§3 Colouring (general) graphs, continued

As announced last week, we will now prove the following theorem.

**Theorem [Erdős 1959]** For every $k, \ell$ there exists a graph $G$ with $\chi(G) > k$ and girth$(G) > \ell$.

The proof is a beautiful application of the probabilistic method. In it we will consider the random graph $G(n, p)$ which is defined by taking vertex set $[n]$ and for each pair of vertices including the corresponding edge with probability $p$ (independently of the choices for other edges). I.e. we flip $\binom{n}{2}$ coins with success probability $p$ to determine the edge set.

Nowadays this is usually called the Erdős-Rényi random graph, although it was actually first defined by E.N. Gilbert in 1959 just before Erdős and Rényi published two very influential papers on the model.

In spite of its simple definition it turns out to show rather rich behaviour and is still a major topic of research. Of course, in modern random graph theory, also many more complicated and more “realistic” models are considered.

**Proof.** We pick arbitrary $k, \ell$ and we fix a constant $0 < \theta < 1/\ell$. We let $n$ be large, to be chosen more precisely later, and we set $p := n^{\theta - 1}$. Let $X$ denote the number of cycles in $G(n, p)$ of length at most $\ell$. Then

$$\mathbb{E}X \leq \sum_{i=3}^{\ell} n(n-1) \ldots (n-i+1) p^i \leq \sum_{i=3}^{\ell} (np)^i = \sum_{i=3}^{\ell} n^{\theta i} \leq \ell n^{\theta \ell} \ll n.$$  

Hence, by Markov’s inequality

$$\mathbb{P}(X > n/2) \leq \frac{\mathbb{E}X}{(n/2)^{n \to \infty}} 0. \quad (1)$$

We now set

$$x := \left\lceil \frac{100 \ln n}{p} \right\rceil.$$  

Note that $x \to \infty$ as $n \to \infty$. We have

$$\mathbb{P}(\alpha(G(n, p)) \geq x) \leq \mathbb{E}[\text{stable sets of size } x]$$

$$= \binom{n}{x} (1 - p)^{\binom{x}{2}}$$

$$\leq n^x (e - p)^{x(x-1)/2}$$

$$= \left( ne^{-p(x-1)/2} \right)^x. \quad (2)$$

(using that $1 + z \leq e^z$ in the third line)

Now note that

$$\frac{p(x-1)}{2} = \frac{(p/2) \cdot \left( \left\lfloor \frac{100 \ln n}{p} \right\rfloor - 1 \right)}{2} \geq 50 \ln n - p/2 > 49 \ln n,$$

the last inequality holding for all sufficiently large $n$. Filling this in in the last line of (2) we see that

$$\mathbb{P}(\alpha(G(n, p)) \geq x) \leq \left( n^{-48} \right)^x n \to \infty 0. \quad (3)$$

By (1) and (3), the probability that $X \leq n/2$ and $\alpha(G(n, p)) < x$ tends to one. In particular, for $n$ sufficiently large, we can find a graph $G$ with these properties (and $n$ vertices). Let us fix one such $G$. Of course this $G$ may still have cycles shorter than $\ell$.

To remedy this, we create a new graph $G^*$ by removing one vertex from each cycle in $G$. The $G^*$ we thus obtain has at least $n/2$ vertices (since $X \leq n/2$) and girth$(G^*) > \ell$. Also note that every stable set in $G^*$ is also a stable set in $G$. In other words, $\alpha(G^*) \leq \alpha(G)$. For the chromatic number of $G^*$ we thus have the lower bound
\[
\chi(G^*) \geq \frac{|V(G^*)|}{\alpha(G^*)} \geq \frac{n/2}{x} \geq \frac{n}{2(100\ln n + 1)} = \frac{n \cdot n^{\theta-1}}{200\ln n + 2p} = \frac{n^\theta}{200\ln n + 2p} \rightarrow \infty.
\]

Hence, if \( n \) was chosen sufficiently large, then \( \chi(G^*) > k \).

