff the clique with vertex set A is a subgraph of $G(f)$. So if $\mathcal{A}$ is a set system, then $G(f) \in \lceil \mathcal{A} \rceil$ iff some $A \in \mathcal{A}$ is PC. So Lemma 37 gives an upper bound on the proportion of complete g-partite graphs contained in a set of the form $\lceil \mathcal{A}^* \rceil - \lceil \mathcal{A} \rceil$.

The main argument

Let $\mathcal{C}$ be the set of all graphs on n vertices that contain a clique of size m (so $\mathcal{C}$ can be thought of as a subset of $\{0,1\}^{[n]^{(2)}}$). Suppose that $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_N = \mathcal{C}$ is a monotone straight-line computation of $\mathcal{C}$ and let $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_N$ be the results of the corresponding computation in $\mathcal{L}$ with $\cap$ and $\cup$ replaced by $\sqcap$ and $\sqcup$ respectively.

Case 1: $\mathcal{B}_N$ is the set of all graphs on $[n]$.
We know by Lemma 34 that $\mathcal{B}_N \subset \mathcal{A}_N \cup \bigcup_{l=1}^N \delta_\sqcup(M_l, N_l)$ where (M_l, N_l) is a pair of the form $(\mathcal{B}_i, \mathcal{B}_j)$. But

$$\delta_\sqcup(\mathcal{B}_i, \mathcal{B}_j) = \lceil (\mathcal{B}_i \cup \mathcal{B}_j)^* \rceil - \lceil \mathcal{B}_i \cup \mathcal{B}_j \rceil = \lceil (\mathcal{B}_i \cup \mathcal{B}_j)^* \rceil - (\lceil \mathcal{B}_i \rceil \cup \lceil \mathcal{B}_j \rceil)\,.$$

By Lemma 37, the proportion of complete $(m-1)$-partite graphs in $\delta_\sqcup(\mathcal{B}_i, \mathcal{B}_j)$ is therefore at most $n^k \left(1 - \frac{(m-1)(m-2)\dots(m-k)}{(m-1)^k}\right)^r$. Set $k = \sqrt{m}$. Then

$$\frac{(m-1)(m-2)\dots(m-k)}{(m-1)^k} \geq \left(1 - \frac{k}{m}\right)^k \approx e^{-\frac{k^2}{m}} = \frac{1}{e} = c.$$

So the proportion of complete $(m-1)$-partite graphs in $\delta_\sqcup(\mathcal{B}_i, \mathcal{B}_j)$ is at most $n^k e^{-cr}$, which implies that $N \geq e^{cr - k \log n}$. So we will need $cr \geq 2k \log n$.

Case 2: $\mathcal{B}_N$ is not the set of all graphs.

First we attain an upper bound on the number of m-cliques contained in $\mathcal{B}_N$. We know that $\mathcal{B}_N$ is a set of the form $\lceil\mathcal{A}\rceil$ for some closed set $\mathcal{A}$. Any m-clique in $\lceil\mathcal{A}\rceil$ must have a vertex set of size m that contains some minimal element A of $\mathcal{A}$ as a subset, with $|A| \geq 2$.

Now by Lemma 35, the number of minimal elements of $\mathcal{A}$ of size i is at most $(r-1)^i$. Each is contained in at most $\binom{n-i}{m-i}$ sets of size m, so the number of m-cliques in $\lceil\mathcal{A}\rceil$ is at most

$$\sum_{i=2}^{k}(r-1)^i\binom{n-i}{m-i} \leq \binom{n}{m}\sum_{i=2}^{k}(r-1)^i\left(\frac{m}{n}\right)^i$$

so if $\frac{(r-1)m}{n} \leq \frac{1}{2}$ then this is at most $\frac{1}{2}\binom{n}{m}$.

It remains to obtain an upper bound for the number of m-cliques in any $\delta_\sqcap(\lceil\mathcal{A}\rceil, \lceil\mathcal{B}\rceil)$ with $\mathcal{A}$, $\mathcal{B}$ closed. Then we will be done by Lemma 34 since $\mathcal{A}_N \subset \mathcal{B}_N \cup_{i=1}^{N} \delta_\sqcap(M_i, N_i)$ and so very many steps will be required to reach $\frac{1}{2}\binom{n}{m}$.

