# Chapter 3

# Erdős–Rényi random graphs

## 3.1 Definitions

Fix $n$, consider the set $V = \{1, 2, \ldots, n\} =: [n]$, and put $N := \binom{n}{2}$ be the number of edges on the full graph $K_n$, the edges are $\{e_1, e_2, \ldots, e_N\}$. Fix also $p \in [0, 1]$ and choose edges according to Bernoulli's trials: an edge $e_i$ is picked with probability $p$ independently of occurrences of other edges. What we obtain as a result of this procedure is usually called the Erdős–Rényi random graph and will be denoted as $\mathcal{G}(n, p)$. The words "random graph" is a misnomer, since actually we are dealing with a probability space $\mathcal{G}(n, p) = (\Omega, \mathcal{F}, \mathrm{P})$, where $\Omega$ is the sample space of all possible graphs on $n$ vertices, $|\Omega| = 2^N = 2^{\binom{n}{2}}$, P is the probability measure that for each graph $G \in \Omega$ assigns probability

$$\mathsf{P}(G) = p^m (1 - p)^{N-m},$$

where $m$ is the number of edges in $G$, and $\mathcal{F}$ are the events, which are natural to interpret in these settings as *graph properties*. For instance, if $A$ is a property that graph is connected then

$$\mathsf{P}(A) = \sum_{G \in A} \mathsf{P}(G),$$

where the summation is through all the connected graphs in $\mathcal{G}(n, p)$.

Another convention is that while talking about some characteristics of a random graph, it is usually meant the average across all ensemble of outcomes. For example, while talking about the clustering coefficient of $\mathcal{G}(n, p)$, it is usually meant

$$C = \frac{\mathsf{E}(\#\{\text{closed paths of length 2}\})}{\mathsf{E}(\#\{\text{paths of length 2}\})},$$

where $\#\{\text{closed paths of length 2}\}$ and $\#\{\text{paths of length 2}\}$ are random variables defined on $\mathcal{G}(n, p)$.

An alternative approach is to fix $n$ and $m$ at the very beginning, where $n$ is the order of the graph, and $m$ is the size of the graph, and pick any of possible graphs on $n$ labeled vertices with exactly $m$ edges with equal probabilities. We thus obtain $\mathcal{G}(n, m) = (\Omega, \mathcal{F}, \mathrm{P})$, where now $|\Omega| = \binom{N}{m}$, and

$$\mathsf{P}(G) = \frac{1}{\binom{N}{m}}.$$

This second approach was used initially by Erdős and Rényi, but the random graph $\mathcal{G}(n, p)$ is somewhat more amenable to analytical investigation due to the independence of edges in the graph. This is not true, obviously, for $\mathcal{G}(n, m)$. Actually, there is a very close connection between the two, and the properties of $\mathcal{G}(n, m)$ and $\mathcal{G}(n, p)$ are very similar in the case $m = \binom{n}{2}p$.

Here are some simple properties of $\mathcal{G}(n, p)$:

- The mean number of edges in $\mathcal{G}(n, p)$ is $\binom{n}{2}p$. We can arrive to this results by noting that the distribution of the number of edges $X$ in $\mathcal{G}(n, p)$ is binomial with parameters $N$ and $p$:

  $$\mathsf{P}(X = k) = \binom{N}{k} p^k (1 - p)^{N-k},$$

  and recalling the formula for the expectation of the binomial random variable. A more straightforward approach is to consider the event $A = \{e \in G\}$, i.e., edge $e$ belongs to the graph $G \in \mathcal{G}(n, p)$. Naturally, the number of edges in $G$ is given by

  $$X = \sum_{e \in K_n} \mathbf{1}_{\{e \in G\}}.$$

  Now apply expectations to the both sides, use the linearity of expectation, the fact that $\mathsf{E}\,\mathbf{1}_A = \mathsf{P}(A)$ and $\mathsf{P}(\{e \in G\}) = p$, and that $|E(K_n)| = \binom{n}{2}$ to get the same conclusion.

- The degree distribution for any vertex in $\mathcal{G}(n,p)$ is binomial with parameters $n-1$ and $p$. I.e., if $D_i$ is the random variable denoting the degree of the vertex $i$, then

$$\mathsf{P}(D_i = k) = \binom{n-1}{k} p^k (1-p)^{k-1},$$

since each vertex can be connected to $n-1$ other vertices. Note that for two different vertices $i$ and $j$ the random variables $D_i$ and $D_j$ are not exactly independent (e.g., if $D_i = n-1$ then obviously $D_j \neq 0$), however, for large enough $n$, they are almost independent and we can assume that the degree distribution of the random graph $\mathcal{G}(n,p)$ approaching the binomial distribution. Since the binomial distribution can be approximated by the Poisson distribution in the case $np \to \lambda$, so we set $p = \frac{\lambda}{n}$ and state the final result that the degree distribution of the Erdős–Rényi random graph is Poisson with parameter $\lambda$:

$$\mathsf{P}(X = k) = \frac{\lambda^k}{k!} e^{-k}.$$

That is why $\mathcal{G}(n,p)$ is sometimes called Poisson random graph.

- The clustering coefficient of $\mathcal{G}(n,p)$ can be formally calculated as

$$C = \frac{\mathsf{E}(\#\{\text{closed paths of length 2}\})}{\mathsf{E}(\#\{\text{paths of length 2}\})} = \frac{\binom{n}{3} p^3 \cdot 6}{\binom{n}{3} p^2 \cdot 3 \cdot 2} = p,$$

where $\binom{n}{3}$ is the number of triples in $n$ vertices. However, even without this formal derivation, recall that the clustering coefficient describes how many triangles in the network. In $\mathcal{G}(n,p)$ model this is equivalent to the fact how often the path of length 2 is closed, and this is exactly $p$, the probability of having edge between two vertices. Note that if we consider $p = \frac{\lambda}{n}$ and let $n \to \infty$, then $C \to 0$.

**Problem 3.1.** What is the mean number of squares in $\mathcal{G}(n,p)$? You probably should start solving this problem by answering the question: How many different necklaces can be made out of $i$ stones.

Our random graphs are considered as models of real-world complex networks, which are usually growing. Hence it is natural to assume that $n \to \infty$. In several cases we already suggested the assumption that $p$ should depend on $n$ and approach zero as $n$ grows. Here is one more reason for this: The growing sequence of random graphs $\mathcal{G}(n,p)$ for fixed $p$ is not particularly interesting.

**Problem 3.2.** Prove that for constant $p$ $\mathcal{G}(n,p)$ is connected whp.

**Theorem 3.1.** *If $p$ fixed $\mathcal{G}(n,p)$ has diameter 2 whp.*

*Proof.* Consider the random variable $X_n$ which is the number of vertex pairs in the graph $G \in \mathcal{G}(n,p)$ on $n$ vertices with no common neighbors. To prove the theorem, we have to show that

$$\mathsf{P}(X_n = 0) \to 1, \quad n \to \infty.$$

Or, switching to the complementary event,

$$\mathsf{P}(X_n \geq 1) \to 0, \quad n \to \infty.$$

We have

$$\mathsf{P}(X_n \geq 1) \leq \mathsf{E}X_n \quad \text{by Markov's inequality.}$$

Now consider

$$X_n = \sum_{u,v \in V} \mathbf{1}_{\{u,\, v \text{ have no common neighbor}\}},$$

apply the expectation

$$\mathsf{E}X_n = \sum_{u,v \in V} \mathsf{P}(\{u, v \text{ have no common neighbor}\}) = \binom{n}{2} (1-p^2)^{n-2},$$

which approaches zero as $n \to \infty$. ■

## 3.2 Triangles in Erdős–Rényi random graphs

Before turning to the big questions about the Erdős–Rényi random graphs, let us consider a toy example, which, however shows the essence of what is usually happening in these random graphs.

Denote $T_{3,n}$ the random variable on the space $\mathcal{G}(n,p)$, which is equal to the number of triangles in a random graph. For example, $T_{3,n}(K_n) = \binom{n}{3}$, and for any graph $G$ with only two edges $T_{3,n}(G) = 0$. First I will establish the conditions when there are no triangles whp. I will use the same method that I used implicitly in proving Theorem 3.1, which is called the first moment method, and which can be formally formulated as follows.

