Extra Notes 4 (3/5 and 7/5)

Theorem E.13. *Let $\varepsilon > 0$. Let E_n be the event that $G(n,\, 1/2)$ contains a clique of size at least $(2+\varepsilon)\log_2 n$. Then $\mathbb{P}(E_n) \to 0$ as $n \to \infty$.*

Proof: Let $k = \lceil (2+\varepsilon)\log_2 n \rceil$. Since $\frac{(2+\varepsilon)^2}{2} = 2 + 2\varepsilon + \frac{\varepsilon^2}{2} > 2 + 2\varepsilon$, if n is sufficiently large then $\frac{k(k-1)}{2} > (2+2\varepsilon)(\log_2 n)^2$. We have

$$E_n = \bigcup_{i=1}^{\binom{n}{k}} E_{n,\,i},$$

where the $\binom{n}{k}$ subsets of $V(G(n,\, 1/2))$ are ordered arbitrarily, and $E_{n,\,i}$ is the event that the subgraph induced by the i:th subset is a clique. By MP,

$$\mathbb{P}(E_{n,\,i}) = \left(\frac{1}{2}\right)^{\binom{k}{2}}$$

and hence, when n is sufficiently large,

$$\mathbb{P}(E_{n,\,i}) < 2^{-(2+2\varepsilon)(\log_2 n)^2}.$$

Thus, for n sufficiently large,

$$\mathbb{P}(E_n) \le \sum_{i=1}^{\binom{n}{k}} \mathbb{P}(E_{n,\,i}) < \binom{n}{k} 2^{-(2+2\varepsilon)(\log_2 n)^2} <$$

$$< n^k (2^{\log_2 n})^{-(2+2\varepsilon)(\log_2 n)} \le n^{1+(2+\varepsilon)(\log_2 n)} n^{-(2+2\varepsilon)(\log_2 n)} = n^{1-\varepsilon(\log_2 n)},$$

which goes to zero as $n \to \infty$, for any fixed $\varepsilon > 0$.

Corollary E.14. *Let $\varepsilon > 0$. Then, as $n \to \infty$,*

$$\mathbb{P}(\omega(G(n,\, 1/2)) > (2+\varepsilon)(\log_2 n)) \to 0, \tag{0.13}$$

$$\mathbb{P}\left(\chi(G(n,\, 1/2)) < \frac{n}{(2+\varepsilon)(\log_2 n)}\right) \to 0. \tag{0.14}$$

Proof: Eq. (0.13) is an immediate consequence of Theorem E.13. For (0.14), first observe that Proposition 16.12 implies that

$$\chi(G(n,\, 1/2)) < \frac{n}{(2+\varepsilon)(\log_2 n)} \Leftrightarrow \alpha(G(n,\, 1/2)) > (2+\varepsilon)(\log_2 n) \Leftrightarrow \omega(\overline{G(n,\, 1/2)}) > (2+\varepsilon)(\log_2 n).$$

But the random graph $G(n,\, 1/2)$ and its complement have the same distribution (since each edge appears with probability $1/2$ in each, independent of all other edges), which implies that the events $\omega(G(n,\, 1/2)) > (2+\varepsilon)(\log_2 n)$ and $\omega(\overline{G(n,\, 1/2)}) > (2+\varepsilon)(\log_2 n)$ have the same probability. The probability of the former goes to zero, by (0.13), so we're done.

Full proof of Theorem 16.8. The idea here is once again to apply a probabilistic method to the random graph $G(n,\, p)$, but the novelty in this proof is that p will depend on n. Before giving the proof, we will need some terminology and lemmas. Note that

Definitions E.15 - E.18 and Lemmas E.19, E.20 can be extended beyond finite sets, as would be the case in any standard text on probability. It suffices for our purposes to consider the simplest case of finite sets, however (this is not a course in probability, after all !).

Definition E.15. A *probability space* is a pair $(\Omega,\, \mu)$ where Ω is a finite set and $\mu\,:\,\Omega \to \mathbb{R}_+$ is a non-negative function such that $\sum_{\omega\in\Omega}\mu(\omega) = 1$. Usually μ is called a *probability measure/distribution* on the set Ω.