Suppose that $K(Z)$ is an m-clique with vertex set Z, and suppose that $K(Z) \in (\lceil\mathcal{A}\rceil \cap \lceil\mathcal{B}\rceil) - (\lceil\mathcal{A}\cap\mathcal{B}\rceil)$. Then Z must contain minimal elements X and Y of $\mathcal{A}$ and $\mathcal{B}$ respectively, and $X \cup Y$ is not an element of $\mathcal{A}\cap\mathcal{B}$. The only way this can happen is if $|X\cup Y| > k$, so either $|X| > \frac{k}{2}$ or $|Y| > \frac{k}{2}$. It follows that the number of possible Z is at most

$$2\sum_{i>\frac{k}{2}}(r-1)^i\binom{n-i}{m-i} \leq 2\binom{n}{m}\sum_{i>\frac{k}{2}}\left(\frac{(r-1)m}{n}\right)^i \leq 4\binom{n}{m}2^{-\frac{k}{2}}.$$

It follows that, in case 2, $N \geq \frac{\frac{1}{2}\binom{n}{m}}{4\binom{n}{m}2^{-\frac{k}{2}}} = \frac{1}{8}2^{\frac{k}{2}}$.

We have shown that $N \geq \min\{e^{\frac{cr}{2}}, \frac{1}{8}2^{\frac{k}{2}}\}$, subject to the restrictions $k^2 = m$, $r \geq ck\log n$ and $rm \leq cn$, so $r \geq c\sqrt{m}\log n$, and so $m^{\frac{3}{2}}\log n \leq cn$, and so $m \leq \left(\frac{cn}{\log n}\right)^{\frac{2}{3}}$. So take that for m, and let $k = \left(\frac{cn}{\log n}\right)^{\frac{1}{3}}$ and $r = ck\log n$; we obtain a bound of $e^{\left(\frac{cn}{\log n}\right)^{\frac{1}{3}}}$ as the monotone circuit complexity of a function of $\binom{n}{2}$ variables.

6 Algebraic methods

Theorem 38 (Frankl, Wilson). *Let $\mathcal{F}$ be a subset of $[n]^{(k)}$—i.e. a set system consisting of subsets of $[n]$ of size k. Let p be a prime and let*

$\lambda_1, \lambda_2, \ldots, \lambda_s$ be residues mod p, none of them congruent to k. Let Λ be the set $\{\lambda_1, \lambda_2, \ldots, \lambda_s\}$

Suppose that $|F \cap F'|$ is congruent to some $\lambda \in \Lambda$ whenever F, $F' \in \mathcal{F}$ with $F \neq F'$. Then $|\mathcal{F}| \leq \binom{n}{s}$.

Proof. Assume wlog $s < k$. For each $i < j$ let $N(i,j)$ be the $\binom{n}{i} \times \binom{n}{j}$ matrix defined on pairs $(A,B) \in [n]^{(i)} \times [n]^{(j)}$ by

$$N(i,j)(A,B) = \begin{cases} 1 & \text{if } A \subset B \\ 0 & \text{if } A \not\subset B \end{cases}.$$

Let V be the vector space over $\mathbb{R}$ spanned by the rows of $N(s,k)$. There are $\binom{n}{s}$ rows, so $\dim V \leq \binom{n}{s}$.

Let $i \leq s$ and look at the matrix $N(i,s)N(s,k)$, which is an $\binom{n}{i} \times \binom{n}{k}$ matrix defined on pairs $(A,B) \subset [n]^{(i)} \times [n]^{(k)}$. We see that

$$\begin{aligned} N(i,s)N(s,k)(A,B) &= \sum_{C \subset [n]^{(s)}} \mathbf{1}_{A \subset C} \mathbf{1}_{C \subset B} \\ &= \begin{cases} \binom{k-i}{s-i} & \text{if } A \subset B \\ 0 & \text{if } A \not\subset B \end{cases} \end{aligned}$$

so $N(i,s)N(s,k)(A,B) = \binom{k-i}{s-i} N(i,k)(A,B)$. It follows that the rows of $N(i,k)$ belong to V.