\section*{4. The Lovász Local Lemma}

In a previous lecture Ross has already stated, but not proved the Lovász Local Lemma (LLL). This first appeared in a paper by Erdős and Lovász in the 70s. Trivia: László Lovász rose to fame when he solved the “weak perfect graph conjecture” at the tender young age of 17. He is not to be confused with László Lovász Jr. who is also a (gifted) mathematician. The LLL is not to be confused with the LLL lattice basis reduction algorithm of Lenstra-Lenstra-Lovász (same Lovász).

In the lemma, we consider a set of probability events \( A_1, \ldots, A_n \) (defined on the same probability space). A graph \( G \) with vertex set \([n]\) is a dependency graph for these events if \( A_j \) is (mutually) independent of \( \{ A_i : ij \in E(G) \} \).

Recall that there is a (subtle) difference between pairwise independent and mutually independent. This distinction is often very important in the context of applications of the LLL. If \( A \) is an event then \( \overline{A} \) denotes its complement.

**Lemma. [LLL, general version]** If \( A_1, \ldots, A_n \) are events with dependency graph \( G \) and \( x_1, \ldots, x_n \in [0, 1) \) are such that \( \mathbb{P}(A_i) \leq x_i \prod_{ij \in E(G)} (1 - x_j) \), then

\[
\mathbb{P}(A_1 \cap \cdots \cap A_n) \geq \prod_{i=1}^n (1 - x_i).
\]

In particular, this probability is positive.

The LLL has turned out to be a very useful tool that has been successfully applied on many different problems. Often the so-called “symmetric LLL”, a corollary that is relatively easy to derive, is used instead of the general version.

**Lemma. [LLL, symmetric version]** If \( A_1, \ldots, A_n \) are events with dependency graph \( G \) and all degrees in \( G \) are \( \leq d \) and \( \mathbb{P}(A_i) \leq p \) for all \( i \), and \( ep(d + 1) \leq 1 \), then \( \mathbb{P}(\overline{A_1} \cap \cdots \cap \overline{A_n}) > 0 \).

**Proof of symmetric LLL, assuming general LLL.** We put \( x_i := \frac{1}{d+1} \) for all \( i \). Then

\[
\mathbb{P}(A_i) \leq p \leq \frac{1}{e(d+1)} \leq \frac{1}{d+1} \left(1 - \frac{1}{d+1} \right)^d \leq x_i \prod_{ij \in E} (1 - x_j).
\]

(We remark that the third inequality is not as trivial as it perhaps seems at first. To show that \( \frac{1}{e} \leq (1 - \frac{1}{d+1})^d \) for every \( d \geq 1 \) you might point out that \( (1 - \frac{1}{d+1})^d \) is decreasing, for instance by differentiating \( d \ln (1 - \frac{1}{d+1}) = d \ln d - d \ln (d+1) \), and then that \( \lim_{d \to \infty} \left(1 - \frac{1}{d+1}\right)^d = e^{-1} \) by a standard limit from first year calculus.)

Before we give the proof of the general LLL, let us illustrate the use of the symmetric LLL with one more application to Ramsey numbers.

Consider the chance experiment where we colour the edges of \( K_n \) red or blue, each with probability 1/2, independently of the choices for all other edges. We fix some \( k \leq n \). For each set \( S \) of vertices of cardinality \( k \), let \( A_S \) denote the event that \( S \) spans a monochromatic clique. (All edges between vertices in \( S \) get the same colour.) So

\[
\mathbb{P}(A_S) = 2^{1 - \binom{k}{2}} := p.
\]

Note that \( A_S, A_T \) are independent if \( |S \cap T| < 2 \). More generally, \( A_S \) is independent of \( \{ A_T : |S \cap T| < 2 \} \). We can thus define a dependency graph with \( \binom{n}{2} \) vertices and degrees \( \leq d := \binom{k}{2} \binom{n-k}{2} \).