**Theorem 3.2** (First moment method)**.** *Let $X_n \geq 0$ be an integer valued random variable. If $\mathsf{E}X_n \to 0$ then $X_n = 0$ whp as $n \to \infty$.*

*Proof.* By Markov's inequality $\mathsf{P}(X \geq 1) \leq \mathsf{E}X$, hence the theorem. ■

**Theorem 3.3.** *Let $\alpha\colon \mathbf{N} \longrightarrow \mathbf{R}$ be a function such that $\alpha(n) \to 0$ as $n \to \infty$; let $p(n) = \frac{\alpha(n)}{n}$ for each $n \in \mathbf{N}$. Then $T_{3,n} = 0$ whp.*

*Proof.* The goal is to show that

$$\mathsf{P}(T_{3,n} = 0) \to 1,$$

as $n \to \infty$. This is the same as $\mathsf{P}(T_{3,n} \geq 1) \to 0$. According to Markov's inequality

$$\mathsf{P}(T_{3,n} \geq 1) \leq \mathsf{E}(T_{3,n}).$$

Let us estimate $\mathsf{E}(T_{3,n})$. For each fixed $n$ the random variable $T_{3,n}$ can be represented as

$$T_{3,n} = \mathbf{1}_{\tau_1} + \ldots + \mathbf{1}_{\tau_k}, \quad k = \binom{n}{3},$$

where $\tau_i$ is the event that the $i$th triple of vertices from the set of all vertices of $\mathcal{G}(n,p)$ forms a triangle. Here we assume that all possible triples are ordered and labeled. Using the linearity of the expectation

$$\mathsf{E}(T_{3,n}) = \mathsf{E}\,\mathbf{1}_{\tau_1} + \ldots + \mathsf{E}\,\mathbf{1}_{\tau_k} = \mathsf{P}(\tau_1) + \ldots + \mathsf{P}(\tau_k) = \binom{n}{3} p^3,$$

since $\mathsf{P}(\tau_i) = p^3$ in the Erdős–Rényi random graphs $\mathcal{G}(n,p)$.

Finally, we have

$$\mathsf{E}(T_{3,n}) = \binom{n}{3} p^3 = \frac{n!}{(n-3)!3!} \frac{\alpha^3(n)}{n^3} = \frac{n(n-1)(n-2)\alpha^3(n)}{6n^3} \sim \frac{\alpha^3(n)}{6} \to 0.$$

■

Now I will establish the conditions when the Erdős–Rényi random graphs have triangles almost always. For this I will use the second moment method.

**Theorem 3.4** (The second moment method)**.** *Let $X_n \geq 0$ be an integer valued random variable. If $\mathsf{E}X_n > 0$ for $n$ large and $\operatorname{Var} X_n/(\mathsf{E}X_n)^2 \to 0$ then $X_n > 0$ whp.*

*Proof.* By Chebyshev's inequality $\mathsf{P}(|X - \mathsf{E}X| \geq \mathsf{E}X) \leq \operatorname{Var} X/(\mathsf{E}X)^2$, from where the result follows. ■

**Theorem 3.5.** *Let $\omega\colon \mathbf{N} \to \mathbf{R}$ be a function such that $\omega(n) \to \infty$ as $n \to \infty$; let $p(n) = \frac{\omega(n)}{n}$ for each $n \in \mathbf{N}$. Then $T_{3,n} \geq 1$ a.a.s.*

*Proof.* Here we start with Chebyshev's inequality

$$\begin{aligned}\mathsf{P}(T_{3,n}=0) &= \mathsf{P}(T_{3,n}\le 0)\\ &= \mathsf{P}(-T_{3,n}\ge 0)\\ &= \mathsf{P}(\mathsf{E}\,T_{3,n}-T_{3,n}\ge \mathsf{E}\,T_{3,n})\\ &\le \mathsf{P}(|\,\mathsf{E}\,T_{3,n}-T_{3,n}|\ge \mathsf{E}\,T_{3,n})\\ &\le \frac{\operatorname{Var}T_{3,n}}{(\mathsf{E}\,T_{3,n})^2}.\end{aligned}$$

From Theorem 3.3 we already know $\mathsf{E}\,T_{3,n}\sim\frac{\omega^3(n)}{6}$. To find $\operatorname{Var}T_{3,n}=\mathsf{E}\,T_{3,n}^2-(\mathsf{E}\,T_{3,n})^2$ note, using the same notations as before, that

$$\begin{aligned}\mathsf{E}(T_{3,n}^2) &= \mathsf{E}(\,\mathbf{1}_{\tau_1}+\ldots+\,\mathbf{1}_{\tau_k})^2 =\\ &= \mathsf{E}\,\mathbf{1}_{\tau_1}^2+\ldots+\mathsf{E}\,\mathbf{1}_{\tau_k}^2+\sum_{i\ne j}\mathsf{E}(\,\mathbf{1}_{\tau_i}\,\mathbf{1}_{\tau_j}) =\\ &= \sum_i \mathsf{E}\,\mathbf{1}_{\tau_1}+\sum_{i\ne j}\mathsf{E}(\,\mathbf{1}_{\tau_i}\,\mathbf{1}_{\tau_j}).\end{aligned}$$

Here the sum in the second term is taken through all the ordered pairs of $i\ne j$, hence there is no "2" in the expression. Recall that $\mathsf{E}(\,\mathbf{1}_{\tau_i}\,\mathbf{1}_{\tau_j})=\mathsf{P}(\tau_i\cap\tau_j)$ is the probability that both triples of the vertices number $i$ and number $j$ belong to $\mathcal{G}(n,p)$. If $\tau_i\cap\tau_j=\emptyset$ then $\mathsf{P}(\tau_i\cap\tau_j)=p^6$; if $\tau_i$ and $\tau_j$ have only one vertex in common then $\mathsf{P}(\tau_i\cap\tau_j)=p^6$, and if they have two vertices in common then $\mathsf{P}(\tau_i\cap\tau_j)=p^5$ (draw some examples to convince yourself). The total number of the pairs of triples $i$ and $j$ with no common vertices is $\binom{n}{3}\binom{n-3}{3}$, with one common vertex is $3\binom{n}{3}\binom{n-3}{2}$, with two common vertices is $3\binom{n}{3}\binom{n-3}{1}$. Summing,

$$\sum_{i\ne j}\mathsf{E}(\,\mathbf{1}_{\tau_i}\,\mathbf{1}_{\tau_j})=\sum_{i\ne j}\mathsf{P}(\tau_i\cap\tau_j)=\binom{n}{3}\binom{n-3}{3}p^6+3\binom{n}{3}\binom{n-3}{2}p^6+3\binom{n}{3}\binom{n-3}{1}p^5.$$

Using the facts that

$$\binom{n}{3}\sim\binom{n-3}{3}\sim\frac{n^3}{6},\quad \binom{n-3}{2}\sim\frac{n^2}{2},\quad n-3\sim n,$$

we find that $\sum_{i\ne j}\mathsf{E}(\,\mathbf{1}_{\tau_i}\,\mathbf{1}_{\tau_j})=(1+o(1))(\mathsf{E}\,T_{3,n})^2$ and $\mathsf{E}\,T_{3,n}\to\infty$, hence,

$$\frac{\operatorname{Var}T_{3,n}}{(\mathsf{E}\,T_{3,n})^2}=\frac{1}{\mathsf{E}\,T_{3,n}}+\frac{\sum_{i\ne j}\mathsf{E}(\,\mathbf{1}_{\tau_i}\,\mathbf{1}_{\tau_j})-(\mathsf{E}\,T_{3,n})^2}{(\mathsf{E}\,T_{3,n})^2}\to 0.$$

■

**Problem 3.3.** Fill in the details of the estimates in the last part of the proof of Theorem 3.5.

Finally, let me tackle the border line case. For this I will use the *method of moments* and the notion of a factorial moment of a random variable $X$ of order $r$ which is defined as $\mathsf{E}\big(X(X-1)\ldots(X-r+1)\big)=:\mathsf{E}(X)_r$.

**Problem 3.4.** Show that if $X$ is a Poisson random variable with parameter $\lambda$ then $\mathsf{E}(X)_r=\lambda^r$.