Now let $M(i,k) = N(i,k)^T N(i,k)$. If $(A,B) \in [n]^{(k)} \times [n]^{(k)}$, then

$$M(i,k)(A,B) = \sum_{C \subset [n]^{(i)}} \mathbf{1}_{C \subset A} \mathbf{1}_{C \subset B} = \binom{|A \cap B|}{i}.$$

Now choose $a_1, a_2, \ldots, a_s$ such that $\prod_{\lambda \in \Lambda}(x - \lambda) = \sum_{i=1}^{s} a_i \binom{x}{i}$ and let $M = \sum_{i=1}^{s} a_i M(i,k)$. The rows of $M(i,k)$ and hence M belong to V. But $M(A,B) \equiv 0 \pmod{p}$ if A, $B \in \mathcal{F}$ with $A \neq B$ since then $A \cap B \equiv \lambda \pmod{p}$ for some $\lambda \in \Lambda$ and so $\sum a_i \binom{|A \cap B|}{i} \equiv 0 \pmod{p}$ (by definition of $a_1, \ldots, a_s$). If $A = B$ then $|A \cap B| \not\equiv \lambda \pmod{p}$ for any $\lambda \in \Lambda$ so $M(A,B) \not\equiv 0 \pmod{p}$. That applies if A, $B \in \mathcal{F}$. So the restriction of M to pairs $(A,B) \in \mathcal{F}$ has rows which are linearly independent over the field $\mathbb{F}_p$ and hence over $\mathbb{R}$, and so is a matrix of rank $|\mathcal{F}|$. But since the rows of M belong to V, M has rank at most $\binom{n}{s}$. So $|\mathcal{F}| \leq \binom{n}{s}$. □

Borsuk's conjecture

Borsuk asked whether every convex body in $\mathbb{R}^d$ can be partitioned into at most $d+1$ sets of smaller diameter.

Larman had the following idea for potential counterexamples. Take a set system $\mathcal{F} \subset [n]^{(k)}$ and associate with $F \in \mathcal{F}$ a 01-sequence in the usual way and regard it as a point in $\mathbb{R}^n$. Then writing F for this point as well, we have $d(F, F') = |F \Delta F'|^{\frac{1}{2}}$. In particular, this distance is maximized when $|F \cap F'|$ is minimized. So look for a set system $\mathcal{F}$ such that however you partition it into $\mathcal{F}_1 \cup \mathcal{F}_2 \cup \cdots \cup \mathcal{F}_m$ if m is not too large then some $\mathcal{F}_i$ contains F and F' with $|F \cap F'|$ minimal (amongst all pairs from $\mathcal{F}$).

Corollary 39. *Let p be a prime and let $n = 4p$. Let $\mathcal{F} \subset [n]^{(2p)}$ be such that for all F, $F' \in \mathcal{F}$, $|F \cap F'| \neq p$. Then $|\mathcal{F}| \leq \binom{n}{p-1}$.*

Proof. The condition implies that $|F \cap F'| \in \{1, 2, \ldots, p-1\} \pmod p$ if $F \neq F'$. □

Disproof of Borsuk's conjecture

Let $m = 4p$ and let $n = \binom{m}{2}$. Identify $[m]^{(2)}$ with $[n]$.

Given any partition of $[m]$ into two sets A and B of equal size, let $G(A, B)$ be the complete bipartite graph with vertex sets A and B. Then, if (A, B) and (C, D) are two such partitions with $|A \cap C| = k$, we have $|G(A, B) \cap G(C, D)| = k^2 + (2p - k)^2$ which is minimized when $k = p$. So by Corollary 39, any subset $\mathcal{F}' \subset \mathcal{F}$ that does not contain a pair $G(A, B)$ and $G(C, D)$ at maximal distance apart has size at most $\binom{m}{p-1}$ (since if you look at the A's you get $\mathcal{A} \subset [m]^{(2p)}$ with $A \neq \mathcal{A}' \implies |A \cap A'| \neq p$). But $|\mathcal{F}| = \frac{1}{2}\binom{m}{2p}$. So the smallest partition has size at least $\binom{m}{2p} / 2\binom{m}{p-1}$.

