TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 286, Number 1, November 1984

THE EVOLUTION OF RANDOM GRAPHS

BY
BÉLA BOLLOBÁS[1]

ABSTRACT. According to a fundamental result of Erdös and Rényi, the structure of a random graph G_M changes suddenly when $M \sim n/2$: if $M = \lfloor cn \rfloor$ and $c < \frac{1}{2}$ then a.e. random graph of order n and since M is such that its largest component has $O(\log n)$ vertices, but for $c > \frac{1}{2}$ a.e. G_M has a giant component: a component of order $(1 - \alpha_c + o(1))n$ where $\alpha_c < 1$. The aim of this paper is to examine in detail the structure of a random graph G_M when M is close to $n/2$. Among others it is proved that if $M = n/2 + s$, $s = o(n)$ and $s \geq (\log n)^{1/2} n^{2/3}$ then the giant component has $(4 + o(1))s$ vertices. Furthermore, rather precise estimates are given for the order of the rth largest component for every fixed r.

1. Introduction. Let n be a natural number and set $N = \binom{n}{2}$ and $V = \{1, 2, \ldots, n\}$. A *graph process* on V is a sequence $(G_t)_0^N$ such that (i) each G_t is a graph on V with t edges, and (ii) $G_0 \subset G_1 \subset \cdots \subset G_N$. Let $\tilde{\mathcal{G}}$ be the set of all $N!$ graph processes. Turn $\tilde{\mathcal{G}}$ into a probability space by giving all members of it the same probability and write $\tilde{G}$ for random elements of $\tilde{\mathcal{G}}$. Furthermore, call G_t the *state* of the process $\tilde{G} = (G_t)_0^N$ at time t. Clearly a (random) graph process is a Markov chain whose states are graphs on V. This Markov chain is a model of the evolution of a random graph (r.g.) with vertex set V.

The evolution of random graphs was first studied by Erdös and Rényi [**5–7**]. They investigated the least values of t for which certain properties are likely to appear, i.e. they studied the stage of the evolution of a r.g. at which a given property first appears. Erdös and Rényi proved the surprising fact that most properties studied in graph theory appear rather suddenly: there are functions $t_1(n) < t_2(n)$ rather close to each other such that almost no G_{t_1} has the property and almost every G_{t_2} has the property. (As customary in the theory of random graphs, the term '*almost every*' (a.e.) means 'with probability tending to 1 as $n \to \infty$'.)

Perhaps the most interesting results of Erdös and Rényi concern the sudden change in the structure of G_t around $t = n/2$.

They proved that if $t \sim cn$ for some constant c, $0 < c < 1/2$, then a.e. G_t is such that its largest component has $O(\log n)$ vertices: if $t \sim cn$ and $c > 1/2$ then the largest component of a.e. G_t has $(1 - \alpha_c + o(1))n$ vertices, where $0 < \alpha_c < 1$; and if $t = \lfloor n/2 \rfloor$ then the maximal size of a component of a.e. G_t has order $n^{2/3}$. (In fact, Erdös and Rényi [**6, 7**] asserted the last statement for $t \sim n/2$ but, as we

Received by the editors December 3, 1982 and, in revised form, August 15, 1983.
1980 *Mathematics Subject Classification*. Primary 05C99, 05C30; Secondary 60J99, 62P99.
[1]Research supported by NSF Grant MCS-8104854.

0002-9947/84 $1.00 + $.25 per page

shall see, that is not true.) Furthermore,

$$\alpha_c = \frac{1}{2c}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(2ce^{-2c})^k$$

so $\alpha_c e^{2c} \to 1$ as $c \to \infty$.

The aim of this note is to prove considerably more precise results about the structure of G_t, especially when t is close to $n/2$. Among others, our results will shed light on the curious double jump described above.

Loosely speaking, we shall show that a.e. $\tilde{G}$ is such that for $t \geq n/2 + (\log n)^{1/2}n^{2/3}$ the graph G_t has a unique component of order at least $n^{2/3}$ (the *giant component* of G_t) and all other components have fewer than $n^{2/3}/2$ vertices. Furthermore, if $n^{\beta_1} < s < n^{\beta_2}$, where $2/3 < \beta_1 < \beta_2 < 1$, then $G_{n/2-s}$ and $G_{n/2+s}$ have remarkably similar structures. A.e. $G_{n/2-s}$ has all its vertices on components of order at most $n^2(\log n)/s$ and so has $G_{n/2+s}$, except for its vertices on its giant component.

The giant component of $G_{n/2+s}$ has $(4+o(1))s$ vertices. Most vertices are on small components which are trees. The distributions of these tree-components of $G_{n/2-s}$ and $G_{n/2+s}$ are very similar, except in $G_{n/2+s}$ there are about $1 - 2s/n^2$ times as many of them as in $G_{n/2-s}$. As an easy corollary of our results, for $t \sim cn$, $c > 1/2$, we obtain precise information about the distribution of the order of the rth largest component of G_t for every fixed r.

2. Definitions and basic facts. In addition to graph processes we shall consider two other models of random graphs. The space $\mathcal{G}(n, M)$ consists of all graphs with $M = M(n)$ edges and with vertex set $V = \{1, 2, \ldots, n\}$; the elements of $\mathcal{G}(n, M)$ are equiprobable. The model $\mathcal{G}(n, p)$ consists of all 2^N graphs with vertex set V, in which the probability of a graph with m edges is $p^m q^{N-m}$, where $q = 1 - p$. Thus $\mathcal{G}(n, p)$ consists of all graphs with vertex set V in which the edges are chosen independently and with the same probability $p = p(n)$, $0 < p < 1$. For basic properties of these models see [**1**, Chapter 7 and **3**]; for undefined terminology in graph theory see [**2**]. As customary, we shall talk of random graphs G_M and G_p, meaning that we consider elements of $\mathcal{G}(n, M)$ and $\mathcal{G}(n, p)$. Note that the probability that a r.g. G_M has Q is the same as the probability that a graph process $\tilde{G} = (G_t)_0^N$ is such that G_t has Q for $t = M$. Therefore no confusion will arise from the two slightly different meanings of the symbol G_M. We say that *almost every* G_M (or G_p) *has a property* Q if the probability that G_M (or G_p) has Q tends to 1 as $n \to \infty$. If M is close to pN then the models $\mathcal{G}(n, M)$ and $\mathcal{G}(n, p)$ are virtually interchangeable. In particular, if $pqN \to \infty$, Q is a convex property (that is if $F \subset G \subset H$ and F and H have Q then so does G) and a.e. G_p has Q then so does a.e. G_M, provided $|M - pn| = O(pqN)^{1/2}$. These remarks imply that most of our results could be formulated for any of our models; nevertheless, there are advantages in considering all three.

Denote by $C(k, d)$ the number of connected labelled graphs with k vertices and $k + d$ edges. Thus $C(k, d) = 0$ if $d \leq 2$, $C(k, -1)$ is the number of labelled trees of order k, so $C(k, -1) = k^{k-2}$, and $C(k, 0)$ is the number of connected unicyclic graphs. For $d \geq 0$ the function $C(k, d)$ is less simple. It was proved by Katz [**9**] and Rényi [**10**] (and it is also a consequence of more general formulae in [**8**]) that

$$C(k,0) = \frac{1}{2}k^{k-1}\sum_{r=3}^{k}\prod_{j=1}^{r-1}\left(1-\frac{j}{k}\right) \sim \left(\frac{\pi}{8}\right)^{1/2} k^{k-1/2}. \tag{1}$$

For $d \geq 1$ Wright [**11–13**] proved a number of results about $C(k,d)$. He showed that for $d = o(k^{1/3})$

$$C(k,d) = f_d k^{d+(3d-1)/2}\{1 + O(d^{3/2}/k)\}, \tag{2}$$

where f_d depends only on d. In particular, $f_0 = (\pi/8)^{1/2}$, $f_1 = 5/24$ and $f_2 = 5\sqrt{2}/128$. We shall not need exact estimates in the vein (2), but we shall need a bound on $C(k,d)$ which is valid for all values of d. To be precise, we need the following inequality from [**4**]: for every $K > 0$ there is a constant $c_0 = c_0(K)$ such that

$$C(k,d) \leq c_0 K^{-d} k^{k+(3d-1)/2} \tag{3}$$

for every d, $-1 \leq d \leq \binom{k}{2} - k$. In fact, we could manage with considerably cruder bounds than (3) but the calculations become more pleasant if $k^{k+(3d-1)/2}$ is multiplied by a factor tending to 0 as $d \to \infty$, rather than by one increasing with d. We could also use the fact that $C(k,d)$ is long concave as a function of d. This was proved by Odlyzko, answering a question of mine, by making use of the proof in [**13**].