**Theorem 3.6.** *Let $X_n$ be a sequence on non-negative integer values random variables. Let $\mathsf{E}(X_n)_r\sim\lambda^r$ for any $r$ for $n\to\infty$, where $\lambda>0$ is some constant. Then*

$$\mathsf{P}(X_n=k)\sim\frac{\lambda^k e^{-\lambda}}{k!}\,.$$

**Theorem 3.7.** *Let $p(n) \sim \frac{c}{n}$ for some constant $c > 0$. Then the random variable $T_{3,n}$ converges in distribution to the random variable $T_3$ that has a Poisson distribution with the parameter $\lambda = \frac{c^3}{6}$.*

*Proof.* To prove this theorem we will use the method of moments. First we note that in the proof of Theorem 3.5 we could use the notation

$$\mathsf{E}\,(T_{3,n})_2 = \mathsf{E}\big(T_{3,n}(T_{3,n}-1)\big) = \sum_{i\neq j} \mathsf{E}(\,\mathbf{1}_{\tau_i}\,\mathbf{1}_{\tau_j}),$$

for the second factorial moment of $T_{3,n}$. Recall that we use $(x)_r = x(x-1)\dots(x-r+1)$. Similarly,

$$\mathsf{E}\,(T_{3,n})_3 = \mathsf{E}\big(T_{3,n}(T_{3,n}-1)(T_{3,n}-2)\big) = \sum_{i\neq j\neq l} \mathsf{E}(\,\mathbf{1}_{\tau_i}\,\mathbf{1}_{\tau_j}\,\mathbf{1}_{\tau_l}),$$

where the summation is along all ordered triples $i \neq j \neq l$ (prove it).

In general, one has

$$\mathsf{E}\,(T_{3,n})_r = \mathsf{E}\big(T_{3,n}(T_{3,n}-1)\dots(T_{3,n}-r+1)\big) = \sum_{i_1\dots i_r} \mathsf{E}(\,\mathbf{1}_{\tau_{i_1}}\dots\,\mathbf{1}_{\tau_{i_r}}). \tag{3.1}$$

Using the proofs of Theorems 3.3 and 3.5 we can conclude that under the hypothesis of the theorem

$$\mathsf{E}\,T_{3,n} \to \lambda, \quad n\to\infty,$$
$$\mathsf{E}\,(T_{3,n})_2 \to \lambda^2, \quad n\to\infty.$$

To prove the theorem we need to show that

$$\mathsf{E}\,(T_{3,n})_r \to \lambda^r, \quad n\to\infty,$$

for any fixed $r$. This, according to the method of moments, would mean that

$$T_{3,n} \xrightarrow{d} T_3,$$

and $T_3$ has a Poisson distribution with the parameter $\lambda$.

For each tuple $i_1,\dots,i_r$ the events $\tau_{i_1},\dots,\tau_{i_r}$ can be classified as such that 1) there are no common vertices for any triples and 2) there is at least one vertex that belongs to at least two events at the same time (cf. proof of Theorem 3.5). Denote the sum of probabilities of the first type events as $\Sigma_1$, and for the second type as $\Sigma_2$. Two facts we will show are 1) $\Sigma_1 \sim \lambda^r$, 2) $\Sigma_2 = o(\Sigma_1) = o(1)$.

We have, assuming that $3r \le n$, that

$$\Sigma_1 = \binom{n}{3}\binom{n-3}{3}\cdots\binom{n-3(r-1)}{3}p^{3r},$$

which means that

$$\frac{\Sigma_1}{\lambda^r} \to 1,$$

since $\lambda^r \sim \left(\binom{n}{3}p^3\right)^r$ (fill in the details).

Represent $\Sigma_2$ as

$$\Sigma_2 = \sum_{s=4}^{3r-1} \Sigma^s,$$

where $\Sigma^s$ gives the total contribution of tuples $i_1,\dots,i_r$ such that $|\tau_{i_1}\cup\ldots\cup\tau_{i_r}| = s$. The total number $t$ of edges of the triangles generated by $\tau_{i_1},\dots,\tau_{i_r}$ is always strictly bigger than $s$ (give examples), hence

$$\mathsf{E}(\,\mathbf{1}_{\tau_{i_1}}\dots\,\mathbf{1}_{\tau_{i_r}}) = p^t \sim \frac{c^t}{n^t} = \frac{1}{n}\,\mathcal{O}\left(\frac{1}{n^s}\right).$$

On the other hand, for each $s$ the number of terms in $\Sigma^s$ is $\binom{n}{s}\theta = \mathcal{O}(n^s)$, where $\theta$ does not depend on $n$, therefore

$$\Sigma_2 = \sum_{s=4}^{3r-1} \Sigma^s = \sum_{s=4}^{3r-1} \mathcal{O}(n^s)\frac{1}{n}\,\mathcal{O}\left(\frac{1}{n^s}\right) = \mathcal{O}\left(\frac{1}{n}\right) = o(1).$$

■

Theorems 3.3, 3.5, and 3.7 cover in full the question about triangles in $\mathcal{G}(n,p)$. Indeed, for any $p = p(n)$ we may have that either $np(n) \to 0$ (Theorem 3.3, here $p = o(1/n)$), $np(n) \to \infty$ (Theorem 3.5, here $1/n = o(p)$), or $np(n) \to c > 0$ (Theorem 3.7, here $p \sim c/n$). In particular, if $p = c > 0$ we immediately get a corollary that any random graph $\mathcal{G}(n,p)$ have at least one triangle whp, no matter how small $p$ is.

Clearly function $1/n$ plays an important role in this discussion.

**Definition 3.8.** *An increasing property is a graph property conserved under the addition of edges. A function $t(n)$ is a threshold function for an increasing property if (a) $p(n)/t(n) \to 0$ implies that $\mathcal{G}(n,p)$ does not possess this property whp, and b if $p(n)/t(n) \to \infty$ implies that it does possess this property whp.*

The number of triangles is an increasing property (by adding an edge we cannot reduce the number of triangles in a graph), and the function

$$t(n) = \frac{1}{n}$$

is a threshold function for this property.

**Problem 3.5.** Is the threshold function unique?

Here are some other examples of increasing properties:

- A fixed graph $H$ is a subgraph in $G$.
- There exists a large components in $G$.
- $G$ is connected.
- The diameter of $G$ is at most $d$.

**Problem 3.6.** Can you specify a graph property that is not increasing?

**Problem 3.7.** Consider the Erdős–Rényi random graph $\mathcal{G}(n,p)$. What is the expected number of spanning trees in $\mathcal{G}(n,p)$? (Cayley's formula is useful here, which says that the number of different trees on $n$ labeled vertices is $n^{n-2}$. Can you think of how to prove this formula?)

**Problem 3.8.** Consider two Erdős–Rényi random graphs $\mathcal{G}(n,p)$ and $\mathcal{G}(n,q)$. What is the probability that $H \in \mathcal{G}(n,p) \subseteq G \in \mathcal{G}(n,q)$?

**Problem 3.9.** Consider the following graph (call it $H$). Prove that for $\mathcal{G}(n,p)$ there exists a threshold function such that $\mathcal{G}(n,p)$ contains a.a.s. no subgraphs isomorphic to $H$ if $p$ is below the threshold and contains $H$ as a subgraph a.a.s. if $p$ is above the threshold.

**Problem 3.10.** Prove that

- If $pn = o(1)$ then $\mathcal{G}(n,p)$ contains no cycles. Hence, all components are trees.
- If $pn^{k/(k-1)} = o(1)$ then there are no trees of order $k$. (Cayley's formula can be useful.)
- If $pn^{k/(k-1)} = c$ then the trees of order $k$ distributed according to the Poisson law with mean $\lambda = c^{k-1}k^{k-2}/k!$.

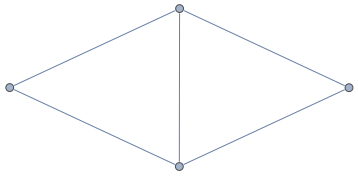

Figure 3.1: Graph $H$

The given exercises can be generalized in the following way. The ratio $2|E(G)|/|V(G)|$ for a graph $G$ is called its *average vertex degree*. A graph $G$ is called *balanced* if its average vertex degree is equal to the maximum average vertex degree over all its induced subgraphs.

**Theorem 3.9.** *For a balanced graph $H$ with $k$ vertices and $l \geq 1$ edges the function $t(n) = n^{-k/l}$ is a threshold function for the appearance of $H$ as a subgraph of $\mathcal{G}(n,p)$.*

**Problem 3.11.** What is the threshold function of appearance of complete graphs of order $k$?

Even more generally, it can be proved that all increasing properties have threshold functions!