Now, $\binom{4p}{2p} / \binom{4p}{p-1}$ is 'roughly' $2^{4p} / \left(\frac{4}{3}\right)^{3p} \cdot 4^p = \alpha^{4p}$ for some $\alpha > 1$. Hence we have shown the following:

Theorem 40. *There is a convex body in $\mathbb{R}^n$ that cannot be partitioned into $c^{\sqrt{n}}$ parts of smaller diamter, with $c > 1$ an absolute constant.*

Theorem 41. *Let p be a prime and n a positive integer. Let G be a graph with vertex set $[n]^{(p^2-1)}$ where F is joined to F' iff $|F \cap F'| \equiv -1 \pmod p$. Then G contains no clique or independent set of size greater than $\binom{n}{p-1}$.*

Proof. If $\mathcal{F}$ is an independent set then $|F \cap F'| \in \{0, 1, \ldots, p-2\} \pmod p$ if $F \neq F'$, so by Theorem 38, $\mathcal{F} \leq \binom{n}{p-1}$. If $\mathcal{F}$ is a clique then for $F \neq F'$ we have $|F \cap F'| \in \{p-1, 2p-1, \ldots, p^2-p-1\}$ which implies by Theorem 38 with respect to some prime $q > p - 1$, that $|\mathcal{F}| \leq \binom{n}{p-1}$. □

If we now set $n = p^3$, we find that the Ramsey number $R\left(\binom{p^3}{p-1}+1\right)$ is greater than $\binom{p^3}{p^2-1}$. Let $k = \binom{p^3}{p-1} \le p^{3p}$. Then

$$\binom{p^3}{p^2-1} \ge p^{p^2-1} \ge \left(\frac{\log k}{\log\log k}\right)^{c\left(\frac{\log k}{\log\log k}\right)^2} \ge e^{c\log\log k\left(\frac{\log k}{\log\log k}\right)^2} = e^{c\frac{(\log k)^2}{\log\log k}}$$

which exceeds any fixed power of k.

Theorem 42. *There is a subset $A \subset \{1, 2, \dots, n\}$ of size $c\sqrt{n}$ containing no quadruple (x, y, z, w) with $x + y = z + w$ except degenerate ones.*

Proof. Let p be an odd prime. Let $\Gamma \subset [p]^2$ be the set

$$\Gamma = \left\{\left(x, x^2 \pmod p\right) : x = 1, 2, \dots, p\right\}.$$

Then

$$(x, x^2)+(y, y^2) = (z, z^2)+(w, w^2) \pmod p \implies \begin{cases} x - z \equiv w - y & (1) \\ x^2 - z^2 \equiv w^2 - y^2 & (2) \end{cases}.$$

If $x \neq z$ then $x \not\equiv z$ and we can divide (2) by (1) to get $x + z \equiv w + y$, giving $x \equiv w$ and $y \equiv z$, and so $x = w$ and $y = z$.

Now map Γ to $[n]$ by the map $(a, b) \mapsto a+2pb-2p$. If a_i, b_i, c_i, $d_i \in [p]$ with $a_1 + 2pa_2 + b_1 + 2pb_2 = c_1 + 2pc_2 + d_1 + 2pd_2$ then $a_1 + b_1 \equiv c_1 + d_1 \pmod{2p}$ from which it follows that $a_1 + b_1 = c_1 + d_1$ and so $a_2 + b_2 = c_2 + d_2$. So the image of Γ is a subset of $[2p^2]$ of size p containing no non-trivial $x + y = z + w$. □

Theorem 43. *For all $n \in \mathbb{N}$, there is a family $\mathcal{A} \subset \mathbb{P}[2n]$ of size 2^n with no non-trivial solutions of $A\Delta B = C\Delta D$.*

Proof. We can identify $\mathbb{P}[2n]$ in the usual way with $\{0, 1\}^{2n}$ and we can identify that with $\{0, 1\}^n \times \{0, 1\}^n$. Now we'd like to find $X \subset \{0, 1\}^{2n}$ with no non-trivial solutions to $x + y = z + w$.

