TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 267, Number 1, September 1981

THE DIAMETER OF RANDOM GRAPHS

BY

BÉLA BOLLOBÁS

ABSTRACT. Extending some recent theorems of Klee and Larman, we prove rather sharp results about the diameter of a random graph. Among others we show that if $d = d(n) \geqslant 3$ and $m = m(n)$ satisfy $(\log n)/d - 3 \log\log n \to \infty$, $2^{d-1}m^d/n^{d+1} - \log n \to \infty$ and $d^{d-2}m^{d-1}/n^d - \log n \to -\infty$ then almost every graph with n labelled vertices and m edges has diameter d.

About twenty years ago Erdös [7], [8] used random graphs to tackle problems concerning Ramsey numbers and the relationship between the girth and the chromatic number of a graph. Erdös and Rényi [9], [10] initiated the study of random graphs for their own sake, and proved many beautiful and striking results. The graph invariants investigated in recent years include the clique number [5], [13], [17], the chromatic number [5], [13], the edge chromatic number [11], the circumference [16], [19], and the degree sequence [4]. The aim of this paper is to give rather precise results concerning the diameter. Recall that the diameter diam G of a connected graph is the maximum of the distances between vertices, and a disconnected graph has infinite diameter. The diameter of a random graph has hardly been studied, apart from the case diam $G = 2$ by Moon and Moser [18], the case diam $G < \infty$ by Erdös and Rényi [9], and the diameter of components of sparse graphs by Korshunov [15]. When I was writing this paper, I learned that Klee and Larman [14] proved some results concerning the case diam $G = d$ for fixed values of d. The main result of Klee and Larman [14] is that if $d \geqslant 3$ is a fixed natural number and $m = m(n)$ satisfies

$$m^d/n^{d+1} - \log n \to \infty \quad \text{and} \quad m^{d-1}/n^d \to 0 \quad \text{as } n \to \infty,$$

then almost every labelled graph with n vertices and m edges has diameter d. As a special case of our results we prove that the conditions above can be weakened to

$$2^{d-1}m^d/n^{d+1} - \log n \to \infty \quad \text{and} \quad 2^{d-2}m^{d-1}/n^d - \log n \to -\infty.$$

However, our main aim is to give precise bounds on $m = m(n)$ ensuring that almost every labelled graph with n vertices and m edges has diameter d, where $d = d(n)$ is a function of n *which may tend to* ∞ as $n \to \infty$ but which does not increase too fast, say $d < \frac{1}{3}(1 - \varepsilon)\log n/\log\log n$.

As in our calculations below we are forced to sum estimates $d(n)$ times and $d(n) \to \infty$, we cannot use estimates of the form $O(n^{-K})$, $o(1)$, and so on. This is the reason why the paper is so inconveniently full of concrete constants rather than

Received by the editors May 12, 1980.

1980 *Mathematics Subject Classification.* Primary 05C99; Secondary 60C05.

constants $c_1, c_2, \ldots$. To compensate for this, our estimates tend to be very crude, we always use constants following from generous calculations, so the reader should not be surprised if he can see the inequalities with better constants.

We shall use the notation and terminology of [1]. We shall denote by $\Gamma_k(x)$ the set of vertices at distance k from x:

$$\Gamma_k(x) = \{y \in G: d(x, y) = k\}$$

and write $N_k(x)$ for the set of vertices within distance k:

$$N_k(x) = \bigcup_{i=0}^{k} \Gamma_i(x).$$

Thus diam $G = d$ if $N_d(x) = V(G)$ for every vertex x and $N_{d-1}(y) \neq V(G)$ for some vertex y. As in [3] we write $\mathcal{G}(n, P(\text{edge}) = p)$ for the discrete probability space consisting of the $2^{\binom{n}{2}}$ labelled graphs of order n in which the probability of a fixed graph with m edges is $p^m(1-p)^{\binom{n}{2}-m}$. Equivalently, in $\mathcal{G}(n, P(\text{edge}) = p)$ the edges are chosen independently and with probability p. A related model is $\mathcal{G}(n, m)$ consisting of all graphs with n labelled vertices and m edges, in which any two graphs have the same probability.