## 3.3 Connectivity of the Erdős–Rényi random graphs

Function $\log n/n$ is a threshold function for the connectivity in $\mathcal{G}(n,p)$.

**Problem 3.12.** Show that the function $\log n/n$ is a threshold function for the disappearance of isolated vertices in $\mathcal{G}(n,p)$.

**Theorem 3.10.** *Let $p = \frac{c \log n}{n}$. If $c \geq 3$ and $n \geq 100$ then*

$$\mathsf{P}(\{\ \mathcal{G}(n,p) \text{ is connected } \}) \to 1.$$

*Proof.* Consider a random variable $X$ on $\mathcal{G}(n,p)$, which is defined as $X(G) = 0$ if $G$ is connected, and $X(G) = k$ is $G$ has $k$ components (note that $X(G) \neq 1$ for any $G$). We need to show $\mathsf{P}(X = 0) \to 1$, which is the same as $\mathsf{P}(X \geq 1) \to 0$, and by Markov's inequality (first moment method) $\mathsf{P}(X \geq 1) \leq \mathsf{E}\, X$.

Represent $X$ as

$$X = X_1 + \ldots + X_{n-1},$$

where $X_j$ is the number of the components that have exactly $j$ vertices. Now suppose that we order all $j$-element subsets of the set of vertices and label them from 1 to $\binom{n}{j}$. Consider the events $K_i^j$ such that the $i$th $j$ elements subset forms a component in $G$. Using the usual notation, we have

$$X_j = \sum_{i=1}^{\binom{n}{j}} \mathbf{1}_{K_i^j}.$$

As a result,

$$\mathsf{E}\, X = \sum_{j=1}^{n-1} \sum_{i=1}^{\binom{n}{j}} \mathsf{E}\, \mathbf{1}_{K_i^j} = \sum_{j=1}^{n-1} \sum_{i=1}^{\binom{n}{j}} \mathsf{P}(K_i^j).$$

Next,

$$\mathsf{P}(K_i^j) \leq \mathsf{P}(\{\text{ there are no edges connecting vertices in } K_i^j \text{ and in } V \setminus K_i^j\ \}) = (1-p)^{j(n-j)}.$$

Here we just disregard the condition that all the vertices in $K_i^j$ have to be connected.

The last inequality yields

$$\mathsf{E}X \le \sum_{j=1}^{n-1}\sum_{i=1}(1-p)^{j(n-j)} = \sum_{j=1}^{n-1}\binom{n}{j}(1-p)^{j(n-j)}. \tag{3.2}$$

The last sum is symmetric in the sense that the terms with $j$ and $n-j$ are equal. Let $j=1$:

$$n(1-p)^{n-1} \le ne^{-p(n-1)} \le e^{-\frac{3(n-1)\log n}{n}}.$$

Take $n$ such that $(n-1)/n \ge 0.9$, hence

$$n(1-p)^{n-1} \le e^{-2.7\log n} = \frac{1}{n^{2.7}}\,.$$

Consider now the quotient of two consecutive terms in (3.2)

$$\frac{\binom{n}{j+1}(1-p)^{(j+1)(n-j-1)}}{\binom{n}{j}(1-p)^{j(n-j)}} = \frac{n-j}{j+1}(1-p)^{n-2j-1}.$$

If $j \le n/8$ then

$$\begin{aligned}\frac{n-j}{j+1}(1-p)^{n-2j-1} &\le (n-1)(1-p)^{\frac{3n}{4}-1} \le (n-1)e^{-\frac{9\log n}{4}+p} \le \\ &\le ne^{-\frac{9\log n}{4}+p} \le (\text{for sufficiently large } n) \le \\ &\le ne^{-2n\log n} = \frac{1}{n}\,.\end{aligned}$$

If $j > \frac{n}{8}$ one has

$$\binom{n}{j} < 2^n, \quad (1-p)^{j(n-j)} \le (1-p)^{\frac{n^2}{16}} \le e^{-\frac{pn^2}{16}} \le e^{-\frac{3n\log n}{16^2}},$$

hence

$$\binom{n}{j}(1-p)^{j(n-j)} \le 2^n n^{-\frac{3n}{16}},$$

which is again, for sufficiently large $n$ is small compared to $n^{-2.7}$. Hence in the sum (3.2) the first term is the biggest, therefore

$$\mathsf{E}X \le \sum_{j=1}^{n-1}\frac{1}{n^{2.7}} < \frac{n}{n^{2.7}} \to 0, \quad n\to\infty.$$

■

More exact (theoretically) results is that

**Theorem 3.11.** *Let $p = \frac{c\log n}{n}$. If $c > 1$ then $\mathcal{G}(n,p)$ is connected whp. If $c < 1$ then $\mathcal{G}(n,p)$ is not connected whp.*

The proof for the part $c > 1$ follows the lines of Theorem 3.10, however the estimates have to be made more accurately. For the part $c < 1$ the starting point is again Markov's inequality. We need to show that $\mathsf{P}\,X > 1 \to 1$ as $n\to\infty$ (cf. Theorem 3.5). For the case $c = 1$, however, we need $p(n) \sim \frac{\log n}{n}$ and there are many possible functions of this form. Here is an example of a statement for one of them:

**Theorem 3.12.** *Let $p(n) = (\log n + c + o(1))n^{-1}$. Then*

$$\mathsf{P}(\{\ \mathcal{G} \text{ is connected }\}) \to e^{-e^{-c}}.$$

*In particular, if $p = \frac{\log n}{n}$, this probability tends to $e^{-1}$.*

## 3.4 The giant component of the Erdős–Rényi random graph

### 3.4.1 Non-rigorous discussion

We know that if $pn = o(1)$ then there are no triangles. In a similar manner it can be shown that there are no cycles of any order in $\mathcal{G}(n,p)$. This means that most components of the random graph are trees and isolated vertices. For $p > c\log n/n$ for $c \geq 1$ the random graph is connected whp. What happens in between these stages? It turns our that a unique giant component appears when $p = c/n$, $c > 1$. We first study the appearance of this largest component in a heuristic manner. We define *the giant component* as the component of $\mathcal{G}(n,p)$, whose order is $\mathcal{O}(n)$.

Let $u$ be the frequency of the vertices that do not belong to the giant component. In other words, $u$ gives the probability that a randomly picked vertex does not belong to the giant component. Let us calculate this probability in a different way. Pick any other vertex. It can be either in the giant component or not. For the original node not to belong to the giant component these two either should not be connected (probability $1-p$) or be connected, but the latter vertex is not in the giant component (probability $pu$). There are $n-1$ vertices to check, hence

$$u = (1 - p + pu)^{n-1}.$$

Recall that we are dealing with $pn \sim \lambda$, hence it is convenient to use the parametrization

$$p = \frac{\lambda}{n}, \quad \lambda > 0.$$

Note that $\lambda$ is the mean degree of $\mathcal{G}(n,p)$. We get

$$\begin{aligned}
u = \left(1 - \frac{\lambda}{n}(1-u)\right)^{n-1} &\implies \\
\log u(n-1)\log\left(1 - \frac{\lambda}{n}(1-u)\right) &\implies \\
\log u \approx -\frac{\lambda(n-1)}{n}(1-u) &\implies \\
\log u \approx -\lambda(1-u) &\implies \\
u = e^{-\lambda(1-u)}, &
\end{aligned}$$

where the fact that $\log(1+x) \approx x$ for small $x$ was used.

Finally, for the frequency of the vertices in the giant component $v = 1 - u$ we obtain

$$1 - v = e^{-\lambda v}.$$

This equation always has the solution $v = 0$. However, this is not the only solution for all possible $\lambda$s. To see this consider two curves, defined by $f(v) = 1 - e^{-\lambda v}$ and $g(v) = v$. Their intersections give the roots to the original equation. Note that, as expected, $f(0) = g(0) = 0$. Note also that $f'(v) = \lambda e^{-\lambda v} > 0$ and $f''(v) = -\lambda^2 e^{-\lambda v} < 0$. Hence the derivative for $v \leq 0$ cannot be bigger than at $v = 0$, which is $f'(0) = \lambda$. Hence (see the figure), if $\lambda \leq 0$ then there is unique trivial solution $v = 0$, but if $\lambda > 1$ then another positive solution $0 < v < 1$ appears.