We can think of $\{0, 1\}^n$ as the additive group of the field with 2^n elements, and then let $X = \{(x, x^3) : x \in \{0, 1\}^n\} \subset \{0, 1\}^n \times \{0, 1\}^n$. As before, if $(x, x^3) + (y, y^3) = (z, z^3) + (w, w^3)$ then $x + y = z + w$ and $x^3 + y^3 = z^3 + w^3$ which implies, unless $x = y$ and $z = w$, that $x^2 + xy + y^2 = z^2 + zw + w^2$. But $(x + y)^2 = x^2 + y^2 = z^2 + w^2$ so $xy = zw$. That is enough to show that $(x, y) = (z, w)$. □

Theorem 44 (Behrend, 1947). *There is a subset $X \subset [n]$ of size $ne^{-c\sqrt{\log n}}$ that contains no arithmetic progression of length 3.*

Proof. Let m, $k \in \mathbb{N}$ and define $B_r \subset [m]^k$ to be the set

$$B_r = \Big\{(x_1, x_2, \dots, x_k) : \sum_{i=1}^{k} x_i^2 = r\Big\}.$$

This set lies on the surface of a sphere so contains no three collinear points, and in particular no arithmetic progression of length 3 (in the obvious sense). There are at most km^2 values of r for which $B_r \neq \emptyset$, so some B_r has size at least $\frac{m^k}{m^2 k}$. Let B be such a B_r. Now map B to $[(2m)^k]$ by the map $\phi(x_1, x_2, \dots, x_k) = x_1 + 2mx_2 + (2m)^2 x_3 + \cdots + (2m)^{k-1} x_k$. Then if $\phi(\mathbf{x})$, $\phi(\mathbf{y})$ and $\phi(\mathbf{z})$ are in arithmetic progression, we have x_1, y_1, z_1 in arithmetic progression mod $2m$, giving x_1, y_1, z_1 in arithmetic progression. Then $2mx_2$, $2my_2$, $2mz_2$ are in arithmetic progression mod $(2m)^2$ giving x_2, y_2, z_2 in arithmetic progression and so on. So $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ are in arithmetic progression, a contradiction.

Let $n = (2m)^k$. Then we have found a subset of $[n]$ with no arithmetic progression of length 3 with size at least $\frac{m^k}{m^2 k}$. Let $k = \sqrt{\log n}$ and $m = \frac{1}{2} e^{\sqrt{\log n}}$. Then the size is at least

$$\frac{n}{2^{\sqrt{\log n}} \frac{1}{4} e^{2\sqrt{\log n}} \sqrt{\log n}} \geq n e^{-c\sqrt{\log n}}.$$

□

7 Topological methods

Theorem 45 (The Borsuk-Ulam Theorem). *Let*

$$S_n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}.$$

Let $f : S_n \to \mathbb{R}^n$ be any continuous function. Then there exists $x \in S_n$ such that $f(x) = f(-x)$.

Corollary 46. *Let A_1, A_2, …, A_{n+1} be subsets of S_n, all either open or closed, such that $S_n = \bigcup_{i=1}^{n+1} A_i$. Then some A_i contains a pair $\{x, -x\}$ of antipodal points.*

Proof. Define $f : S_n \to \mathbb{R}^n$ by $f(x) = \big(d(x, A_1), d(x, A_2), \dots, d(x, A_n)\big)$. Then f is continuous so, by the Borsuk-Ulam Theorem, we can find x such that $f(x) = f(-x)$. We know that $x \in A_i$ for some i.

If $i \leq n$ and A_i is closed then $d(-x, A_i) = 0$ so $-x \in A_i$.

If $i \leq n$ and A_i is open then $d(-x, A_i) = 0$. Then a small perturbation $-y$ of $-x$ is in A_i, and $y \in A_i$ since A_i is open and the perturbation small.

This shows that we are done if either $x \in A_i$ for some $i \leq n$ or $-x \in A_i$ for some $i \leq n$. So we can assume that this is not the case. Then $x \in A_{n+1}$ and $-x \in A_{n+1}$. □

Theorem 47. *For every n there is a finite triangle-free graph with chromatic number n.*

Proof. First, we will find an infinite graph G. Let $V(G) = S_n$ and join x to y iff $\langle x, y\rangle < -\frac{1}{2}$—i.e. the angle between x and y is greater than 120°. Clearly G is triangle-free. Now, let $S_n = A_1 \cup A_2 \cup \cdots \cup A_{n+1}$. Then, by Corollary 46, some $\bar{A}_i$ contains an antipodal pair $\{x, -x\}$. Choose y and z in A_i, close enough to x and $-x$ for $\langle y, z\rangle < -\frac{1}{2}$, and then we have an edge inside A_i. So $\chi(G) > n + 1$.