A component of a graph is said to be a (k,d)-*component* if it has k vertices and $k+d$ edges. We denote by $X(k,d)$ the *number of* (k,d)-*components* of a random graph. Note that in $\mathcal{G}(n,p)$ the expectation of $X(k,d)$ is

$$E_p(X(k,d)) = \binom{n}{k} C(k,d) p^{k+d}(1-p)^{k(n-k)+\binom{k}{2}-k-d}. \tag{4}$$

In particular, the expected number of tree-components of G_p is

$$E_p(X(k,-1)) = \binom{n}{k} k^{k-2} p^{k-1}(1-p)^{kn-k^2/2-3k/2+1}. \tag{5}$$

For the sake of convenience we shall omit the integrality signs throughout the paper. It is easily seen that the validity of the arguments will remain unaffected. Furthermore, our inequalities are asserted to hold if n is sufficiently large. Finally, $c_1, c_2, \ldots$ denote positive constants.

3. Gaps in the sequences of components. The key result in proving the sudden emergence of the giant component and in estimating its order rather precisely is that from shortly after time $n/2$ most graph processes never have a component of order between $n^{2/3}/2$ and $n^{2/3}$. The restriction $t \leq 2n/3$ in the result below is only for the sake of convenience, it can be easily removed.

THEOREM 1. *Let $s_0 = (\frac{6}{7}\log n)^{1/2} n^{2/3}$ and $t_0 = n/2 + s_0$. Then a.e. graph process $\tilde{G} = (G_t)_0^\infty$ is such that if $t_0 \leq t \leq 2n/3$ and G_t has a component of order k then either $k < n^{2/3}/2$ or else $k > n^{2/3}$.*

PROOF. Throughout the proof we shall assume that $n^{2/3}/2 \leq k \leq n^{2/3}$. Denote by $E_t(k)$ the expected number of components of order k in G_t: $E_t(k) =$

$\sum\{E_t(X(k,d))\colon\ -1 \le d \le \binom{k}{2} - k\}$. In order to prove our theorem it suffices to show that

$$\sum_{t=t_0}^{2n/3} \sum_{k=n^{2/3}/2}^{n^{2/3}} E_t(k) = o(1).$$

For $d > \binom{k}{2} - k$ we have $X(k,d) = 0$ and for $k + d \le t$ and $-1 \le d \le \binom{k}{2} - k$ the expected number of (k,d)-components of G_t is

$$E_t(X(k,d)) = \binom{n}{k} C(k,d) \binom{\binom{n-k}{2}}{t-k-d} \Big/ \binom{N}{t}. \tag{6}$$

Indeed, having chosen the $k+d$ edges of a (k,d)-component, the remaining $t-k-d$ edges have to be chosen from a set of $\binom{n-k}{2}$ edges. Set

$$E_t'(k) = \sum_{d=-1}^{10k} E_t(X(k+d)) \quad \text{and} \quad E_t''(k) = E_t(k) = E_t'(k).$$

We shall estimate separately the sums of the $E_t'(k)$'s and $E_t''(k)$'s. In the first case we shall make use of (3), but in the second case it will suffice to bound $C(k,d)$ by the total number of graphs with k labelled vertices and $k+d$ edges. In both cases the initial difficulty is that (6) is considerably more unpleasant than (4) with $p = t/N$.

(i) Suppose $\frac{1}{2}n^{2/3} \le k \le n^{2/3}$, $t_0 \le t \le 2n/3$ and $-1 \le d \le 10k$. Then

(7)
$$\begin{aligned}
\binom{\binom{n-k}{2}}{t-k-d} \Big/ \binom{N}{t} &= \frac{(t)_{k+d}\binom{n-k}{2}_{t-k-d}}{(N)_t} \\
&\le (1+o(1)) \left(\frac{t}{N}\right)^{k+d} \exp\left\{-\frac{(k+d)^2}{2t}\right\} \frac{\binom{n-k}{2}_{t-k-d}}{(N)_{t-k-d}} \\
&\le c_1 \left(\frac{t}{N}\right)^{k+d} \exp\left\{-\frac{(k+d)^2}{2t} - \frac{kn-k^2/2}{N}(t-k-d) \right. \\
&\qquad\qquad \left. - \frac{(kn-k^2/2)^2}{2N^2}(t-k-d)\right\} \\
&\le c_2 \left(\frac{t}{N}\right)^{k+d} \exp\left\{-\frac{kn-k^2/2}{N}t + (k+d)\left[\frac{kn-k^2/2}{N} - \frac{k+d}{2t}\right] - \frac{k^2n^2}{2N^2}t\right\} \\
&\le c_3 \left(\frac{t}{N}\right)^{k+d} \exp\left\{-\frac{kn-k^2/2}{N}t + 2(k+d)k/n - \frac{(k+d)^2}{2t} - 2k^2t/n^2\right\} \\
&\le c_3 \left(\frac{t}{N}\right)^{k+d} \exp\left\{-\frac{kn-k^2/2}{N}t\right\}.
\end{aligned}$$

The last inequality holds since for fixed values of n, k and d the minimum of $(k+d)^2/2t + 2k^2t/n^2$ is attained at $t = (k+d)n/2k$ and the minimum is exactly $2(k+d)k/n$.

Relations (3), (6) and (7) imply that

$$E_t'(k) \le c_4 \binom{n}{k} \sum_{d=-1}^{10k} 2^{-d} k^{k+(3d-1)/2} \left(\frac{t}{N}\right)^{k+d} \exp\left\{-\frac{kn-k^2/2}{N}t\right\}$$
$$\le c_5 \binom{n}{k} k^{k-2} \left(\frac{t}{N}\right)^{k-1} \exp\left\{-\frac{kn-k^2/2}{N}t\right\}.$$

Setting $s = t - n/2$ and $\varepsilon = 2s/n$ (so that t/N is well approximated by $(1+\varepsilon)/n$ and $\varepsilon \le \frac{1}{3}$) we find that

$$E_t'(k) \le c_6 n k^{-5/2} \exp\left\{-\frac{k^2}{2n} + k + \varepsilon k - \frac{\varepsilon^2}{2}k + \frac{\varepsilon^3}{3}k - k + \frac{k^2}{2n} - \varepsilon k + \frac{\varepsilon k^2}{2n}\right\}$$
$$\le c_6 n k^{-5/2} \exp\left\{-\frac{2\varepsilon^2}{5}(1-\varepsilon)k\right\} = c_6 n k^{-5/2} \exp\left\{-\frac{8}{5}s^2\left(1-\frac{2s}{n}\right)k/n^2\right\}.$$

Consequently

$$\sum_{k=n^{2/3}/2}^{n^{2/3}} E_t'(k) \le c_7 n^{-2/3} \exp\left\{-\frac{4}{5}s^2\left(1-\frac{2s}{n}\right)n^{-4/3}\right\} n^2 s^{-2}$$

and so

$$\sum_{s=s_0}^{n/6} \sum_{k=n^{2/3}/2}^{n^{2/3}} E_t'(k) \le c_8 n^{4/3} s_0^{-2} \exp\left\{-\frac{4}{5}s_0^2\left(1-\frac{2s_0}{n}\right)n^{-4/3}\right\} n^{4/3}/s_0$$
$$\le c_9 n^{8/3} s_0^{-3} \exp\left\{-\frac{4}{5}s_0^2 n^{-4/3}\right\}$$
$$\le c_{10} n^{2/3} (\log n)^{-3/2} \exp\left\{-\frac{4}{5}\left(\frac{6}{7}\log n\right)\right\} \quad .$$
$$= o(1).$$