Definition E.16. Let $(\Omega,\, \mu)$ be a probability space. A function $X\,:\,\Omega \to \mathbb{R}$ is called a *(real-valued) random variable on* Ω. Even though, formally, a random variable is just a function, it is common to use letters $X,\, Y, \ldots$ for random variables, instead of $f,\, g, \ldots$, in order to reflect the "randomness".

Definition E.17. Let $(\Omega,\, \mu)$ be a probability space, X a random variable on Ω and $A \subseteq \mathbb{R}$. The *probability of the event "$X \in A$"* is given by

$$\mathbb{P}(X \in A) := \sum_{\omega\in\Omega} \mu(w) \cdot \delta(\omega), \quad \text{where } \delta(\omega) = \begin{cases} 1, & \text{if } X(\omega) \in A, \\ 0, & \text{if } X(\omega) \notin A. \end{cases}$$

Definition E.18. Let $(\Omega,\, \mu)$ be a probability space and X a random variable on Ω. The *expected/average value* of X, denoted $\mathbb{E}[X]$, is given by

$$\mathbb{E}[X] := \sum_{\omega\in\Omega} X(\omega) \cdot \mu(\omega).$$

Lemma E.19. (Linearity of Expectation) *Let $X_1,\, \ldots,\, X_n$ be random variables on the same probability space $(\Omega,\, \mu)$ and let $X = \sum_{i=1}^{n} X_i$ be their pointwise sum (as functions). Then*

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]. \tag{0.15}$$

Proof: By Definition E.18,

$$\begin{aligned} \mathbb{E}[X] &= \sum_{\omega\in\Omega} X(\omega) \cdot \mu(\omega) = \sum_{\omega\in\Omega} \left(\sum_{i=1}^{n} X_i(\omega) \right) \cdot \mu(\omega) = \\ &= \sum_{i=1}^{n} \left(\sum_{\omega\in\Omega} X_i(\omega) \cdot \mu(\omega) \right) = \sum_{i=1}^{n} \mathbb{E}[X_i], \quad \text{v.s.v.} \end{aligned}$$

Lemma E.20. (Markov's Inequality) *Let X be a non-negative valued random variable on a probability space $(\Omega,\, \mu)$. Then, for any $\lambda > 0$,*

$$\mathbb{P}(X > \lambda \cdot \mathbb{E}[X]) \leq \frac{1}{\lambda}. \tag{0.16}$$

Proof: First note that the inequality is trivial if $\mathbb{E}[X] = 0$, since in that case X, being non-negative, must be identically zero (i.e.: $\mathbb{P}(X > 0) = 0$).

So we may assume that $\mathbb{E}[X] \neq 0$. Let $A := (\lambda \cdot \mathbb{E}[X], \infty)$, so we must show that $\mathbb{P}(X \in A) \leq 1/\lambda$. We have

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega) \cdot \mu(\omega) = \sum_{\omega : X(\omega) \leq \lambda \cdot \mathbb{E}[X]} X(\omega)\mu(\omega) + \sum_{\omega : X(\omega) > \lambda \cdot \mathbb{E}[X]} X(\omega) \cdot \mu(\omega).$$

Since X is non-negative valued, the first sum is, at the very least, non-negative. In the second sum, $X(\omega) > \lambda \cdot \mathbb{E}[X]$ for each ω by definition. Hence the second sum is at least $\lambda \cdot \mathbb{E}[X] \cdot \sum_{\omega : X(\omega) > \lambda \cdot \mathbb{E}[X]} \mu(\omega) = \lambda \cdot \mathbb{E}[X] \cdot \mathbb{P}(X \in A)$. Thus, we've shown that

$$\mathbb{E}[X] \geq 0 + \lambda \cdot \mathbb{E}[X] \cdot \mathbb{P}(X \in A) \cdots \Rightarrow \cdots \mathbb{P}(X \in A) \leq \frac{1}{\lambda}, \quad \text{v.s.v.}$$

Proof of Theorem 16.8: Fix a positive integer t and fix a real number θ such that $0 < \theta < 1/t$. We consider $G(n, p)$ with $p = n^{\theta - 1}$. Roughly speaking, the proof consists of three parts. In the first part, we show that $G(n, p)$ contains relatively few cycles of length at most t with high probability (that is, with probability tending to 1 as $n \to \infty$). In the second part, we show that $G(n, p)$ also has low independence number, and hence high chromatic number, with high probability. Hence, both of the events described in parts 1 and 2 occur simoultaneously with high probability, which proves that some graph G satisfying both conditions exists. Finally, in part 3, we show that some modification G^* of G satisfies the exact statement of the theorem.