Technically, we only showed that if $\lambda \leq 1$ then there is no giant component. If $\lambda > 1$ then we use the following argument. Since $\lambda$ gives the mean degree, then, starting from a randomly picked vertex, it will have $\lambda$ neighbors on average. Its neighbors will have $\lambda^2$ neighbors (there are $\lambda$ of them and each has $\lambda$ adjacent edges) and so on. After $s$ steps we would have $\lambda^s$ vertices within the distance $s$ from the initial vertex. If $\lambda > 1$ this number will grow exponentially and hence most of the nodes have to be connected into the giant component. Their frequency can be found as the nonzero solution to $1 - v = e^{-\lambda v}$. Of

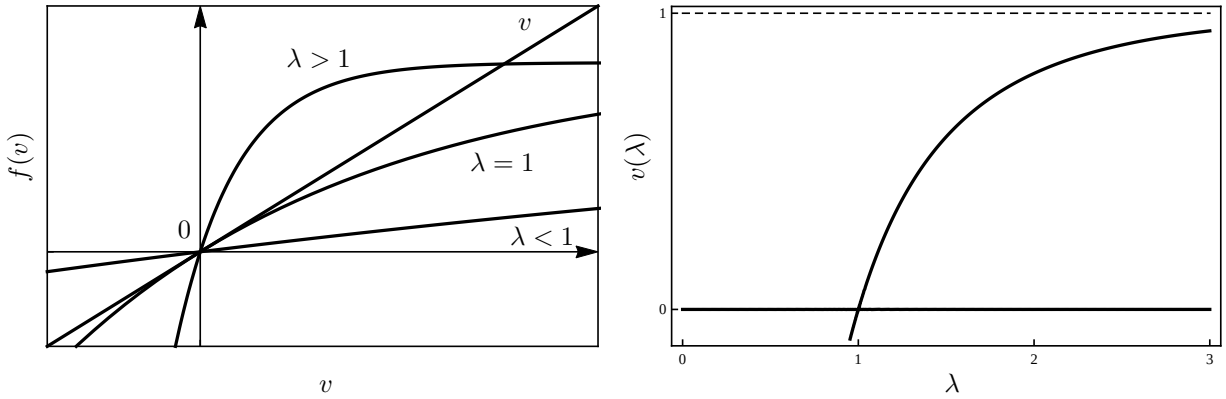


Figure 3.2: The analysis of the number of solutions of the equation $1 - v = e^{-\lambda v}$. The left panel shows that if $v \geq 0$ then it is possible to have one trivial solution $v = 0$ in the case $\lambda \leq 1$, and two solutions if $\lambda > 1$. The right panel shows the solutions (thick curves) as the functions of $\lambda$

course this type of argument is extremely rough, but the fact is that it can be made absolutely rigorous within the framework of the branching processes.

Moreover, the last heuristic reasoning can be used to estimate the diameter of $\mathcal{G}(n, p)$. Obviously, the process of adding new vertices cannot continue infinitely, it has to stop when we reach all $n$ vertices:

$$\lambda^s = n.$$

From where we have that

$$s = \frac{\log n}{\log \lambda}$$

approximates the diameter of the Erdős–Rényi random graph. It is quite surprising that the exact results are basically the same: It can be proved that for $np > 1$ and $np < c \log n$, the diameter of the random graph (understood as the diameter of the largest connected component) is concentrated on at most four values around $\log n / \log np$.

Finally we note that since the appearance of the giant component shows this threshold behavior (if $\lambda < 1$ there is no giant component a.a.s., if $\lambda > 1$ the giant component is present a.a.s.) one often speaks of a *phase transition*.

## 3.5 Branching processes

### 3.5.1 Generating functions

For the following we will need some information on probability generating functions. This section serves as a short review of the pertinent material.

Let $(a_k)_{k=0}^{\infty} = a_0, a_1, a_2 \ldots$ be a sequence of real numbers. If the function

$$\varphi(s) = a_0 + a_1 s + a_2 s^2 + \ldots = \sum_{k=0}^{\infty} a_k s^k$$

converges in some interval $|s| < s_0$, then $\varphi$ is called the *generating function* for the sequence $(a_k)_{k=0}^{\infty}$. The variable $s$ here is a dummy variable. If $(a_k)_{k=0}^{\infty}$ is bounded then $\varphi(s)$ is convergent at least for some $s$ other than zero.

**Example 3.13.** Let $a_k = 1$ for any $k = 0, 1, 2, \ldots$. Then (prove this formula)

$$\varphi(s) = 1 + s + s^2 + s^3 + \ldots = \frac{1}{1-s}, \quad |s| < 1.$$

Let $(a_k)_{k=0}^{\infty} = 1, 2, 3, 4, \ldots$ then

$$\varphi(s) = 1 + 2s + 3s^2 + 4s^3 + \ldots = \frac{1}{(1-s)^2} \quad |s| < 1.$$

Let $(a_k)_{k=0}^{\infty} = 0, 0, 1, 1, 1, 1, \ldots$ then

$$\varphi(s) = s^2 + s^3 + s^4 + \ldots = \frac{s^2}{1-s}, \quad |s| < 1.$$

Let $a_k = \binom{n}{k}$ then

$$\varphi(s) = \sum_{k=0}^{\infty} \binom{n}{k} s^k = \sum_{k=0}^{n} \binom{n}{k} s^k = (1+s)^n, \quad s \in \mathbf{R}.$$

Let $a_k = \frac{1}{k!}$ then

$$\varphi(s) = 1 + s + \frac{s^2}{2!} + \frac{s^3}{3!} + \ldots = e^s, \quad s \in \mathbf{R}.$$

**Problem 3.13.** Consider the Fibonacci sequence

$$0, 1, 1, 2, 3, 5, 8, 13, \ldots,$$

where we have

$$a_k = a_{k-1} + a_{k-2}, \quad k = 2, 3, 4, \ldots.$$

Find the generating function for this sequence.

If we have $A(s)$ then the formula to find the elements of the sequence is

$$a_k = \frac{A^{(k)}(0)}{k!}.$$

Let $X$ be a discrete random variable that assumes integer values $0, 1, 2, 3, \ldots$ with the corresponding probabilities

$$\mathsf{P}(X = k) = p_k.$$

The generating function for the sequence $(p_k)_{k=0}^{\infty}$ is called *probability generating function* and often abbreviated pgf:

$$\varphi(s) = \varphi_X(s) = p_0 + p_1 s + p_2 s^2 + \ldots$$

**Example 3.14.** Let $X$ be a Bernoulli's random variable, then its pgf is

$$\varphi_X(s) = 1 - p + ps.$$

Let $Y$ be a Poisson random variable, then its pgf is

$$\varphi_Y(s) = \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!} e^{-\lambda} = e^{-\lambda(1-s)}.$$

Note that for any pgf

$$\varphi(1) = p_0 + p_1 + p_2 + \ldots = 1,$$

hence pgf converges in some interval containing 1.

Let $\varphi(s)$ be a pgf of $X$ that assumes values $0, 1, 2, 3, \ldots$. Consider $P'(1)$:

$$\varphi'(1) = \sum_{k=0}^{\infty} k p_k = \mathsf{E}X,$$

if the corresponding expectation exists (there are random variables with no average). Therefore, if we know pgf then it is straightforward to find the mean value.

Similarly,

$$\mathsf{E}\big(X(X-1)\big) = \sum_{k=1}^{\infty} k(k-1) p_k = \varphi''(1),$$

or, in general,

$$\mathsf{E}\,(X)_r = \varphi^{(r)}(1).$$

Hence any moment can be found using the pgf. For instance,

$$\operatorname{Var} X = \varphi''(1) + \varphi'(1) - \big(\varphi'(1)\big)^2.$$

**Problem 3.14.** Show that

$$\mathsf{E}(X^r) = \left(s\frac{\mathrm{d}}{\mathrm{d}s}\right)^r \varphi_X(s)\bigg|_{s=1}.$$

**Problem 3.15.** Prove for the geometric random variable that is defined as

$$\mathsf{P}(X = k) = (1-p)^k p, \quad k = 0, 1, 2, \ldots$$

that

$$\mathsf{E}X = \frac{1-p}{p}, \quad \operatorname{Var} X = \frac{1-p}{p^2}.$$

(This random variable can be interpreted as the number of the first win in a series of trials with the probability of success $p$ in one trial.)