(ii) Suppose $\frac{1}{2}n^{2/3} \le k \le n^{2/3}$, $t_0 \le t \le 2n/3$ and $10k \le d \le \binom{k}{2} - k$. In this case rather crude estimtes will suffice:

$$\binom{\binom{n-k}{2}}{t-k-d} \Big/ \binom{N}{t} \le \binom{N}{t-k-d} \Big/ \binom{N}{t} \le \left(\frac{t}{N}\right)^{k+d} \le \left(\frac{4}{n}\right)^{k+d}$$

and

$$C(k,d) \le \binom{\binom{k}{2}}{k+d} \le \left(\frac{ek^2}{2(k+d)}\right)^{k+d}.$$

Therefore by (6)

$$E_t(X(k,d)) \le \left(\frac{en}{k}\right)^k \left(\frac{ek^2}{2(k+d)}\right)^{k+d} \left(\frac{4}{n}\right)^{k+d} \le \left(\frac{k}{n}\right)^d \le n^{-d/3}.$$

This implies that $E_t''(k) = o(n^{-3})$ and so

$$\sum_{t=t_0}^{2n/3} \sum_{k=n^{2/3}/2}^{n^{2/3}} E_t''(k) = o(n^{-1}).$$

Let us call a component *small* if it has fewer than $n^{2/3}/2$ vertices and *large* if it has more than $n^{2/3}$ vertices. Theorem 1 states that a.e. graph process is such that for $t_0 \le t \le 2n/3$ every component of G_t is either small or large. This will enable us to prove that a.e. graph process such that shortly after time $n/2$ it has a unique large component (a 'giant' component), and all other components are small. In the sequel we shall need a fairly good upper bound for the variance of the number of small components.

THEOREM 2. *Suppose* $\Lambda \subset \mathbf{N} \times \{\mathbf{N} \cup \{0,-1\}\}$, $-1 < \varepsilon = \varepsilon(n) \le n/4$ *and* $p = (1+\varepsilon)/n$. *Set*

$$X_i = \sum\{k^i X(k,d)\colon\ (k,d) \in \Lambda\}, \quad \mu_i = E_p(X_i)$$

and $\tilde{k} = \max\{k\colon\ (k,d) \in \Lambda\}$. *If* $-1 < \varepsilon \le -(1+\varepsilon)^2/n$ *then* $\sigma^2(X_i) \le \mu_{2i}$ *and if* $-(1+\varepsilon)^2/n \le \varepsilon$ *and* $\tilde{k}^2(\varepsilon + (1+\varepsilon)^2/n) \le n$ *then*

$$\sigma^2(X_i) \le \mu_{2i} + \frac{2}{n}(\varepsilon + (1+\varepsilon)^2/n)\mu_{i+1}^2.$$

PROOF. By relation (4)

$$E_p(X_i) = \sum_{(k,d)\in\Lambda} k^i \binom{n}{k} C(k,d) p^{k+d} (1-p)^{k(n-k)+\binom{k}{2}-k-d}.$$

Also, if (k_1,d_1) and (k_2,d_2) are distinct elements of Λ then

$$\begin{aligned} E_p(X(k_1,d_1)(k_2,d_2)) &= \binom{n}{k_1}\binom{n-k_1}{k_2} C(k_1,d_1)C(k_2,d_2) \\ &\quad \cdot p^{k_1+k_2+d_1+d_2}(1-p)^{(k_1+k_2)(n-k_1-k_2)+\binom{k_1+k_2}{2}-k_1-k_2-d_1-d_2} \\ &= E_p(X(k_1,d_1))E_p(X(k_2,d_2))\frac{(n)_{k_1+k_2}}{(n)_{k_1}(n)_{k_2}}(1-p)^{-k_1k_2}. \end{aligned} \tag{8}$$

Similarly, for $(k,d) \in \Lambda$,

$$E_p(X(k,d)^2) \le E_p(X(k,d)) + E_p(X(k,d))^2 \frac{(n)_{2k}}{(n)_k(n)_k}(1-p)^{-k^2}. \tag{9}$$

Relations (8) and (9) give

$$\begin{aligned} E_p(X_i^2) \le E_p(X_i) + \sum\Big\{ & k_1^i k_2^i E(X(k_1,d_1))E(X(k_2,d_2)) \\ & \cdot \frac{(n)_{k_1+k_2}}{(n)_{k_1}(n)_{k_2}}(1-p)^{-k_1k_2}\colon\ (k_1,d_1),(k_2,d_2) \in \Lambda \Big\} \end{aligned} \tag{10}$$

Since $1 - y \le (1-x)e^{x-y}$ whenever $0 \le x \le y \le 1$, we find that

$$\begin{aligned} \frac{(n)_{k_1+k_2}}{(n)_{k_1}(n)_{k_2}} &= \prod_{i=0}^{k_1-1}\left(1 - \frac{k_2+i}{n}\right) \Big/ \left(1 - \frac{i}{n}\right) \\ &\le \exp\left\{\sum_{i=1}^{k_1-1}\left(\frac{i}{n} - \frac{k_2+i}{n}\right)\right\} = \exp\{-k_1k_2/n\}. \end{aligned} \tag{11}$$

Furthermore,

$$1-p=1-\frac{1+\varepsilon}{n}\geq \exp\left\{-\frac{1+\varepsilon}{n}-\frac{(1+\varepsilon)^2}{n^2}\right\},$$

so by (11)

$$\frac{(n)_{k_1+k_2}}{(n)_{k_1}(n)_{k_2}}(1-p)^{-k_1k_2}\leq \exp\left\{-\frac{k_1k_2}{n}+\frac{(1+\varepsilon)k_1k_2}{n}+\frac{(1+\varepsilon)^2k_1k_2}{n^2}\right\}$$
$$=\exp\left\{\frac{k_1k_2}{n}\left(\varepsilon+\frac{(1+\varepsilon)^2}{n}\right)\right\}=1+\delta(k_1,k_2).$$

Now if $\varepsilon+(1+\varepsilon)^2/n\leq 0$ then $\delta(k_1,k_2)\leq 0$ and so by (10)

$$\sigma^2(X_i)\leq E_p(X_{2i})=\mu_{2i}.$$

Finally, if $-(1+\varepsilon)^2/n\leq\varepsilon$ and $\tilde{k}^2(\varepsilon+(1+\varepsilon)^2/n)\leq n$ then

$$\delta(k_1,k_2)\leq\frac{2k_1k_2}{n}\left(\varepsilon+\frac{(1+\varepsilon)^2}{n}\right).$$

Therefore (10) gives

$$\sigma^2(X_i)\leq E(X_{2i})+\left(\frac{2\varepsilon}{n}+\frac{2(1+\varepsilon)^2}{n^2}\right)\sum\{k_1^{i+1}k_2^{i+1}E(X(k_1,d_1))E(X(k_2,d_2))\colon (k_1,d_1),(k_2,d_2)\in\Lambda\}$$
$$=\mu_i+\left(\frac{2\varepsilon}{n}+\frac{2(1+\varepsilon)^2}{n^2}\right)\mu_{i+1}^2.\quad\square$$

Armed with Theorem 2, we can locate the maximal order of a component of G_p having fewer than $n^{2/3}$ vertices, provided p is too close to the critical value $1/n$. As for $p>1/n$ this will turn out to be the order of the second largest component, we denote it by $S(G_p)$:

$$S(G_p)=\max\{k\colon\ k=0\text{ or }k\leq n^{2/3}\text{ and }G_p\text{ has a component of order }k\}.$$

THEOREM 3. *Let* $-\frac{1}{4}<\varepsilon=\varepsilon(n)<\frac{1}{4}$, $p=(1+\varepsilon)/n$ *and define*

$$g_\varepsilon(k)=\log n-\frac{5}{2}\log k+k(\log(1+\varepsilon)-\varepsilon)+2\log(1/\varepsilon).$$

Let $k_0=k_0(n)$ *and* $k_2=k_2(n)<n^{2/3}$ *be such that*

$$g_\varepsilon(k_0)\to\infty\quad\text{and}\quad g_\varepsilon(k_2)\to-\infty.$$

Then a.e. G_p *is such that* $S(G_p)<k_2$ *and if* $n|\varepsilon|^3(\log n)^{-2}\to\infty$ *then a.e.* G_p *is such that* $k_0<S(G_p)$.

