14 The Probabilistic Method

Probabilistic Graph Theory

Theorem 14.1 (Szele) *For every positive integer n, there exists a tournament on n vertices with at least $n!2^{-(n-1)}$ Hamiltonian paths.*

Proof: Construct a tournament by randomly orienting each edge of K_n in each direction independently with probability $\frac{1}{2}$. For any permutation σ of the vertices, let X_σ be the indicator random variable which has value 1 if σ is the vertex sequence of a Hamiltonian path, and 0 otherwise. Let X be the random variable which is the total number of Hamiltonain paths. Then $X = \sum X_\sigma$ and we have

$$\mathbb{E}(X) = \sum_{\sigma} \mathbb{E}(X_\sigma) = n!2^{-(n-1)}.$$

Thus, there must exist at least one tournament on n vertices which has $\geq n!2^{-(n-1)}$ Hamiltonain paths as claimed. $\square$

Theorem 14.2 *Every graph G has a bipartite subgraph H for which $|E(H)| \geq \frac{1}{2}|E(G)|$.*

Proof: Choose a random subset $S \subseteq V(G)$ by independently choosing each vertex to be in S with probability $\frac{1}{2}$. Let H be the subgraph of G containing all of the vertices, and all edges with exactly one end in S. For every edge e, let X_e be the indicator random variable with value 1 if $e \in E(H)$ and 0 otherwise. If $e = uv$, then e will be in H if $u \in S$ and $v \notin S$ or if $u \notin S$ and $v \in S$, so $\mathbb{E}(X_e) = \mathbb{P}(e \in E(H)) = \frac{1}{2}$ and we find

$$\mathbb{E}(|E(H)|) = \sum_{e \in E(G)} \mathbb{E}(X_e) = \tfrac{1}{2}|E(G)|.$$

Thus, there must exist at least one bipartite subgraph H with $|E(H)| \geq \frac{1}{2}|E(G)|$. $\square$

Theorem 14.3 (Erdös) *If $\binom{n}{t} < 2^{\binom{t}{2}-1}$, then $R(t,t) > n$. In particular, $R(t,t) > 2^{\frac{t}{2}}$ for $t \geq 3$.*

Proof: Construct a random *red/blue* colouring of the edges of K_n by colouring each edge independently either *red* or *blue* with probability $\frac{1}{2}$. For any fixed set S of t vertices, let

A_S be the event that the induced subgraph on S is *monochromatic* (either all its edges are *red* or all its edges are *blue*). Now $\mathbb{P}(A_R) = 2^{1-\binom{t}{2}}$ so the probability that at least one of the events A_S occurs is at most $\binom{n}{t}2^{1-\binom{t}{2}} < 1$. Therefore, with positive probability, none of the A_S events occur, so there is a *red/blue* edge-colouring without a monochromatic K_t and $R(t,t) \geq n$. If $t \geq 3$ then setting $n = \lfloor 2^{\frac{t}{2}} \rfloor$ we have

$$\binom{n}{t}2^{1-\binom{t}{2}} = \frac{n(n-1)\dots(n-t+1)}{t!}2^{1-\frac{t(t-1)}{2}} < \frac{2^{1+\frac{t}{2}}}{t!} \cdot \frac{n^t}{2^{\frac{t^2}{2}}} < 1$$

This completes the proof. □

Theorem 14.4 *Every loopless graph G satisfies*

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{deg(v)+1}.$$

Proof: Choose an ordering $v_1, v_2, \dots, v_n$ of $V(G)$ uniformly at random. Let S be the set of all vertices which appear before all of their neighbors in our ordering. Observe that S is an independent set. Now, for each vertex v, let X_v be the indicator random variable which has value 1 if $v \in S$ and 0 otherwise. Then $\mathbb{E}(X_v) = \mathbb{P}(v \in S) = \frac{1}{deg(v)+1}$ so $\mathbb{E}(|S|) = \sum_{v \in V(G)} \mathbb{E}(X_v) = \sum_{v \in V(G)} \frac{1}{deg(v)+1}$ and we conclude that there must exist at least one ordering for which the set S is an independent set with the required size. □

Dominating Set: If G is a graph, a set of vertices $S \subseteq V(G)$ is a *dominating set* if every vertex is either in S or has a neighbor in S.