To make this into a finite graph, let $\delta > 0$ be small and let Δ be a δ-net of S_n—i.e. for all $x \in S_n$, there is some $y \in \Delta$ such that $d(x, y) < \delta$. Join x, $y \in \Delta$ iff $\langle x, y\rangle < -\frac{2}{3}$. Now let $\Delta = B_1 \cup B_2 \cup \cdots \cup B_{n+1}$. Let $A_i \subset S_n$ be $\{x : \exists y \in B_i \text{ with } d(x, y) < \delta\}$. Then some $\bar{A}_i$ contains an antipodal pair $\{z, -z\}$. But we can find v, $w \in B_i$ close to z, $-z$, so $\langle v, w\rangle < -\frac{2}{3}$ and so the colouring is not proper. □

Theorem 48 (Kneser's conjecture, proved by Lovasz). *Let G be the graph with vertex set $[n]^{(k)}$ in which A and B are joined iff $A \cap B = \emptyset$. Then $\chi(G) = \max\{n - 2k + 2, 1\}$.*

Proof (due to Greene). First consider the colouring

$$\phi(G) = \min\{\min A, n - 2k + 2\}.$$

This is an $(n - 2k + 2)$-colouring. If $\phi(A) = \phi(B) < n - 2k + 2$ then $\min A = \min B$ and so $A \cap B \neq \emptyset$. If $\phi(A) = \phi(B) = n - 2k + 2$ then A, $B \subset [n - 2k + 2, n]$, but this set has size $2k - 1$ and so A and B cannot be disjoint.

To show that $\chi(G) = n - 2k + 2$, let $d = n - 2k + 1$ and let $X \subset S_d$ be a set of n points in general position, so no $d + 1$ points of X are contained in a d-dimensional subspace of $\mathbb{R}^{d+1}$. Identify $[n]$ with X, or, in other words, take $V(G) = X^{(k)}$.

Let $\phi : X^{(k)} \to [d]$ be a colouring of $X^{(k)}$. For each $x \in S_d$, let $H_x = \{y \in S_d : \langle x, y\rangle > 0\}$. For each $i \leq d$, let A_i be the set of all points $x \in S_d$ such that H_x contains a k-tuple from X coloured with colour i. Let $A_{d+1} = S_d - (A_1 \cup A_2 \cup \cdots \cup A_d)$. Note that A_i is open for $i \leq d$, and so A_{d+1} is closed.

By the Borsuk-Ulam Theorem, some A_i must contain an antipodal pair $\{x, -x\}$. If $i \leq d$, then both H_x and H_{-x} contain k-tuples coloured with colour i. But H_x and H_{-x} are disjoint, so these k-tuples are disjoint and the colouring is not proper.

If $i = d+1$ then neither H_x nor H_{-x} contains any k-tuple from X. Hence $S_d - (H_x \cup H_{-x}) = \{y : \langle x, y\rangle = 0\}$ contains at least $n - 2(k-1) = d+1$ points, contradicting the fact that the points of X are in general position. □

Next, we look at another equivalent formulation of the Borsuk-Ulam Theorem.

Theorem 49. *Let $B_n = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$. Then there is no continuous map $f : B_n \to S_{n-1}$ that is antipodal on the boundary—i.e. such that $f(x) = -f(-x)$ for every $x \in S_{n-1}$.*

Before proving this, we observe that it implies the Borsuk-Ulam Theorem.

Proof (of Theorem 45). Suppose that we could find some continuous map $f : S_n \to \mathbb{R}^n$ that did not identify two antipodal points. Then we can define $g : S_n \to \mathbb{R}^n$ by $g(x) = (f(x) - f(-x))/\|f(x) - f(-x)\|$. Note that g is continuous and $g(-x) = -g(x)$. Also, the image of g lies inside S_{n-1}.

Now define $h : B_n \to S_{n-1}$ by $h(\mathbf{x}) = g\big((\mathbf{x}, \sqrt{1 - \|\mathbf{x}\|^2})\big)$. Then h is continuous, and antipodal on the boundary of B_n. □

Exercise. Show that Theorem 45 implies Theorem 49.