PROOF. (i) As in the proof of Theorem 1, for $k\leq n^{2/3}$

$$\sum_{d\geq -1}E_p(X(k,d))\leq c_1E_p(X(k,-1)).$$

Hence the expected number of components of G_p with order between k_2 and $n^{2/3}$ is at most

$$\begin{aligned}
c_1 \sum_{k=k_2}^{n^{2/3}} & E_p(X(k,-1)) \\
&\le c_2 n \sum_{k=k_2}^{n^{2/3}} k^{-5/2} \exp\{k - k^2/2n\}(1+\varepsilon)^k \left(1 - \frac{1+\varepsilon}{n}\right)^{kn-k^2/2} \\
&\le c_2 n \sum_{k=k_2}^{n^{2/3}} k^{-5/2} \exp\left\{k - \frac{k^2}{2n} + k\log(1+\varepsilon) - k(1+\varepsilon) - \frac{k^2}{2n}(1+\varepsilon)\right\} \\
&\le c_3 n \sum_{k=k_2}^{n^{2/3}} k^{-5/2} \exp\{k(\log(1+\varepsilon) - \varepsilon)\} \\
&\le c_4 n k_2^{-5/2} \exp\{k_2(\log(1+\varepsilon) - \varepsilon)\}\varepsilon^{-2} \\
&= c_4 \exp\{g_\varepsilon(k_2)\}.
\end{aligned}$$

By our choice of k_2 we have $g_\varepsilon(k_2) \to -\infty$ so almost no G_p has a component whose order is between k_2 and $n^{2/3}$.

(ii) In the proof of the second inequality we may and shall assume that $n|\varepsilon|^3(\log n)^{-2} \to \infty$ and $k_0 \to \infty$. Set $k_\varepsilon = \lfloor 8(\log n)\varepsilon^{-2} \rfloor$, $\Lambda = \{(k,-1)\colon k_0 \le k \le k_\varepsilon\}$ and let X_0 be as in Theorem 2. Then $g_\varepsilon(k_\varepsilon) \to -\infty$ and

$$\mu_0 = E(X_0) = \sum_{k=k_0}^{k_\varepsilon} E(T_k) \sim \frac{n}{\sqrt{2\pi}} \sum_{k=k_0}^{k_\varepsilon} k^{-5/2} \exp\{k(\log(1+\varepsilon) - \varepsilon)\}.$$

Clearly $k_\varepsilon > k_0 + \varepsilon^2$ so by the choice of k_0 we have $\mu_0 \to \infty$. Furthermore, we may suppose that μ_0 does not grow too fast, say $\mu_0 = o(n|\varepsilon|^3(\log n)^{-2})$. Then

$$\mu_0 \varepsilon k_\varepsilon^2/n = O(\mu_0 \varepsilon (\log n)^2 \varepsilon^{-4} n^{-1}) = o(1).$$

Theorem 2 implies that

$$\sigma^2(X_0) \le \mu_0 + 3\varepsilon\mu_1^2/n \le \mu_0 + 3\varepsilon\mu_0^2 k_\varepsilon^2/n = \mu_0(1+o(1)),$$

so by Chebyshev's inequality $P(X > 0) \to 1$. $\square$

Let us state some explicit bounds for $S(G_p)$ implied by Theorem 3.

COROLLARY 4. *Let* $p = (1+\varepsilon)/n$.
(i) *If* $0 < \varepsilon < \frac{1}{4}$ *is fixed then a.e.* G_p *satisfies* $S(G_p) \le 3(\log n)\varepsilon^{-2}$,
(ii) *If* $n^{-\gamma_0} \le \varepsilon = o((\log n)^{-1})$ *and* $\omega(n) \to \infty$ *then*

$$|S(G_p) - (2\,\log n + 6\,\log\varepsilon - 5\,\log\log n)\varepsilon^{-2}| \le \omega(n)\varepsilon^{-2}$$

for a.e. G_p.

PROOF. (i) All we have to check is that $g_\varepsilon(3(\log n)\varepsilon^{-2}) \to -\infty$.
(ii) Straightforward calculations show that

$$k_0 = (2\,\log n + 6\,\log\varepsilon - 5\,\log\log n - \omega(n))\varepsilon^{-2}$$

and

$$k_2 = (2 \log n + 6 \log \varepsilon - 5 \log\log n + \omega(n))\varepsilon^{-2}$$

satisfy the conditions in Theorem 3. $\square$

The corollary above enables us to prove an analog of Theorem 1 for $t \geq \frac{2}{3}n$.

THEOREM 5. *A.e. graph process* $\tilde{G} = (G_t)_0^N$ *is such that for* $t \geq 5n/8$ *the graph* G_t *has no small component of order at least* $100 \log n$.

PROOF. By Corollary 4 a.e. graph process is such that for some t satisfying $3n/5 < t < 5n/8$ the graph G_t has no small component whose order is at least $75 \log n$. Since the union of two components of order at most $a = \lceil 100 \log n \rceil$ has order at most $2a < n^{2/3}$, the assertion of the theorem will follow if we show that a.e. graph process is such that for $t \geq 3n/5$ the graph G_t has no component whose order is at least a and at most $2a$. Furthermore, since for $t \geq 2n \log n$ a.e. G_t is connected, it suffices to prove this for $t \leq 2n \log n$.

Let $\frac{3}{5}n \leq t \leq 2n \log n$ and set $c = 2t/n$. Then the expected number of components of G_t having order at least a and at most $2a$ is

$$\begin{aligned} E\left(\sum_{k=a}^{2a}\sum_{d\geq -1} X(k,a)\right) &= \sum_{k=a}^{2a}\binom{n}{k}\sum_{d\geq -1} C(k,d)\binom{\binom{n-k}{2}}{t-k-d}\Big/\binom{N}{t} \\ &\leq c_0 \sum_{k=a}^{2a}(en)^k k^{-5/2}\binom{\binom{n-k}{2}}{t-k+1}\Big/\binom{N}{t} \\ &\leq c_1 \sum_{k=a}^{2a}(en)^k k^{-5/2}\left(\frac{t}{N}\right)^{k-1}\left(\frac{N-kn}{N}\right)^{t-k} \\ &\leq c_2 n \sum_{k=a}^{2a} e^k k^{-5/2} c^k e^{-ck} \\ &= c_2 n \sum_{k=a}^{2a} k^{-5/2}(ce^{1-c})^k, \end{aligned}$$

where c_0, c_1 and c_2 are absolute constants. Since for $c = 1.2$ we have $c - 1 - \log c > 0.0176$, the last expression is at most

$$c_2 n (\log n)^{-3/2} n^{-1.7} = o(n^{-1/2}). \quad \square$$

4. The emergence of the giant component. Given a graph process $\tilde{G} = (G_t)_0^\infty$, denote by w_t the number of components of G_t and let

$$V = \bigcup_{i=1}^{w_t} U_i(G_t)$$

be the partition of V into vertex sets of its components. Note that for every t either G_t and G_{t+1} define the same partition of V or else $w_{t+1} = w_t - 1$ and the partition defined by G_t is a refinement of the partition defined by G_{t+1}. Hence if $\tilde{G}$ is such that for $t \geq t_0$ the graph G_t does not have a component whose order is between $n^{2/3}/2$ and $n^{2/3}$ then for $t_0 \leq t \leq t'$ the graph $G_{t'}$ has at most as many components of order at least $n^{2/3}$ as G_t.