By the symmetric LLL there exists a colouring such that no \( S \) is monochromatic if
\[ e \cdot \left( \binom{k}{2} \cdot \binom{n}{k-2} + 1 \right) \cdot 2^{1-\binom{k}{2}} < 1. \quad (4) \]

I.e. if \( n \) satisfies (4) then \( R(k, k) > n \).

**Exercise.** Use (4) to give an alternative derivation of the lower bound \( R(k, k) = \Omega \left(k \cdot 2^{k/2}\right)\). (This was already proved in Ross’ first lecture using different arguments.)

Recall \( f(m) = \Omega(g(m)) \) means that there exists a constant \( c \) such that \( f(m) \geq cg(m) \) for all \( m \).

**Hint:** You may want to use that \( \binom{n}{k} \geq \left(\frac{a}{b}\right)^b \) for \( 1 \leq b \leq a \), but in that case you have to prove that first.

---

**Proof of the general LLL.** We suppose \( A_1, \ldots, A_n \), the dependency graph \( G \) and the numbers \( x_1, \ldots, x_n \) as in the statement. The key is to establish the following claim.

**Claim.** \( \forall i, S \ni i : \mathbb{P}(A_i \mid \bigcap_{j \in S} \overline{A}_j) \leq x_i \).

Let us first show how the claim implies the lemma. We have

\[
\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 \mid A_1) \cdot \mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1})
\]

\[
\geq (1 - \mathbb{P}(A_1)) \cdot (1 - \mathbb{P}(A_2 \mid A_1)) \cdots (1 - \mathbb{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1}))
\]

It remains to prove the claim, which we will do by induction on \( s = |S| \). The base case, when \( s = 0 \), is trivial. Let us thus assume \( s \geq 1 \) and the claim holds for all \( s' < s \). We write

\[
S_1 := \{ j \in S : i j \in E(G) \}, \quad S_2 := S \setminus S_1.
\]

If \( S_1 = \emptyset \) then we are done since

\[
\mathbb{P}(A_i \cap \bigcap_{j \in S} \overline{A}_j) = \mathbb{P}(A_i) \leq x_i \prod_{j \in E} (1 - x_j) \leq x_i.
\]

We thus assume \( S_1 \neq \emptyset \). Note that

\[
\mathbb{P}(A_i \mid \bigcap_{j \in S} \overline{A}_j) = \frac{\mathbb{P}(A_i \cap \bigcap_{j \in S} \overline{A}_j)}{\mathbb{P}(\bigcap_{j \in S} \overline{A}_j)}.
\]

(5)

(as can be seen by writing out the definition of conditional probability.)

Using that \( A_i \) is independent of \( A_j : j \in S_2 \) we see that the numerator of the RHS of (5) satisfies

\[
\mathbb{P}(A_i \cap \bigcap_{j \in S_1} \overline{A}_j \cap \bigcap_{j \in S_2} \overline{A}_j) \leq \mathbb{P}(A_i) \prod_{j \in S_2} (1 - x_j).
\]

(6)

We now consider the denominator of the RHS of (5). Let us write \( S = \{ j_1, \ldots, j_k \} \). We find that

\[
\mathbb{P}(\bigcap_{j \in S_1} \overline{A}_j \cap \bigcap_{j \in S_2} \overline{A}_j) = \mathbb{P}(\bigcap_{j \in S_1} \overline{A}_j \cap \bigcap_{j \in S_2} \overline{A}_j) \geq \prod_{j \in E} (1 - x_j).
\]

(7)

Combining (5), (6) and (7) we find

\[
\mathbb{P}(A_i \mid \bigcap_{j \in S} \overline{A}_j) \leq \frac{x_i \prod_{j \in E} (1 - x_j)}{\prod_{j \in E} (1 - x_j)} = x_i,
\]

as claimed. Thus, the claim and thereby also the LLL is proved.