Theorem 50 (Tucker's Lemma). *Let T be a triangulation of B_n that is antipodally symmetric on the boundary. That is, if $\sigma \in T$ with $\sigma \subset S_{n-1}$ then $-\sigma \in T$. Let λ be a labelling of the vertices of T with labels from the set $\{1, 2, \ldots, n\} \cup \{-1, -2, \ldots, -n\}$. Then if this labelling satisfies $\lambda(-x) = -\lambda(x)$ for every vertex x, there must be some edge (i.e. 1-simplex) with vertices labelled i and $-i$.*

We shall prove Tucker's lemma in an alternative formulation. Let L be the "octahedral triangulation" consisting of all convex hulls of up to n of the vectors $\pm e_i$ (where $(e_i)_{i=1}^n$ is the standard basis of $\mathbb{R}^n$) such that you never use both e_i and $-e_i$.

Then a simplicial map from T to L is exactly a map ϕ that takes the vertices of T to $\{e_1, e_2, \ldots, e_n\} \cup \{-e_1, -e_2, \ldots, -e_n\}$ such that if x and y are vertices of the same simplex in T then $\{\phi(x), \phi(y)\} \neq \{e_i, -e_i\}$ for any i.

So under the hypotheses on T, Tucker's lemma says that there is no simplicial map $\phi : T \to L$ such that $\phi|_{S_{n-1}}$ is antipodal.

We first show that Tucker's lemma implies Theorem 49.

Proof (of Theorem 49). Assume Theorem 49 is false and let $f : B_n \to S_{n-1}$ be a continous map, antipodal on the boundary. We shall construct a counterexample to Tucker's lemma.

Since f is uniformly continuous, we can find $\delta > 0$ such that whenever $d(x, y) < \delta$ we have $d(f(x), f(y)) < \frac{1}{\sqrt{n}}$. Let T be a triangulation of B_n of mesh less than δ (i.e. with every simplex having diameter less than δ). We now define a simplicial map $\phi : T \to L$ as follows. For each vertex $x \in T$, let $\phi(x)$ be the nearest point of $\{\pm e_1, \pm e_2, \ldots, \pm e_n\}$ to $f(x)$, choosing i minimal in the case of a tie. Note that $\phi(-x) = -\phi(x)$ for $x \in S_{n-1}$. Also,

$$\phi(x) = \left\{ \begin{array}{c} e_i \\ -e_i \end{array} \right\} \implies f(x)_i \left\{ \begin{array}{c} \geq \frac{1}{\sqrt{n}} \\ \leq -\frac{1}{\sqrt{n}} \end{array} \right.$$

so if x and y belong to the same simplex then $\{\phi(x), \phi(y)\} \neq \{e_i, -e_i\}$ by choice of δ. □

Given a simplicial complex T, a *k-chain* is a set of k-simplices belonging to T. We shall think of these sets as formal sums of simplices over $\mathbb{Z}_2$. Given a simplex σ, the *boundary* $\partial\sigma$ is the set of all $(k-1)$-facets of σ. Given a k-chain $\sigma_1 + \sigma_2 + \cdots + \sigma_r = A$, we define ∂A to be $\partial\sigma_1 + \partial\sigma_2 + \cdots + \partial\sigma_r$. We have the following simple facts: $\partial(A + A') = \partial A + \partial A'$ and $\partial\partial A = 0$ (since this is true for a k-simplex σ). Given a simplicial map $\phi : T \to K$ and a k-simplex $\sigma \in T$, we define $\phi_k(\sigma)$ to be $\phi(\sigma)$ if this is a k-simplex and 0 otherwise. If $A = \sigma_1 + \sigma_2 + \cdots + \sigma_r$ then $\phi_k(A)$ is defined to be $\phi_k(\sigma_1) + \phi_k(\sigma_2) + \cdots + \phi_k(\sigma_r)$. It is easy to check (simplex by simplex) that $\phi_{k-1}(\partial A) = \partial(\phi_k(A))$.