**Problem 3.16.** Prove the equality

$$\mathsf{E}X = \sum_{k=0}^{\infty} \mathsf{P}(X > k)$$

for the integer values random variable $X$ by introducing the generating function $\psi(s)$ for the sequence

$$q_k = \sum_{j=k+1}^{\infty} p_j.$$

Let $X$ assume values $0, 1, 2, 3, \ldots$. For any $s$ the expression $s^X$ is a well defined new random variable, which has the expectation

$$\mathsf{E}(s^X) = \sum_{k=0}^{\infty} k s^k = \varphi(s).$$

If random variables $X$ and $Y$ are independent then so $s^X$ and $s^Y$, therefore,

$$\mathsf{E}(s^{X+Y}) = \mathsf{E}(s^X)\,\mathsf{E}(s^Y),$$

which means for the probability generating functions

$$\varphi_{X+Y}(s) = \varphi_X(s)\varphi_Y(s),$$

i.e., the pgf of the sum of two independent random variables can be found as the product of the corresponding pgf-s. The next step is to generalize this expression on the sum of $n$ random variables. Let

$$S_n = X_1 + X_2 + \ldots + X_n,$$

where $X_i$ are i.i.d. random variables with pgf $\varphi_X(s)$. Then

$$\varphi_{S_n}(s) = \big(\varphi_X(s)\big)^n.$$

**Example 3.15.** The pgf for the binomial random variable $S_n$ can be easily found by noting that

$$S_n = X_1 + \ldots + X_n,$$

where each $X_i$ has Bernoulli's distribution. Therefore,

$$\varphi_{S_n}(s) = (1 - p + ps)^n,$$

which also can be proved directly. (Prove this formula directly and find $\mathsf{E}S_n$ and $\operatorname{Var} S_n$.)

**Example 3.16.** We have that for Poisson random variable its generating function is $e^{-\lambda(1-s)}$. Consider now two independent Poisson random variables with parameters $\lambda_1$ and $\lambda_2$ respectively. We have

$$\varphi_X(s)\varphi_Y(s) = e^{-\lambda_1(1-s)}e^{-\lambda_2(1-s)} = e^{-(\lambda_1+\lambda_2)(1-s)} = \varphi_{X+Y}(s).$$

In other words for the sum of two independent Poisson random variables we found that it also has Poisson distribution with parameter $\lambda_1 + \lambda_2$.

It is important to mention that pgf determines the random variable. More precisely, if two random variable have pgfs $\varphi_1(s)$ and $\varphi_2(s)$, both pgfs converge in some open interval containing 1 and $\varphi_1(s) = \varphi_2(s)$ in this interval, then the two random variable have identical distributions. A second important fact is that if a sequence of pgfs converges to a limiting pgf, then the sequence of the corresponding probability distributions converges to the limiting probability distribution (note that the pgf for a binomial random variable converges to the pgf of the Poisson random variable).

### 3.5.2 Branching processes

Branching processes are central to the mathematical analysis of random networks. Here I would like to define the Galton–Watson branching process and study its basic properties.

It is convenient to define the Galton–Watson process in terms of individuals and their descendants. Let $X_0 = 1$ be the initial individual. Let $Y$ be the random variable with the probability distribution $\mathsf{P}(Y = k) = p_k$. That random variable describes the number of descendants of each individual. The number of descendants in generation $X_{n+1}$ depends only on $X_n$ and is given by

$$X_{n+1} = \sum_{j=1}^{X_n} Y_j,$$

where $Y_j$ are i.i.d. random variables such that $Y_j \sim Y$. Sequence $(X_0, X_1, \ldots, X_n, \ldots)$ defines the Galton–Watson branching process.

Let $\varphi(s)$ be the probability generating function of $Y$; i.e.,

$$\varphi(s) = \sum_{k=0}^{\infty} p_k s^k.$$

Let us find the generating function for $X_n$:

$$\varphi_n(s) = \sum_{k=0}^{\infty} \mathsf{P}(X_n = k) s^k, \quad n = 0, 1, \ldots$$

We have

$$\varphi_0(s) = s, \quad \varphi_1(s) = \varphi(s).$$

Further,

$$\begin{aligned}
\varphi_{n+1}(s) &= \sum_{k=0}^{\infty} \mathsf{P}(X_{n+1} = k) s^k \\
&= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \mathsf{P}(X_{n+1} = k \mid X_n = j)\, \mathsf{P}(X_n = j) s^k \\
&= \sum_{j=0}^{\infty} \mathsf{P}(X_n = j) \sum_{k=0}^{\infty} \mathsf{P}(Y_1 + \ldots + Y_j = k) s^k \\
&= \sum_{j=0}^{\infty} \mathsf{P}(X_n = j) \big(\varphi(s)\big)^j \\
&= \varphi_n(\varphi(s)),
\end{aligned}$$

where the properties of the generating function of the sum of independent random variables were used. Now, using the relation

$$\varphi_{n+1}(s) = \varphi_n(\varphi(s)),$$

we find $\mathsf{E}X_n$ and $\operatorname{Var} X_n$. Assume that $\mathsf{E}Y = \mu$ and $\operatorname{Var} Y = \sigma^2$ exist.

$$\mathsf{E}X_n = \frac{d}{ds}(\varphi_{n-1}(s))|_{s=1} = \varphi'_{n-1}(1)\varphi'(1) = \varphi'_{n-2}(1)\big(\varphi'(1)\big)^2 = \mu^n.$$

Hence the expectation grows if $\mu > 1$, decreases if $\mu < 1$ and stays the same if $\mu = 1$.

**Problem 3.17.** Show that

$$\operatorname{Var} X_n = \begin{cases} \sigma^2 \mu^{n-1} \dfrac{\mu^n - 1}{\mu - 1}, & \mu \neq 1, \\ n\sigma^2, & \mu = 1. \end{cases}$$

Now I would like to calculate the *extinction probability*:

$$\mathsf{P}(\{X_n = 0 \text{ for some } n\}).$$

To do this, consider

$$q_n = \mathsf{P}(X_n = 0) = \varphi_n(0).$$

Also note that $p_0$ has to be bigger than zero. Then, since the relation $\varphi_{n+1}(s) = \varphi_n(\varphi(s))$ can be rewritten (why?) as

$$\varphi_{n+1}(s) = \varphi(\varphi_n(s)),$$

I have

$$q_{n+1} = \varphi(q_n).$$

Function $\varphi(s)$ is strictly increasing, with $q_1 = p_0 > 0$, which implies that $q_{n+1} > q_n$ and all $q_n$ are bounded by 1. Therefore, there exists

$$\pi = \lim_{n \to \infty} q_n,$$

and $0 < \pi \leq 1$. Since $\varphi(s)$ is continuous, we obtain that

$$\pi = \varphi(\pi),$$

which actually gives the equation to find the extinction probability $\pi$ (due to the fact that $q_n$ are defined to be probabilities of extinction at generation $n$ or prior to it). Actually, it can be proved (exercise!) that $\pi$ is the smallest root of the equation

$$\varphi(s) = s.$$

Note that this equation always has root 1. Now assume that $p_0 + p_1 < 1$. Then $\varphi''(s) > 0$ and $\varphi(s)$ is a convex function, which can intersect the 45° line at most at two points. On of these points is 1. The other one is less than 1 only if $\varphi'(1) = \mu > 1$. Therefore, we have proved

**Theorem 3.17.** *For the Galton–Watson branching process the probability of extinction is given by the smallest root to the equation*

$$\varphi(s) = s.$$

*This root is* 1*, i.e., the process dies out for sure if* $\mu < 1$ (*subcritical process*)*, or if* $\mu = 1$ *and* $p_1 \neq 1$ (*critical process*)*, and this root is strictly less than 1 if* $\mu > 1$ (*supercritical process*)*.*

**Problem 3.18.** How the results above change if $X_0 = i$, where $i \geq 2$?

We showed that if the average number of descendants $\leq 1$ then the population goes extinct with probability 1. On the other hand, if the average number of offspring is $> 1$, then still there is nonzero probability that the process will die out ($\pi$), however, with probability $1 - \pi$ we find that $X_n \to \infty$.

**Example 3.18.** Let $Y \sim \text{Poisson}(\lambda)$, then $\varphi(s) = e^{\lambda(s-1)}$, and the extinction probability can be found as the smallest root of

$$s = e^{\lambda(s-1)}.$$

This is exactly the equation for the probability that a randomly chosen node in the Erdős–Rényi model $\mathcal{G}(n, p)$ does not belong to the giant component! And of course this is not a coincidence.