PART I: Let $X = X_{n,p}$ be the number of simple cycles of length at most t in $G(n, p)$. Thus, formally speaking, X is a random variable on the probability space $G(n, p)$. We need an upper bound for $\mathbb{E}[X]$. Firstly,

$$X = \sum_{i=3}^{t} X_i, \tag{0.17}$$

where X_i is the number of simple cycles of length i in $G(n, p)$. Let n_i be the number of simple cycles of length i in the complete graph K_n. Then in turn,

$$X_i = \sum_{j=1}^{n_i} X_{i,j}, \tag{0.18}$$

where the n_i possible simple cycles have been ordered arbitrarily and $X_{i,j}$ is the so-called *indicator variable* of the event that the j:th possible simple cycle is present in $G(n, p)$ - that is,

$$X_{i,j} = \begin{cases} 1, & \text{if the } j\text{:th simple cycle is present in } G(n, p), \\ 0, & \text{otherwise.} \end{cases} \tag{0.19}$$

Applying linearity of expectation to both (0.17) and (0.18), we have

$$\mathbb{E}[X] = \sum_{i=3}^{t} \sum_{j=1}^{n_i} \mathbb{E}[X_{i,j}]. \tag{0.20}$$

Firstly, since each $X_{i,j}$ is an indicator variable,

$$\mathbb{E}[X_{i,j}] = \mathbb{P}(X_{i,j} = 1) = \mathbb{P}[j\text{th simple cycle of length } i \text{ present}] = p^i = n^{(\theta-1)i}, \tag{0.21}$$

where the second-last equality is because a simple cycle of length i contains i different edges and, by definition of $G(n, p)$, each edge is present, independently of all others, with probability p.

Furthermore,

$$n_i = \frac{P(n, i)}{2i} = \frac{n(n-1)\dots(n-i+1)}{2i} < n^i, \tag{0.22}$$

where the first equality comes from the facts that:

- there are $P(n, i)$ possible choices of an ordered sequence of i vertices, and each such choice defines a simple cycle of length i,
- for each simple cycle, there are $2i$ choices of the ordered sequence of vertices, corresponding to i possible start/end points and two possible orientations.

Substituting (0.21) and (0.22) into (0.20) we find that

$$\mathbb{E}[X] < \sum_{i=3}^{t} n^{\theta i}. \tag{0.23}$$

For any fixed t, since $\theta t < 1$ we have that, for any $\varepsilon > 0$,

$$\mathbb{E}[X] < \varepsilon n, \quad \text{for all sufficiently large } n. \tag{0.24}$$

Hence, by Markov's inequality, for any $\varepsilon > 0$,

$$\mathbb{P}(X < \varepsilon n) \to 1, \quad \text{as } n \to \infty. \tag{0.25}$$

PART II: For any $x \in \mathbb{N}$, let F_x denote the event that $G(n, p)$ contains an independent set of size x. Then

$$F_x = \bigcup_{j=1}^{\binom{n}{x}} F_{x,j}, \tag{0.26}$$

where the x-element subsets of the n vertices have been ordered arbitrarily and $F_{x,j}$ is the event that the j:th subset forms an independent set. By definition of $G(n, p)$ we have, for each j, that

$$\mathbb{P}(F_{x,j}) = (1-p)^{\binom{x}{2}}. \tag{0.27}$$

Note that, for any $p \in (0, 1)$, $1 - p < e^{-p}$. Substituting this and (0.27) into (0.26) and using just a union bound we find that

$$\mathbb{P}(F_x) \le \binom{n}{x} e^{-p\frac{x(x-1)}{2}} < \left(n\, e^{-\frac{p(x-1)}{2}}\right)^x. \tag{0.28}$$