Proof (of Tucker's lemma). This comes in three steps. Let T be a triangulation of B_n, let K be the restriction of T to S_{n-1}, assume that K is antipodally symmetric and let $\phi : T \to L$ be a simplicial map. We also make the following extra assumption about T:

For $0 \leq k \leq n-1$, let

$$H_k^+ = \{x \in S_{n-1} : x_{k+2} = x_{k+3} = \cdots = x_n = 0, x_{k+1} \geq 0\}.$$

and

$$H_k^- = \{x \in S_{n-1} : x_{k+2} = x_{k+3} = \cdots = x_n = 0, x_{k+1} \leq 0\}.$$

We shall assume that for each $H_k^{\pm}$ there is a subcomplex of T that triangulates it.

Call such a triangulation *proper* and note that Tucker for proper triangulations still implies Theorem 49.

For each k, let $A_k^{\pm}$ be the k-chain from T that triangulates $H_k^{\pm}$ and let $C_k^{\pm} = \phi_k A_k^{\pm}$. Let A_k be the k-chain from T that triangulates $H_k^+ \cup H_k^-$ and let $C_k = \phi_k(A_k)$. Note in particular that A_{n-1} is simply the $(n-1)$-chain of all $(n-1)$-simplices in K. Let A_n be the chain of all n-simplices of T. We shall show the following facts:

(i) either C_{n-1} is 0 or it consists of all $(n-1)$-simplices of L, i.e. either every $(n-1)$-simplex of L has an odd number of preimages or every $(n-1)$-simplex of L has an even number of preimages. We say that ϕ has odd or even *degree* respectively;

(ii) for any simplicial map from T to L, its restriction to A_{n-1} has even degree;

(iii) any antipodally symmetric map $\phi : K \to L$ has odd degree.

The proofs of these facts are as follows:

(i) If this is false then there must be neighbouring $(n-1)$-simplices $\sigma, \tau \in L$ with $\sigma \in C_{n-1}$ and $\tau \notin C_{n-1}$. Then their common facet belongs to ∂C_{n-1}.

But $\partial C_{n-1} = \phi_{n-2}(\partial A_{n-1})$, and $\partial A_{n-1} = 0$, since any $(n-2)$-simplex of K is a facet of exactly two $(n-1)$-simplices.

(ii) Let $\phi : T \to L$ be a simplicial map. Then $\phi_{n-1}(A_{n-1}) = \phi_{n-1}(\partial A_n)$. But this is $\partial(\phi_n A_n) = 0$ since $\phi_n A_n = 0$, since L contains no n-simplices.

(iii) Assume, for a contradiction, that ϕ has even degree, i.e. that $C_{n-1} = 0$. We know that $A_{n-1} = A_{n-1}^+ + A_{n-1}^-$, so $C_{n-1} = \phi_{n-1}A_{n-1} = C_{n-1}^+ + C_{n-1}^-$. Since ϕ is antipodally symmetric and A_{n-1}^+ is antipodal to A_{n-1}^-, C_{n-1}^+ is antipodal to C_{n-1}^-, but also $C_{n-1}^- = C_{n-1}^+$, so C_{n-1}^+ is antipodally symmetric. But $C_{n-2} = \partial C_{n-1}^+$ so C_{n-2} is the boundary of an antipodally symmetric $(n-1)$-chain.

Now assume, as an inductive hypothesis, that C_k is the boundary of an antipodally symmetric $(k+1)$-chain D_{k+1}. We can write $D_{k+1} = E_{k+1} + E'_{k+1}$ with E_{k+1} antipodal to E'_{k+1} (from each pair σ, $-\sigma$ in D_{k+1}, put one in E_{k+1} and the other in E'_{k+1}). Now $C_k = C_k^+ + C_k^- = \partial E_{k+1} + \partial E'_{k+1}$ and so $C_k^+ + \partial E_{k+1} = C_k^- + \partial E'_{k+1}$, but C_k^+ and C_k^- are antipodal, as are ∂E_{k+1} and $\partial E'_{k+1}$, so $C_k^+ + \partial E_{k+1}$ is antipodally symmetric. Also,

$$\partial(C_k^+ + \partial E_{k+1}) = \partial C_k^+ + \partial\partial E_{k+1} = C_{k-1} + 0.$$

Hence C_0 is the boundary of a symmetric 1-chain which implies that it consists of an even number of pairs of antipodal points. But it is also $\phi_0 A_0$ and so consists of exactly one pair of antipodal points, a contradiction. $\square$

Typeset in LaTeX 2ϵ by Paul A. Russell.