Moreover, if in such a $\tilde{G}$ the graph G_t has a component which contains all large components of G_{t_0} then G_t has a unique component of order at least $n^{2/3}$ and so has $G_{t'}$ for every $t' \geq 1$. Therefore in order to establish the existence of the giant component, we have to show that the large components of G_{t_0} are contained in a single component of G_t for some $t > t_0$. Furthermore, we have to estimate the number of vertices on the giant component or, equivalently, the number of vertices on the small components. We start with the second task. For the sake of convenience we take $\varepsilon = pn - 1 = n^{-\gamma}$.

THEOREM 6. *Let $0 < \gamma < \frac{1}{3}$, $\varepsilon = n^{-\gamma}$, $p = (1+\varepsilon)/n$ and $\omega(n) \to \infty$. For a graph G, denote by $Y_1(G)$ the number of vertices on the small (k,d)-components with $d \geq 1$, by $Y_0(G)$ the number of vertices on the small unicyclic components and by $Y_{-1}(G)$ the number of vertices on the small tree-components. Then a.e. G_p is such that*

$$Y_1 \leq \omega(n) n^{5\gamma - 1}, \quad Y_0 \leq \omega(n) n^{2\gamma}$$

and

$$Y_{-1} = n - 2n^{1-\gamma} + O(\omega(n) n^{(1+\gamma)/2} + (\log n) n^{1-2\gamma}).$$

PROOF. Set $k_3 = 9(\log n) n^{2\gamma}$. Then for $k_3 \leq k \leq n^{2/3}$ we have

$$\varepsilon^2 k/2 - k\varepsilon^3/3 - k^2\varepsilon/(2n) \geq \varepsilon^2 k/3 \geq 3 \log n,$$

so

$$\begin{aligned}
\sum_{k=k_3}^{n^{2/3}} \sum_{d=-1}^{\binom{k-1}{2}} E(kX(k,d)) &= O\left(\sum_{k=k_3}^{n^{2/3}} E(kX(k,-1)) \right) \\
&= O\left(n \sum_{k=k_3}^{n^{2/3}} k^{-3/2} \exp\left\{ -\varepsilon^2 k/2 + k\varepsilon^3/3 + k^2\varepsilon/(2n) \right\} \right) \\
&= O\left(n^{-2} \sum_{k=k_3}^{n^{2/3}} k^{-3/2} \right) = o(n^{-2}).
\end{aligned}$$

This shows that in our proof we may replace Y_i by

$$\tilde{Y}_i = \sum_{k \leq k_3} kX(k,i), \qquad i = -1, 0, 1.$$

(i) Note first that by (3)

$$\begin{aligned}
E(\tilde{Y}_1) &= O\left(\sum_{k \leq k_3} k \binom{n}{k} C(k,1) p^{k+1} (1-p)^{kn - k^2/2} \right) \\
&= O\left(n^{-1} \sum_{k \leq k_3} k^{3/2} \exp\{-\varepsilon^2 k/2\} \right),
\end{aligned}$$

since for $1 \leq k \leq k_3$ we have

$$k\varepsilon^3/3 + k^2\varepsilon/(2n) = O(1).$$

Hence

$$E(\tilde{Y}_1) = O(n^{-1}\varepsilon^{-5}) = O(n^{5\gamma-1}),$$

implying the first assertion.

(ii) The expectation of $\tilde{Y}_0$ is easily estimated:

$$\begin{aligned} E(\tilde{Y}_0) &= \sum_{k\le k_3} kE(X(k,0)) \\ &= \sum_{k\le k_3} k\binom{n}{k}C(k,k)\left(\frac{1+\varepsilon}{n}\right)^k\left(1-\frac{1-\varepsilon}{n}\right)^{kn-k^2/2+3k/2} \\ &\sim \frac{1}{\sqrt{2\pi}}\sqrt{\pi/8}\sum_{k\le k_3}\exp\{-k\varepsilon^2/2\} \\ &\sim \frac{1}{4}\int_0^\infty \exp\{-x\varepsilon^2/2\}\,dx \sim \frac{1}{2}n^{2\gamma}. \end{aligned}$$

Therefore, by Chebyshev's inequality, $\tilde{Y}_0 \le \omega(n)n^{2\gamma}$ for a.e. random graph G_p.

(iii) We shall estimate $E(\tilde{Y}_{-1})$ rather precisely and then we shall make use of Theorem 2 to conclude that for a.e. G_p the variable Y_{-1} is rather close to its expectation.

Set $p' = (1-\varepsilon)/n$ and write E' for the expectation in $\mathcal{G}(n,p')$. We shall exploit the fact that E and E' are rather closely related.

First of all, calculations analogous to those in (i) and (ii) show that

$$E'(n-\tilde{Y}_0-\tilde{Y}_{-1}) = o(n^{2\gamma}) \quad \text{and} \quad E'(Y_0) \sim \tfrac{1}{2}n^{2\gamma}.$$

Hence

$$E'(\tilde{Y}_{-1}) = n - \tfrac{1}{2}n^{2\gamma} - o(n^{2\gamma}).$$

In order to pass from $E'(\tilde{Y}_{-1})$ to $E(\tilde{Y}_{-1})$, note that for $k \le k_3$ we have

$$\begin{aligned} E(X(k,-1))/E'(X(k,-1)) &= \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^{k-1}\left(\frac{1-(1+\varepsilon)/n}{1-(1-\varepsilon)/n}\right)^{kn-k^2/2+3k/2-1} \\ &= \exp\left\{2(k-1)\varepsilon - (2kn-k^2)\varepsilon/n + O((\log n)n^{-2\gamma})\right\} \\ &= \exp\left\{-2\varepsilon + k^2\varepsilon/n + O((\log n)n^{-2\gamma})\right\} \\ &= 1 - 2\varepsilon + O(k^2\varepsilon/n) + O((\log n)n^{-2\gamma}). \end{aligned}$$

Consequently

$$\begin{aligned} E(\tilde{Y}_{-1}) &= (1-2\varepsilon)E'(\tilde{Y}_{-1}) + O((\log n)n^{1-2\gamma}) \\ &\quad + O\left(\frac{\varepsilon}{n}\sum_{k\le k_3} k^3\binom{n}{k}E'(X(k,-1))\right) \\ &= n - 2n^{1-\gamma} + O(n^{2\gamma}) + O((\log n)n^{1-2\gamma}) \\ &\quad + O\left(\varepsilon\sum k^{1/2}\exp\{-k\varepsilon^2/2\}\right) \\ &= n - 2n^{1-\gamma} + O(n^{2\gamma} + (\log n)n^{1-2\gamma}). \end{aligned} \tag{12}$$

Finally, in order to apply Theorem 2 we have to estimate the following expectation:

$$E\left(\sum_{k\le k_3} k^2 X(k,-1)\right) = O\left(n\sum_{k\le k_3} k^{-1/2}\exp\{-k\varepsilon^2/2\}\right) = O(n^{1+\gamma}).$$

By Theorem 2 we have $\sigma^2(\tilde{Y}_{-1}) = O(n^{1+\gamma})$ so by Chebyshev's inequality a.e. G_p satisfies

$$|\tilde{Y}_{-1} - E(\tilde{Y}_{-1})| \le \omega(n)n^{(1+\gamma)/2}. \tag{13}$$

Relations (12) and (13) imply that

$$\tilde{Y}_{-1} = n - 2n^{1-\gamma} + O(\omega(n)n^{(1+\gamma)/2} + (\log n)n^{1-2\gamma}),$$

as claimed. □

Let us establish now the emergence of the giant component shortly after time $n/2$.