Now take $x = \lceil \frac{2+\varepsilon}{p} \ln n \rceil + 1$, for any fixed $\varepsilon > 0$. Then

$$e^{-\frac{p(x-1)}{2}} \le e^{-\frac{p}{2}\frac{2+\varepsilon}{p}\ln n} = \left(e^{\ln n}\right)^{-\frac{2+\varepsilon}{2}} = n^{-\left(1+\frac{\varepsilon}{2}\right)}.$$

Substituting into (0.28) we have

$$\mathbb{P}(F_x) \le (n \cdot n^{-\left(1+\frac{\varepsilon}{2}\right)})^x = n^{-\frac{\varepsilon x}{2}},$$

which obviously goes to zero as $n \to \infty$.So as $n \to \infty$, the probability that $G(n,\, p)$ contain an independent set of size $\lceil \frac{2+\varepsilon}{p} \ln n \rceil + 1$ tends to zero. In other words, for any $\varepsilon > 0$,

$$\mathbb{P}\left(\alpha(G(n,\, p)) > \lceil \frac{2+\varepsilon}{p} \ln n \rceil \right) \to 0, \;\; \text{as } n \to \infty. \tag{0.29}$$

PART III: Fix $0 < \varepsilon < 1$. From (0.25) and (0.29) it follows that, for any given t and $\theta \in (0,\, 1/t)$, there exists, for all sufficiently large n, a graph $G = G_n$ on n vertices satisfying the following two conditions:

(a) G has at most εn simple cycles of length at most t,
(b) $\alpha(G) \leq (2+\varepsilon)n^{1-\theta}(\ln n)$.

Take such a G, pick any vertex from each of its simple cycles of length at most t and remove all these vertices from G, along with all their adjacent edges. Let G^* be the remaining graph (note that G^* could be disconnected, even if G were originally connected). Since we've removed at most εn vertices, G^* has at least $n(1-\varepsilon)$ vertices. It has girth strictly greater than t, by construction. Any independent set of vertices in G^* was already independent in G, hence $\alpha(G^*) \leq \alpha(G)$. But then, by Proposition 16.12(ii),

$$\chi(G^*) \geq \frac{|V(G^*)|}{\alpha(G^*)} \geq \frac{n(1-\varepsilon)}{(2+\varepsilon)n^{1-\theta}\ln n} = \frac{1-\varepsilon}{2+\varepsilon}\frac{n^{\theta}}{\ln n}.$$

Since $\theta > 0$, for any fixed $\varepsilon \in (0,\, 1)$ this goes to infinity as n does so. In particular, for n sufficiently large, it will be greater than t. So we have proven that, for n sufficiently large, the graph $G^* = G_n^*$ satisfies

$$\min\{\chi(G^*),\, \text{girth}(G^*)\} > t, \quad \text{v.s.v.}$$

Remark on Theorem 16.13. It is worth noting the following:

For any graph G, there exists an ordering of its vertices for which the greedy algorithm would use exactly $\chi(G)$ colors.

Proof: Consider any $\chi(G)$-coloring, and label the colors as $1, , 2, \ldots, \chi(G)$, in any order. Let $c(v)$ denote the color assigned to vertex v. Thus c is a function $c : V(G) \to \{1,\, 2,\, \ldots,\, \chi(G)\}$. Now order the vertices in such a way that

$$i \leq j \text{ if and only if } c(v_i) \leq c(v_j).$$

In other words, first write down the vertices that get color 1, in any internal order, then write down those that get color 2 and so on.

If we apply the greedy algorithm to this ordering, then the color it assigns to a vertex v will always be less than or equal to $c(v)$ since the vertices that are assigned a given color in the coloring c form an independent set, so there will be no reason for the greedy algorithm to seek a higher-numbered color when coloring them. Thus the greedy algorithm will use at most $\chi(G)$ colors, hence exactly this number of colors, by definition of $\chi(G)$.