Throughout the paper n is assumed to tend to infinity. Thus $f(n) \to \infty$ and $\phi(n) = o(1)$ mean that $f(n) \to \infty$ as $n \to \infty$ and $\phi(n) \to 0$ as $n \to \infty$. Furthermore, we say that almost every (a.e.) graph in $\mathcal{G}(n, P(\text{edge}) = p)$ has property P if the probability that a graph does not have P tends to 0 as $n \to \infty$.

We start with a simple and rather crude lemma which we shall use instead of the de Moivre-Laplace theorem. The strength of the lemma is that the estimates are in terms of concrete functions.

LEMMA 1. *Let S_n have binomial distribution with parameters n and p, that is*

$$P(S_n = k) = b(k; n, p) = \binom{n}{k} p^k q^{n-k},$$

where $0 < p < 1$ and $q = 1 - p$.

(i) *Suppose the integer t satisfies $1 < t < s/2$, where $s = npq$. Then*

$$P(S_n \geqslant pn + t) \leqslant (q + s/t) \frac{e^{1/12n}}{(2\pi s)^{1/2}} \exp\{-t^2/2s + t^3/s^2 + 4t/s\}.$$

(ii) *Suppose $0 < p < \frac{1}{3}$, $0 < \varepsilon < \frac{1}{21}$ and $\varepsilon pn > 40$. Then*

$$P(|S_n - pn| \geqslant \varepsilon pn) < \frac{1}{\varepsilon (pn)^{1/2}} e^{-\varepsilon^2 pn/3}.$$

(iii) *If $u > e$ then*

$$P(S_n \geqslant upn) = P(|S_n - pn| \geqslant (u-1)pn) < \frac{u}{u-1}(e/u)^{upn}.$$

If $v > e^4$ and $vpn > 1$ then

$$P\left(S_n \geqslant \frac{2v}{\log v} pn\right) < e^{-vpn}. \quad \square$$

As the routine proof is similar to the proof of the de Moivre-Laplace theorem in [**12**, Chapter VII, §3], we omit it.

In the next four lemmas we shall suppose that c is a positive constant, $0 < p = p(n) < 1$, $d = d(n)$ is a natural number, $d \geqslant 2$, $p^d n^{d-1} = \log(n^2/c)$ and $pn/\log n \to \infty$ as $n \to \infty$. As we are interested in large values of n, we may and shall assume that $n \geqslant 100, pn > 100 \log n$, $(pn)^{d-2} < n/10$ and $p(pn)^{d-2} < 1/10$. Note that

$$p = n^{1/d-1}(\log n^2/c)^{1/d}$$

and

$$d = (\log n + \log\log n + \log 2 + O(1/\log n))/\log(pn),$$

so the maximum of d is $(1 + o(1))\log n/\log\log n$. Clearly

$$(pn)^{d-1} = \frac{\log(n^2/c)}{pn} n = o(n)$$

and $p(pn)^{d-2} = o(1)$.

LEMMA 2. *Let x be a fixed vertex, let $1 \leqslant k \leqslant d - 1$, and suppose K satisfies*

$$9 \leqslant K < (pn/\log n)^{1/2}/21.$$

Denote by $\Omega_k \subset \mathcal{G}(P(edge) = p)$ the set of graphs for which $a = |\Gamma_{k-1}(x)|$ and $b = |N_{k-1}(x)|$ satisfy

$$\frac{1}{2}(pn)^{k-1} \leqslant a \leqslant \frac{3}{2}(pn)^{k-1}$$

and

$$b \leqslant 2(pn)^{k-1}.$$

Set

$$\alpha_k = K\left(\log n/ (pn)^k\right)^{1/2},$$

$$\beta_k = p(pn)^{k-1}, \qquad \gamma_k = \frac{2(pn)^{k-1}}{n}.$$

Then

$$P(|\,|\Gamma_k(x)| - apn| \geqslant (\alpha_k + \beta_k + \gamma_k)apn|\Omega_k) \leqslant n^{-K^2/9}.$$