**Problem 3.19.** Find the probability generating function for the random variable

$$T = \sum_{i=0}^{\infty} X_i = 1 + \sum_{i=1}^{\infty} X_i.$$

Use the fact that

$$T = 1 + \sum_{j=1}^{X_1} T_j,$$

where $T, T_1, \ldots, T_{X_1}$ are i.i.d. random variables. Show that

$$\mathsf{E}T = \frac{1}{1-\mu}\,.$$

### 3.5.3 Rigorous results for the appearance of the giant component

```
Add the discussion on the branching processes
and relation to the appearance of the giant component.
```

### 3.5.4 Rigorous results on the diameter of the Erdős–Rényi graph

## 3.6 The evolution of the Erdős–Rényi random graph

Here I summarize the distinct stages of the evolution of the Erdős–Rényi random graph.

Stage I: $p = o(1/n)$

The random graph $\mathcal{G}(n,p)$ is the disjoint union of trees. Actually, as you are asked to prove in one of the exam problems, there are no trees of order $k$ if $pn^{k/(k-1)} = o(1)$. Moreover, for $p = cn^{-k/(k-1)}$ and $c > 0$, the probability distribution of the number of trees of order $k$ tends to the Poisson distribution with parameter $\lambda = c^{k-1}k^{k-2}/k!$. If $1/(pn^{k/(k-1)}) = o(1)$ and $pkn - \log n - (k-1)\log\log n \to \infty$, then there are trees of any order a.a.s. If $1/(pn^{k/(k-1)}) = o(1)$ and $pkn - \log n - (k-1)\log\log n \sim x$ then the trees of order $k$ distributed asymptotically by Poisson law with the parameter $\lambda = e^{-x}/(kk!)$.

Stage II: $p \sim c/n$ for $0 < c < 1$

Cycles of any given size appear. All connected components of $\mathcal{G}(n,p)$ are either trees or unicycle components (trees with one additional edge). Almost all vertices in the components which are trees ($n - o(n)$). The largest connected component is a tree and has about $\alpha^{-1}(\log n - 2.5\log\log n)$ vertices, where $\alpha = c - 1 - \log c$. The mean of the number of connected components is $n - p\binom{n}{2} + \mathcal{O}(1)$, i.e., adding a new edge decreases the number of connected components by one. The distribution of the number of cycles on $k$ vertices is approximately a Poisson distribution with $\lambda = c^k/(2k)$.

Stage III: $p \sim 1/n + \mu/n$, the double jump

Appearance of the giant component. When $p < 1/n$ then the size of the largest component is $\mathcal{O}(\log n)$ and most of the vertices belong to the components of the size $\mathcal{O}(1)$, whereas for $p > 1/n$ the size of the unique largest component is $\mathcal{O}(n)$, the remaining components are all small, the biggest one is of the order of $\mathcal{O}(\log n)$. All the components other than the giant one are either trees or unicyclic, although the giant component has complex structure (there are cycles of any period). The natural question is how the biggest component grows so quickly. Erdős and Rényi showed that it actually happens in two steps, hence the term"double jump." If $\mu < 0$ then the largest component has the size $(\mu - \log(1+\mu))^{-1}\log n + \mathcal{O}(\log\log n)$. If $\mu = 0$ then the largest component has the size of order $n^{2/3}$, and for $\mu > 0$ the giant component has the size $\alpha n$ for some constant $\alpha$.

Stage IV: $p \sim c/n$ where $c > 1$

Except for one giant component all the components are small, and most of them are trees. The evolution of the random graph here can be described as merging the smaller components with the giant one, one after another. The smaller the component, the larger the chance of "survival." The survival time of a tree of order $k$ is approximately exponentially distributed with the mean value $n/(2k)$.

Stage V: $p = c\log n/n$ with $c \geq 1$

The random graph becomes connected. For $c = 1$ there are only the giant component and isolated vertices.

Stage VI: $p = \omega(n)\log n/n$ where $\omega(n) \to \infty$ as $n \to \infty$. In this range the random graph is not only connected, but also the degrees of all the vertices are asymptotically equal.

Here is the numerical illustration of the evolution of the random graph. To present it I fixed $n = 1000$ the number of vertices and generated $\binom{n}{2}$ random variables, uniformly distributed in $[0,1]$. Hence each edge gets its only number $p_j \in [0,1]$, $j = 1, \ldots, \binom{n}{2}$. For any fixed $p \in (0,1)$ I draw only the edges for which $p_j \leq p$. Therefore, in this manner I can observe how the evolution of the random graph occurs for $\mathcal{G}(n,p)$.

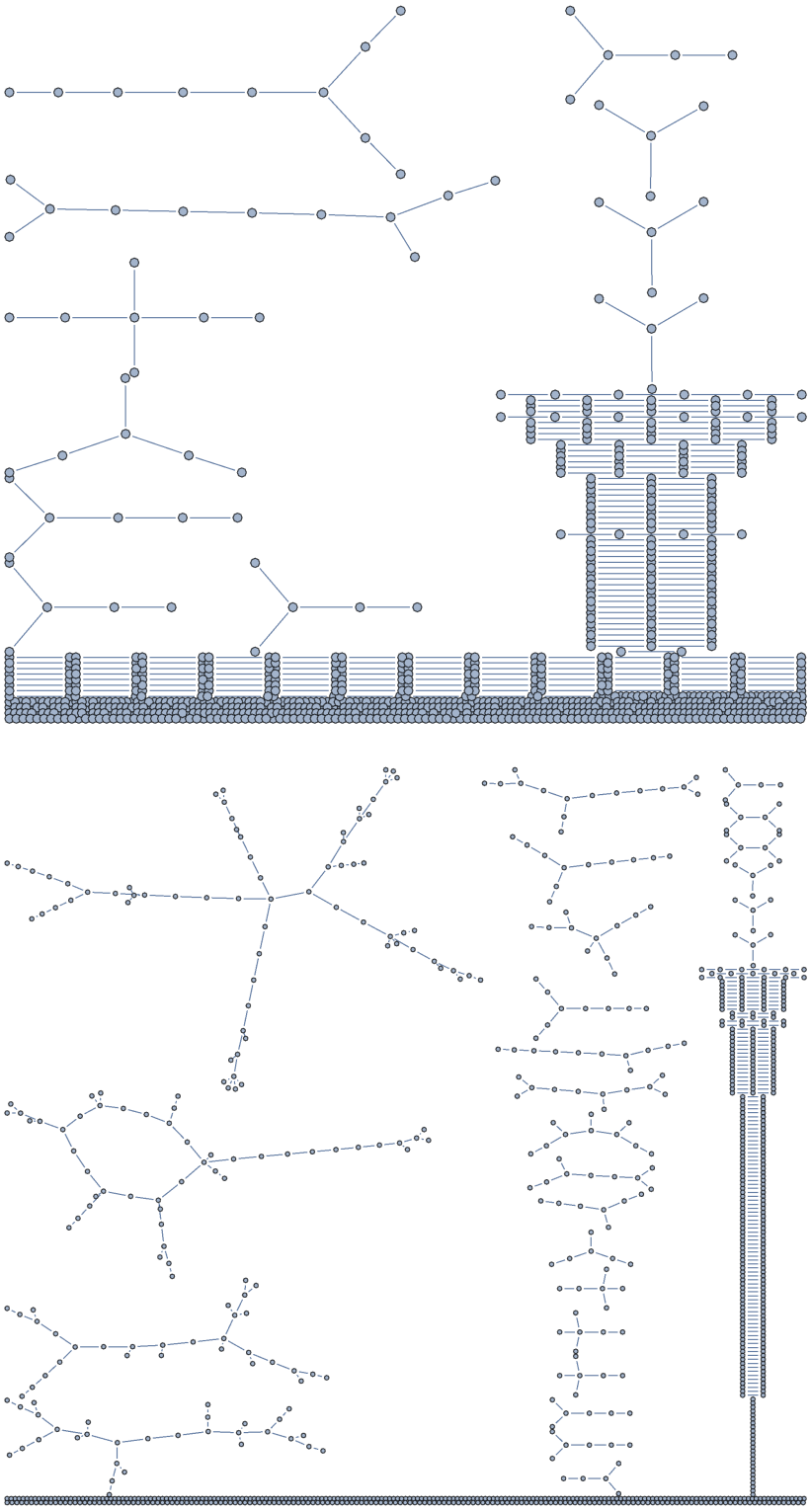

Figure 3.3: Stages I and II. Graphs $G(1000, 0.0005)$ and $G(1000, 0.00095)$ are shown