THEOREM 7. *A.e. graph process $\tilde{G} = (G_t)_0^N$ is such that for every $t \ge t_1 = n/2 + (\log n)^{1/2} n^{2/3}$ the graph G_t has a unique component of order at least $n^{2/3}$. The other components of G_t have at most $n^{2/3}/2$ vertices each.*

PROOF. As before, we call a component *small* if it has fewer than $n^{2/3}/2$ vertices, and *large* if it has at least $n^{2/3}$ vertices. By Theorem 1 a.e. graph process $\tilde{G}$ is such that for $t \ge t_0 = n/2 + s_0$ every component of G_t is either small or large, where t_0 is as defined in Theorem 1. Let $C_1, C_2, \ldots, C_l$ be the large components of G_{t_0} in such a $\tilde{G}$ and suppose in some G_t, $t \ge t_0$, all the C_i's are contained in the same component. Then this component of G_t is the unique large component and for $t' \ge t$ the graph $G_{t'}$ has also a unique large component. Indeed, as t increases from t_0, a vertex $x \in V$ can become a vertex of a large component only if that component contains a C_i, for the union of two small components contains fewer than $n^{2/3}$ vertices. Consequently our theorem will follow if we show that a.e. $\tilde{G}$ is such that for some t between t_0 and t_1 all large components of G_{t_0} are contained in the same component of G_t. Imitating the proof of Theorem 6, one can show that a.e. $\tilde{G}$ is such that G_{t_0} has at least $2s_0 > (\log n)^{1/2} n^{2/3}$ vertices on its large components. (In fact, the expected number of these vertices is about $4s_0$.) Therefore one can find disjoint subsets $V_1, V_2, \ldots, V_m$ of $V(G_{t_0}) = V$ such that

$$m \ge (\log n)^{1/2}/2, \quad |V_i| \ge n^{2/3}, \qquad i = 1, \ldots, m,$$

each V_i is contained in some $V(C_j)$ and

$$\bigcup_{i=1}^m V_i = \bigcup_{j=1}^l V(C_j),$$

where $C_1, \ldots, C_l$ are the large components of G_{t_0}.

Set $p = n^{-4/3}$ and denote by H_p the random graph obtained from G_{t_0} by adding to it edges independently and with probability p. Then a.e. H_p has at most $t_0 + n^{2/3} < t_1$ edges so it suffices to show that $\bigcup_{i-1}^m V_i$ is contained in a single component in a.e. H_p.

What is the probability that for a given pair (i, j), $1 \leq i < j \leq m$, some edge of H_p joins V_i to V_j? It is clearly at least

$$1 - (1-p)^{n^{4/3}} \geq 1 - e^{-1} > 1/2.$$

Hence the probability that in H_p all V_i's are contained in the same component is at least the probability that an element of $G(m, \frac{1}{2})$ is connected. Since $m \to \infty$, this probability tends to 1. (See [**5** or **2**, p. 140] for considerably stronger results.) This completes our proof. $\square$

Combining Theorems 3, 6 and 7, we obtain rather precise information about the orders of the largest components. Since the property of having at most x vertices on the small tree components is monotone, and so are the properties of having at least y vertices on the large components and at most z vertices on the small components which are trees or unicyclic graphs, using [**2**, Theorem 8, p. 133] we may pass from the model $\mathcal{G}(n,p)$ to the model $\mathcal{G}(n, M)$. Denote by $L_r(G)$ the rth largest order of a component of a graph G.

THEOREM 8. *Let $0 < \gamma < \frac{1}{3}$, $s = \frac{1}{2}n^{1-\gamma}$ and $t = n/2 + s$. Then for every $m \in \mathbf{N}$ and $\omega(n) \to \infty$ a.e. G_t is such that*

$$L_1(G) = 4s + O(\omega(n)n/s^{1/2} + (\log n)s^2/n)$$

and

$$k_0 < L_m(G) \leq L_{m-1}(G) \leq \cdots \leq L_2(G) < k_2,$$

where k_0 and k_2 are as in Theorem 3. $\square$

Before extending the range of t in the theorem above, we investigate the distribution of the small components of G_p in the case when p is not as close to $1/n$ as has been required so far.

5. Components of order less than $n^{2/3}$. First we consider the small components of G_p with $p < 1/n$.

THEOREM 9. *Let $0 < \varepsilon = \varepsilon(n) = o(1)$ be such that $\varepsilon n^\eta \to \infty$ for every fixed $\eta > 0$ and set $p = (1-\varepsilon)/n$. Given $\lambda \in \mathbf{R}^+$, choose $l_\lambda = l_\lambda(n)$ in such a way that*

$$\mu(n, \varepsilon, l_\lambda) = (2/\pi)^{1/2} n l_\lambda^{-5/2}((1-\varepsilon)e^\varepsilon)^{l_\lambda}\varepsilon^{-2} \to \lambda$$

and denote by $Z = Z(G_p)$ the number of components of G_p having at least l_λ vertices. Then the distribution of Z tends to P_λ, the Poisson distribution with mean λ.

PROOF. The assumptions imply easily that $l_\lambda \to \infty$, $l_\lambda = o(n^\eta)$ for every $\eta > 0$ and the expected number of vertices on components containing cycles is bounded. Hence we may assume that Z is the number of tree-components of order at least l_λ. Since also $l_\lambda \varepsilon^2 \to \infty$, a trite calculation shows that

$$\begin{aligned} E(Z) &= \sum_{k=l_\lambda}^{n} E(X(k,-1)) \sim \frac{n}{\sqrt{2\pi}} \sum_{k=l_\lambda}^{n} k^{-5/2}((1-\varepsilon)e^\varepsilon)^k \\ &\sim \frac{n}{\sqrt{2\pi}} l_\lambda^{-5/2}((1-\varepsilon)e^\varepsilon)^k/(1-(1-\varepsilon)e^\varepsilon) \sim \mu(n,\varepsilon,l_\lambda) \sim \lambda. \end{aligned}$$

Furthermore, by making use of $l_\lambda = o(n^n)$ it is easily shown that for every $r \in \mathbf{N}$ the rth factorial moment

$$E_r(Z) = E(Z(Z-1)\cdots(Z-r+1))$$

converges to λ^r. Hence $Z \xrightarrow{d} P_\lambda$, as claimed. $\square$

COROLLARY 10. *Let ε, p and l_λ be as in Theorem* 9, *and let $\omega(n) \to \infty$. Then for every $m \in \mathbf{N}$ a.e. G_p satisfies*

$$l_1 - \omega(n)/\varepsilon \le L_m(G_p) \le \cdots \le L_1(G_p) \le l_1 + \omega(n)/\varepsilon.$$

PROOF. It is easily seen that for every fixed λ we have

$$l_\lambda \sim (\log n)/(\varepsilon - \log(1/(1-\varepsilon))) \sim 2(\log n)/\varepsilon \quad \text{and} \quad l_1 - \omega(n)/\varepsilon < l_\lambda < l_1 + \omega(n)/\varepsilon.$$

Hence the assertion follows from Theorem 9. $\square$

If $p = c/n$ for some constant $c < 1$ then we need not be able to find numbers $l_\lambda = l_\lambda(n)$ ensuring that $\mu(n, 1-c, l_\lambda) \to \lambda$. Nevertheless, with some simple changes the results above carry over to this case without any difficulty. In fact, our task is easier since the tree-components we have to consider have only $O(\log n)$ vertices. Thus one arrives at the following result of Erdös and Rényi [**6**, p. 49].

If $0 < c < 1$ is a constant, $p = c/n$, $\alpha = c - 1 - \log c$,

$$k_0 = \frac{1}{\alpha}\left\{\log n - \frac{5}{2}\log\log n - l_0\right\} \in \mathbf{N} \quad \text{and} \quad l_0 = O(1)$$

then the number of components of G_p with at least k_0 vertices has asymptotically Poisson distribution with mean

$$\lambda \sim \frac{1}{c\sqrt{2\pi}} \frac{\alpha^{5/2}}{1 - e^{-\alpha}} e^{l_0}.$$

Though the simple method of means is sufficient to prove this assertion, we would like to point out that recently Barbour [**1**] applied a more sophisticated and powerful method to prove that the number of certain components has asymptotically Poisson distribution.