Theorem 14.5 (Alon) *Let G be a simple n-vertex graph with minimum degree δ. Then G has a dominating set of size at most $n\,\frac{1+\ln(\delta+1)}{\delta+1}$.*

Proof: Choose a subset $S \subseteq V(G)$ by selecting each vertex to be in S independently with probability p where $p = \frac{\ln(\delta+1)}{\delta+1}$. Note that $\mathbb{E}(|S|) = pn$. Let T be the set of vertices which are not in S and have no neighbor in S, and observe that $S \cup T$ is a dominating set. For every vertex v, let X_v be the indicator random variable which has value 1 if $x \in T$ and 0 otherwise. Note that $\mathbb{E}(X_v) = \mathbb{P}(v \in T) \leq (1-p)^{\delta+1}$. Thus we have

$$\begin{aligned}
\mathbb{E}(|S \cup T|) &= \mathbb{E}(|S|) + \mathbb{E}(|T|) \\
&= pn + \sum_{v \in V(G)} X_v \\
&\leq pn + n(1-p)^{\delta+1} \\
&= n\tfrac{\ln(\delta+1)}{\delta+1} + n\left(1 - \tfrac{\ln(\delta+1)}{\delta+1}\right)^{\delta+1} \\
&\leq n\,\tfrac{1+\ln(\delta+1)}{\delta+1}.
\end{aligned}$$

Here the last inequality follows from $(1+\frac{a}{m})^m \leq e^a$ which is an easy consequence of the Taylor expansion of e^a. Since there must exist a set $S \cup T$ with size at most the expected value, this completes the proof. □

Theorem 14.6 (Erdös) *For every g, k there exists a graph with chromatic number $\geq k$ and no cycle of length $\leq g$.*

Proof: For typographical reasons, we set $\theta = \frac{1}{2g}$. Now, let $n \geq 5$ be an integer large enough so that $2g\sqrt{n} \leq \frac{n}{2}$ and $\frac{n^\theta}{6\ln n + 2} \geq k$ and set $p = n^{-1+\theta}$. Form a random graph on n vertices by choosing each possible edge to occur independently with probability p. Let G be the resulting graph, and let X be the number of cycles of G with length $\leq g$. Now we have

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{i=3}^{g} \frac{n(n-1)\ldots(n-i+1)}{2i} p^i \\
&\leq \sum_{i=3}^{g} n^i p^i \\
&\leq g\sqrt{n}
\end{aligned}$$

So, by Markov's inequality, we have $\mathbb{P}(X \geq \frac{n}{2}) \leq \mathbb{P}(X \geq 2g\sqrt{n}) \leq \frac{1}{2g\sqrt{n}}\mathbb{E}(X) \leq \frac{1}{2}$. Thus, $X \leq \frac{n}{2}$ with probability $\geq \frac{1}{2}$. Set $t = \lceil \frac{3}{p} \ln n \rceil$. Then (using the identity $1 - p < e^{-p}$ for the third inequality) we find

$$
\begin{aligned}
\mathbb{P}(\alpha(G) \geq t) &\leq \binom{n}{t}(1-p)^{\binom{t}{2}} \\
&\leq n^t(1-p)^{\frac{t(t-1)}{2}} \\
&\leq \left(ne^{-p(t-1)/2}\right)^t \\
&\leq \left(ne^{-p(3/2p)\ln n}\right)^t \\
&= \left(\tfrac{1}{\sqrt{n}}\right)^t \\
&\leq \tfrac{1}{\sqrt{n}} \\
&< \tfrac{1}{2}
\end{aligned}
$$

Since $\mathbb{P}(X \leq \frac{n}{2}) \geq \frac{1}{2}$ and $\mathbb{P}(\alpha(G) < t) > \frac{1}{2}$ there is a specific G with n vertices for which $X \leq \frac{n}{2}$ and $\alpha(G) < t$. Form the graph H from G by removing from G a vertex from each cycle of length $\leq g$. Then H has no cycles of length $\leq g$, $|V(H)| \geq \frac{n}{2}$, and $\alpha(H) \leq \alpha(G) \leq t$ so we have

$$
\begin{aligned}
\chi(H) &\geq \tfrac{|V(H)|}{\alpha(H)} \\
&\geq \tfrac{n}{2t} \\
&\geq \tfrac{n}{(6/p)\ln n+2} \\
&\geq \tfrac{n}{6n^{1-\theta}\ln n+2} \\
&\geq \tfrac{n^\theta}{6\ln n+2} \\
&\geq k
\end{aligned}
$$

So H satisfies the theorem. □

Crossing Number and Applications

Crossing Number: The *crossing number* of a graph G, denoted $cr(G)$, is the minimum number of crossings in a drawing of G in the plane with the property that edges only meet vertices at their endpoints and exactly two edges cross at each crossing point.