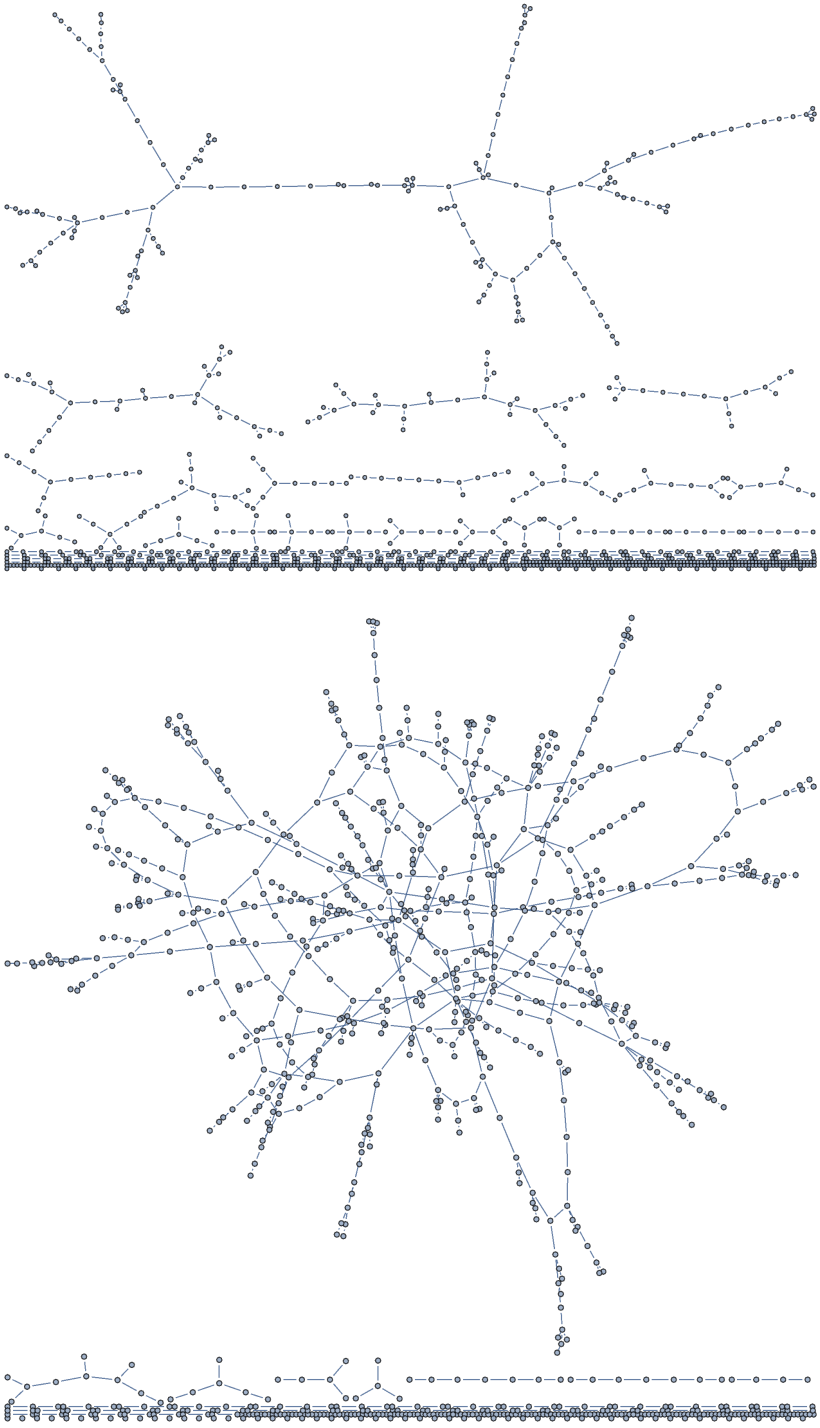

Figure 3.4: Stages III and IV. Graphs $G(1000, 0.001)$ and $G(1000, 0.0015)$ are shown. The giant component is born

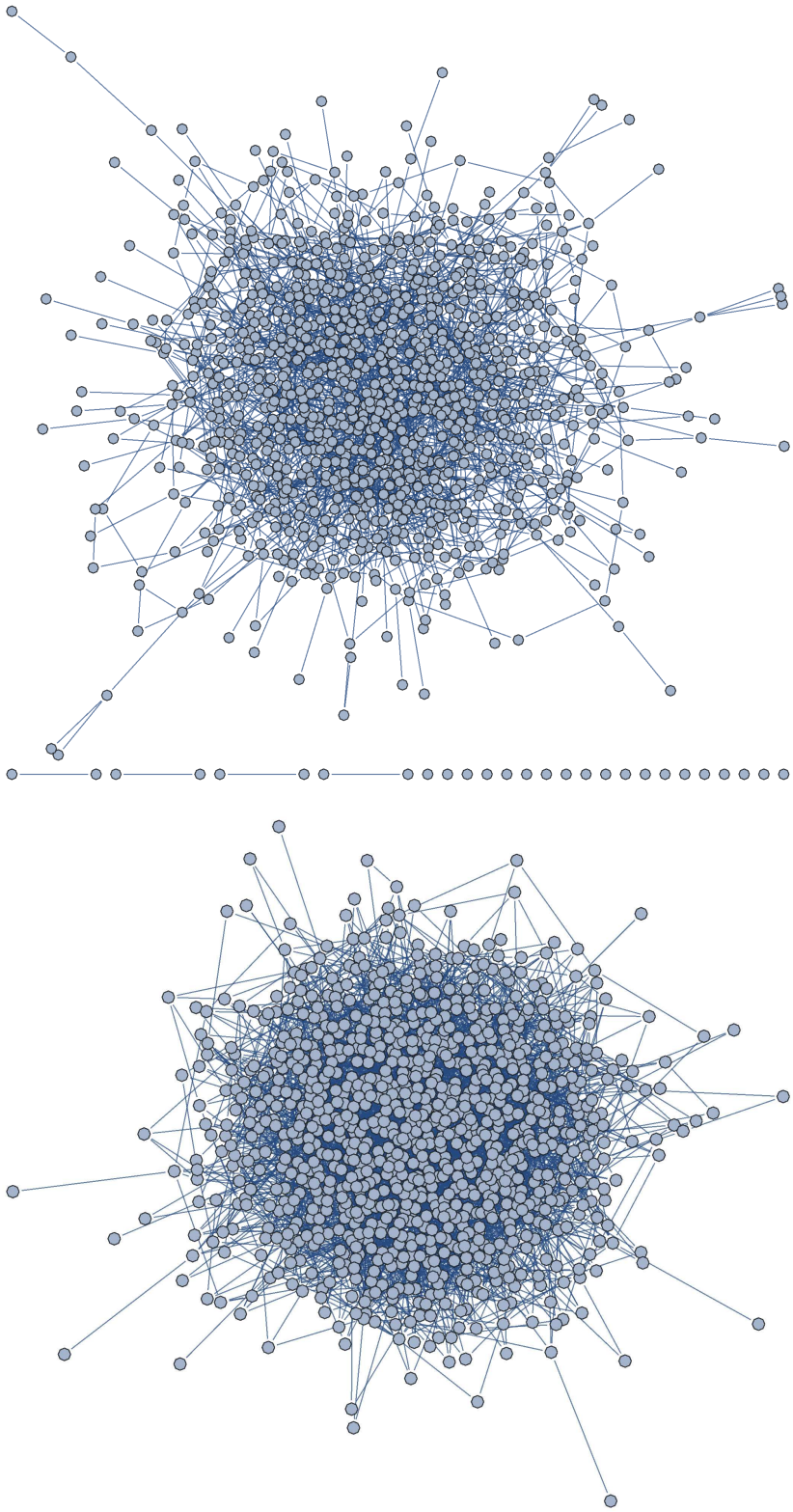



Figure 3.5: Stages IV and V. Graphs $G(1000, 0.004)$ and $G(1000, 0.007)$ are shown. The final graph is connected (well, almost)