Let us turn to the case $p = (1+\varepsilon)/n > 1/n$. The proof of Theorem 6 is easily adapted to show that if $\varepsilon = o(1)$ then the distribution of the number of components of order less than $n^{2/3}$ in G_p is almost the same as the distribution in $G_{p'}$ with $p' = (1-\varepsilon)/n$. Furthermore, it is easily seen that ε can be rather small for a slightly weaker version of Theorem 6 to remain valid. As in Theorem 8, we state the result for G_t rather than G_p.

THEOREM 11. *Let $t = n/2 + s$, $s = o(n)$, $sn^{-2/3}(\log n)^2 \to \infty$, $\varepsilon = 2s/n$ and for $\lambda > 0$ choose $l_\lambda = l_\lambda(n)$ in such a way that*

$$(2/\pi)^{1/2} n l_\lambda^{-5/2}((1-\varepsilon)e^\varepsilon)^{l_\lambda}\varepsilon^{-2} \to \lambda.$$

Then for $\omega(n) \to \infty$ and $m \in \mathbf{N}$ a.e. G_t is such that

$$L_1(G_t) = (4 + o(1))s,$$
$$l_1 - \omega(n)/\varepsilon \le L_m(G_t) \le \cdots \le L_2(G_t) \le l_1 + \omega(n)/\varepsilon,$$
$$L_i(G) = (1 + o(1))l_1, \qquad i = 2, \ldots, m.$$

If $\varepsilon n^\eta \to \infty$ for every $\eta > 0$ then $Z^* \xrightarrow{d} P_\lambda$, where $Z^* = \max\{m-1\colon L_m \ge l_\lambda\}$. $\square$

6. The small components after time $n/2$. As t increases, our task of locating the orders of the components of G_t becomes easier. Indeed, if $c > 1$ is a constant and $t - cn/2 = o(n)$ then the expected number of vertices on components containing cycles and having fewer than $n^{2/3}$ vertices is bounded and so is the expected number of vertices on components whose order is between $c_1 \log n$ and $n^{2/3}$. Hence if $\omega(n) \to \infty$ then a.e. G_t is such that with the exception of $\omega(n)$ vertices all vertices belong to the giant component or else to trees of order at most $c_1 \log n$. Now the expectation of the number of vertices on these trees is about $c_2 n$ and by a slight variant of Theorem 2 the variance is $O(n(\log n)^2)$. Therefore we are led to the following result.

THEOREM 12. *Let $c > 1$ be a constant and let $t = \lfloor cn/2 \rfloor$, $\omega(n) \to \infty$. Then a.e. G_t is such that, with the exception of at most $\omega(n)$ vertices, all vertices of G_t belong to the giant component or to components which are trees. Furthermore,*

$$\left| L_1(G_t) - n\left\{1 - \frac{1}{c}\sum_{k=1}^{\infty} \frac{k^{k-1}}{k!}(ce^{-c})^k\right\}\right| \le \omega(n) n^{1/2} \log n$$

and if $k_0 = \frac{1}{\alpha}\{\log n - \frac{5}{2}\log\log n - l_0\} \in \mathbf{N}$, where $\alpha = c - 1 - \log c$ and $l_0 = O(1)$, then $Z^* = \max\{m - 1\colon L_m(G_t) \ge k_0\}$ has asymptotically Poisson distribution with mean

$$\lambda \sim \frac{1}{c\sqrt{2\pi}} \frac{\alpha^{5/2}}{1 - e^{-\alpha}} e^{l_0}. \quad \square$$

In the range of t covered by Theorem 11 the giant component increases about four times as fast as t. Another way of proving Theorem 12 would be to establish this fact first and then use it to deduce the assertion about the size of the giant component. What is the expectation of the increase of $L_1(G_t)$ as t changes to $t+1$? The probability that the $(t+1)$st edge will join the giant component to a component of order L_j is about $L_1 L_j / \binom{n}{2}$. Hence the expectation is about

$$\begin{aligned}
&(2L_1/n^2) \sum_{k=1}^{n^{2/3}} k^2 \binom{n}{k} k^{k-2} \left(\frac{1+\varepsilon}{n}\right)^{k-1} \left(1 - \frac{1+\varepsilon}{n}\right)^{kn - k^2/2} \\
&\sim (2L_1/n^2) \sum_{k=1}^{n^{2/3}} \frac{n}{\sqrt{2\pi}} k^{1/2} \exp\{-\varepsilon^2 k/2\} \\
&\sim \left(\frac{2}{\pi}\right)^{1/2} (L_1/n) \int_0^\infty x^{1/2} e^{-x\varepsilon^2/2}\, dx \\
&= \left(\frac{2}{\pi}\right)^{1/2} (L_1/n) \Gamma\left(\frac{1}{2}\right) (\varepsilon^2/2)^{-1/2} = 2L_1/(\varepsilon n) = L_1/s.
\end{aligned}$$

From this one can deduce that if t is $o(n)$ but not too small then $L_1(t) = L_1(n/2 + s) \sim c_1 s$. By considering the crude order of $L_1(G_t)$ as t ceases to be $o(n)$, one can show easily that the constant c_1 is 4.

Using Theorem 2, for $t \ge (1+\varepsilon)n/2$ it is easy to obtain fairly precise information about the distribution of the orders of the components of G_t. We shall do a little more than that: we shall prove some results about all graphs G_t of a graph process after time $(1+\varepsilon)n/2$. Let us start with a rather crude result.

THEOREM 13. *Let $\varepsilon > 0$ be fixed and for $c_0 = c_0(n) \geq 1+\varepsilon$ set $c_1 = 3/(c_0 - 1 - \log c_0)$. Then a.e. graph process $\tilde{G} = (G_t)_0^N$ is such that for $t \geq c_0 n/2$ the graph G_t does not contain a component whose order is between $c_1 \log n$ and $n^{2/3}$.*

PROOF. Let $c_0 \leq c \leq 3 \log n$ and set $p = c/n$. By inequality (3) the expected number of components of G_p whose order k satisfies $k_1 = \lfloor c_1 \log n \rfloor \leq k \leq k_2 = \lfloor n^{2/3} \rfloor$ is at most

$$\sum_{k=k_1}^{k_2} \binom{n}{k} \sum_{d \geq -1} \left(\frac{2}{3}\right)^{d+1} k^{k+(3d-1)/2} \left(\frac{c}{n}\right)^{k+d} \left(1 - \frac{c}{n}\right)^{k(n-k)}$$
$$= O(1) \sum_{k=k_1}^{k_2} \binom{n}{k} k^{k-2} \left(\frac{c}{n}\right)^{k-1} \left(1-\frac{c}{n}\right)^{k(n-k)}$$
$$= o(n) c e^{1-c^{k_1}} = o(n^{-2}).$$

Since the property of containing a component of order k is convex and a.e. graph process becomes connected by time $n \log n$, this implies our assertion. □

THEOREM 14. (i) *Let $c_0 > 1$ be fixed. Then a.e. $\tilde{G}$ is such that for $t \geq c_0 n/2$ the graph G_t does not contain a (k,d)-component with $d \geq 1$ and $k \leq n^{2/3}$.*

(ii) *For $\omega(n) \to \infty$ a.e. $\tilde{G}$ is such that for $t \geq \omega(n) n$ every component of G_t, with the exception of its giant component is a tree.*

PROOF. Since for $t \geq n \log n$ a.e. G_t is connected, it suffices to restrict our attention to the range $c_0 n/2 \leq t \leq n \log n$. Furthermore, by Theorem 11 it suffices to consider components of order at most $k_1 = \lfloor c_1 \log n \rfloor$, where

$$c_1 = 3/(c_0 - 1 - \log c_0).$$

Note that the expected number of (k,d)-components with $4 \leq k \leq k_1$ and $d \geq 1$ a graph process contans between times $t_0 = \lfloor c_0 n/2 \rfloor$ and $t_1 = \lfloor 2n \log n \rfloor$ is at most

$$2 \sum_{t=t_0}^{t_1} \sum_{k=4}^{k_1} \binom{n}{k} \sum_{d \geq 1} \left(\frac{2}{3}\right)^{d+1} k^{k+(3d-1)/2} \left(\frac{2t}{n^2}\right)^{k+d} \left(1 - \frac{2t}{n^2}\right)^{k(n-k)} = O(1).$$