Lemma 14.7 *If $G = (V, E)$ is simple, then $cr(G) \geq |E| - 3|V|$.*

Proof: We proceed by induction on $|E|$. As a base, note that the result is trivial when $|E| \leq 1$. For the inductive step, consider a drawing of G in the plane with a minimum number of crossings. If there are no crossings, then the formula follows from Lemma **??**. Otherwise, choose an edge e which is crossed, and consider the graph $G - e$. By keeping the same drawing (except for the deleted edge e) we find that $cr(G) \geq cr(G - e) + 1$. Now, by induction we have $cr(G - e) \geq |E \setminus \{e\}| - |V| = |E| - 1 - |V|$ and combining these inequalities yields the desired result. $\square$

Theorem 14.8 *If* $G = (V, E)$ *is simple, then* $cr(G) \geq \frac{|E|^3}{64|V|^2}$.

Proof: Let $|V| = n$ and $|E| = m$ and fix a drawing of G in the plane with a minimum number of crossings. Set $p = \frac{4n}{m}$ (note that this is ≤ 1) and choose a random subset S of vertices by selecting each vertex to be in S independently with probability p. Let $H = G[S]$ and let X be the random variable which counts the number of crossings of H in the induced drawing. By the previous lemma, we have that $X \geq |E(H)| - 3|V(H)|$ so we find

$$\begin{aligned} 0 &\leq \mathbb{E}(X - |E(H)| + 3|V(H)|) \\ &= \mathbb{E}(X) - \mathbb{E}(|E(H)|) + 3\mathbb{E}(|V(H)|) \\ &= p^4 cr(G) - p^2 m + 3pn \\ &= p^4 \left(cr(G) - \frac{m^3}{64n^2} \right). \end{aligned}$$

Thus $cr(G) \geq \frac{m^3}{64n^2}$ as desired. $\square$

Theorem 14.9 (Szemerédi, Trotter) *Let* P *be a collection of points in* $\mathbb{R}^2$ *and let* L *be a collection of lines in* $\mathbb{R}^2$. *Then the number of incidences between* P *and* L *is at most* $4\left(|P|^{\frac{2}{3}}|L|^{\frac{2}{3}} + |P| + |L|\right)$.

Proof: Let q be the number of incidences between P and L. Removing a line from L which does not contain a point in P only improves the bound, so we may assume that every line contains at least one point. Construct a graph G drawn in the plane by declaring each point in P to be a vertex, and treating each segment of a line between two consecutive points in P to be an edge. Now we have $|V(G)| = |P|$, $|E(G)| = q - |L|$, and the number of crossings is at most $|L|^2$. If $|E(G)| \leq 4|V(G)|$ then we have $q \leq 4|P| + |L|$ and we are finished.

Otherwise, by the previous theorem we find $|L|^2 \geq cr(G) \geq \frac{(q-|L|)^3}{64|P|^2}$ so $4|P|^{\frac{2}{3}}|L|^{\frac{2}{3}} \geq q - |L|$ and this completes the proof. □

Sum sets & Product sets: If A, B are subsets of $\mathbb{R}$, then we let $A + B = \{a + b : a \in A \text{ and } b \in B\}$ and $A \cdot B = \{ab : a \in A \text{ and } b \in B\}$.

Theorem 14.10 (Elekes) *Let A be a finite subset of $\mathbb{R}$. Then $|A + A| \cdot |A \cdot A| \geq \frac{1}{64}|A|^{\frac{5}{2}}$.*

Proof: Let $P = (A + A) \times (A \cdot A) \subseteq \mathbb{R}^2$ and note that $|A|^2 \leq |P| \leq |A|^4$. For every $a, b \in A$ let $\ell_{a,b}$ be the line $\{(x, y) : y = a(x - b)\}$ and let $L = \{\ell_{a,b} : a, b \in A\}$. If $\ell_{a,b} \in L$, then for every $c \in A$ the line L is incident with the point $(a + c, c - b) \in P$, so L is incident with $\geq |A|$ points. Thus, by the Szemerédi Trotter theorem we have

$$\begin{aligned} |A|^3 &= |A| \cdot |L| \\ &\leq 4\left(|A|^{\frac{4}{3}}|P|^{\frac{2}{3}} + |P| + |A|^2\right) \\ &\leq 16|A|^{\frac{4}{3}}|P|^{\frac{2}{3}}. \end{aligned}$$

Rearranging we get $\frac{1}{64}|A|^{\frac{5}{2}} \leq |P|$ which completes the proof. □