On the other hand, it is easily seen (cf. Theorem 9c of [**6**]) that for $t_0 \leq t \leq t_1 = \lfloor n \log n \rfloor$ and $1 \leq k \leq k_1$ the life-time of a component of order k in G_t has approximately exponential distribution with mean $n/(2k)$. In particular, a component of order k in G_t will be a component of $G_{t+1}, G_{t+2}, \ldots, G_{t+l}$ with probability at least $\frac{1}{2}$, where $l = \lfloor n/(3k_1) \rfloor$. Consequently the probability that $\tilde{G}$ is such that there is a time t with $t_0 \leq t \leq t_1$ for which G_t has a (k,d)-component with $4 \leq k \leq k_1$ and $d \geq 1$ is at most

$$O(1)(2/l) = O((\log n)/n).$$

This proves (i). Assertion (ii) is proved analogously. □

Let $\tilde{k} = O(\log n)$, $\Lambda = \{(k,d)\colon 1 \leq k \leq \tilde{k},\ d = -1 \text{ or } 0\}$, and define the X_i as in Theorem 2. Then for every fixed i there is a constant d_i such that for

$n/2 \le t \le n \log n$ in G_t we have

$$\mu_i = E_t(X_i) \le 2 \sum_{k=1}^{\bar{k}} \binom{n}{k} k^{k-2+i} \left(\frac{2t}{n^2}\right)^{k-1} \left(1 - \frac{2t}{n^2}\right)^{k(n-k)} \le d_i n e^{-2t/n}$$

and, using a slight variant of Theorem 2,

$$\sigma^2(X_i) \le \mu_{2i} + \frac{4t}{m}\mu_{i+1}^2/n \le d_i \mu_{2i}.$$

Consequently by Chebyshev's inequality

$$P_t(|X_i(G_t) - \mu_i| \ge u) \le \frac{d_i \mu_{2i}}{u^2}. \tag{14}$$

This relation enables us to deduce uniform bounds for the X_i. Here we state only a rather simple result about $w_t = w(G_t)$, the number of components of G_t.

THEOREM 15. *Suppose* $0 \le \varepsilon$, $\omega(n) \to \infty$ *and* $c(n) = \log n - \omega(n) \to \infty$. *Then a.e.* $\tilde{G}$ *is such that for every* t *satisfying* $n \le 2t \le c(n)n$ *we have*

$$|w_t - n\beta(2t/n)| < \varepsilon n \beta(2t/n), \tag{15}$$

where

$$\beta(c) = \frac{1}{c} \sum_{k=1}^{\infty} \frac{k^{k-2}}{k!} (ce^{1-c})^k.$$

PROOF. Let $\eta > 0$ and set $t_j = \lfloor (1 + j\eta)n/2 \rfloor$, $j = 0, 1, \ldots, m = \lfloor c(n)/\eta \rfloor$. Put $\bar{k} = \lceil (3 \log n)/(\eta - \log(1 + \eta)) \rceil$ and let Λ and X_i be as above. Then a.e. $\tilde{G}$ is such that for $t_1 \le t \le t_m$ we have $w_t = w(G_t) = X_0(G_t) + 1$. This implies that it suffices to estimate $X_0(G_t)$ instead of $w(G_t)$.

It is easily seen that if $\eta > 0$ is sufficiently small then for $n/2 \le t' \le t \le c(n)n/2$

$$|E_{t'}(X_0) - n\beta(2t/n)| < \frac{\varepsilon}{5} n \beta_{2t/n} \tag{16}$$

and

$$|\beta(2t/n) - \beta(2t'/n)| < \frac{\varepsilon}{t} \beta(2t/n). \tag{17}$$

Note that $\beta(c) \ge e^{1-c}$. Hence by (14) and (16)

$$P_{t_j}\left(|X_0(G_{t_j}) - E_{t_j}(X_0)| \ge \frac{\varepsilon}{5} E_{t_j}(X_0)\right) \le \frac{25 d_0}{\varepsilon^2 E_{t_j}(X_0)} \le \frac{50 d_0}{\varepsilon^2} n^{-1} e^{j\eta}.$$

Since $\sum_{j=1}^{m} e^{j\eta} = o(1)$, a.e. $\tilde{G}$ is such that

$$|X_0(G_{t_j}) - E_{t_j}(X_0)| < \frac{\varepsilon}{5} E_{t_j}(X_0) \tag{18}$$

for every $j = 1, \ldots, m$. A.e. $\tilde{G}$ is such that $w_{t+1}(\tilde{G}) \le w_t(\tilde{G})$ if $t \ge t_1$ so a.e. G is such that $X_0 = X_0(G_t)$ is a monotone decreasing function of t for $t \ge t_1$. Therefore relations (16), (17) and (18) imply that a.e. $\tilde{G}$ satisfies (15) if $t_1 \le t \le c(n)n/2$.

Finally, if η is sufficiently small then a.e. $\tilde{G}$ is such that $|w_t - n/2| < \varepsilon n/5$ and $|\beta(2t/n) - \frac{1}{2}| < \varepsilon/5$ whenever $n/2 \le t \le t_1$, so (15) holds in this range as well.

In Theorem 15 the approximation of w_t becomes more precise as t grows. From inequality (14) one can also obtain approximations to the same degree for every value of t. For example, it is easy to show that if $\omega(n) \to \infty$ then $\tilde{G}$ is such that

$$\left| w_t(\tilde{G}) - \frac{n^2}{2t} \sum_{k=1}^{\infty} \frac{k^{k-2}}{k!} \left(\frac{2t}{n} e^{-2t/n} \right)^k \right| < \omega(n) n^{2/3}$$

for every $t \geq n/2$.

REFERENCES

1. A. D. Barbour, *Poisson convergence and random graphs*, Math. Proc. Cambridge Philos. Soc. **92** (1982), 349–359.
2. B. Bollobás, *Graph theory—An introductory course*, Graduate Texts in Math., Springer-Verlag, New York, Heidelberg and Berlin, 1979.
3. ——, *Random graphs*, Combinatorics (H. N. V. Temperley, ed.), London Math. Soc. Lecture Notes Series, vol. 52, Cambridge Univ. Press, New York, 1981, pp. 80–102.
4. ——, *The evolution of sparse graphs*, Graph Theory and Combinatorics (B. Bollobás, ed.), Academic Press, London, 1984.
5. P. Erdös and A. Rényi, *On random graphs*. I, Publ. Math. Debrecen **6** (1959), 290–297.
6. ——, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci. **5** (1960), 17–61.
7. ——, *On the evolution of random graphs*, Bull. Inst. Internat. Statist. Tokyo **38** (1961), 343–347.
8. C. W. Ford and G. E. Uhlenbeck, *Combinatorial problems in the theory of graphs*, Proc. Nat. Acad. Sci. U.S.A. **43** (1957), 163–167.
9. L. Katz, *Probability of indecomposability of a random mapping function*, Ann. Math. Statist. **26** (1955), 512–517.
10. A. Rényi, *On connected graphs*. I, Publ. Math. Inst. Hungar. Acad. Sci. **4** (1959), 385–387.
11. E. M. Wright, *Asymptotic enumeration of connected graphs*, Proc. Roy. Soc. Edinburgh Sect. A **68** (1968/69), 298–308.
12. ——, *The number of connected sparsely-edged graphs*. II: *Smooth graphs and blocks*, J. Graph Theory **2** (1978), 299–305.
13. ——, *The number of connected sparsely-edged graphs*. III: *Asymptotic results*, J. Graph Theory **4** (1980), 393–407.

DEPARTMENT OF MATHEMATICS, LOUISIANA STATE UNIVERSITY, BATON ROUGE, LOUISIANA 70803