PROOF. In order to determine the sets $\Gamma_{k-1}(x)$ and $N_{k-1}(x)$, we have to test which vertices are adjacent to x, then which vertices are adjacent to $\Gamma_1(x)$, and so on, up to $\Gamma_{k-2}(x)$. At each stage we have to test pairs of vertices, at least one of which belongs to $N_{k-2}(x)$. Hence the probability of a given vertex $y \notin N_{k-1}(x)$ being joined to some vertices in $\Gamma_{k-1}(x)$, conditional on Ω_k, is exactly $p_a = 1 - (1-p)^a$. Clearly

$$pa\left(1 - \frac{pa}{2}\right) \leqslant p_a \leqslant pa.$$

Conditional on Ω_k, the random variable $|\Gamma_k(x)|$ has binomial distribution with parameters $n_k = n - b$ and p_a. Since $4n/5 < n_k \leqslant n$, $ap(n - n_k) \leqslant \gamma_k apn$ and

$(ap - p_a)n_k \leqslant \beta_k p_a n_k$, by Lemma 1(ii) we have

$$\begin{aligned} P(\big|\, |\Gamma_k(x)| - apn\big| \geqslant (\alpha_k + \beta_k + \gamma_k)apn|\Omega_k) &\leqslant P(\big|\, |\Gamma_k(x)| - apn_k\big| \geqslant (\alpha_k + \beta_k)apn_k|\Omega_k) \\ &\leqslant P(\big|\, |\Gamma_k(x)| - p_a n_k\big| \geqslant \alpha_k p_a n_k|\Omega_k) \\ &\leqslant \frac{1}{\alpha_k (p_a n_k)^{1/2}} \exp\{-\alpha_k^2 p_a n_k/3\} < \exp\{-\alpha_k^2 p_a n_k/3\} \\ &\leqslant \exp\{-\alpha_k^2 (pn)^k/9\} = n^{-K^2/9}. \end{aligned}$$

Lemma 1(ii) could be applied since

$$0 < p_a \leqslant pa \leqslant \tfrac{3}{2}p(pn)^{k-1} \leqslant \tfrac{3}{2}p(pn)^{d-2} < \tfrac{1}{3},$$

$$0 < \alpha_k = K\big(\log n/\,(pn)^k\big)^{1/2} \leqslant K(\log n/\,(pn))^{1/2} < \tfrac{1}{21},$$

and

$$\alpha_k p_a n_k \geqslant K\big(\log n/\,(pn)^k\big)^{1/2} \frac{3}{10}(pn)^k > 3K \log n > 40. \quad \square$$

LEMMA 3. *Let $K \geqslant 11$ be a constant and define α_k, β_k, γ_k as in Lemma 2. Set*

$$\delta_k = \exp\left(2 \sum_{l=1}^{k} (\alpha_l + \beta_l + \gamma_l)\right) - 1.$$

If n is sufficiently large then with probability at least $1 - n^{-K-2}$ for every vertex x and every natural number k, $1 \leqslant k \leqslant d - 1$, we have

$$\big|\, |\Gamma_k(x)| - (pn)^k\big| < \delta_k (pn)^k.$$

PROOF. The conditions imply that $\delta_{d-1} \to 0$ as $n \to \infty$. In particular, we may assume that $\delta_{d-1} < \frac{1}{4}$. Furthermore, if n is sufficiently large, the conditions of Lemma 2 are satisfied for every k, $1 \leqslant k \leqslant d - 1$. We assume that this is the case.

Let x be fixed and denote by Ω_k^* the set of graphs for which

$$\big|\, |\Gamma_l(x)| - (pn)^l\big| \leqslant \delta_l (pn)^l, \qquad 0 \leqslant l < k.$$

Clearly $\Omega_k^* \subset \Omega_{k-1}^* \subset \Omega_k$.

We shall prove by induction that

$$1 - P(\Omega_k^*) \leqslant 3kn^{-K^2/a}$$

for every k, $0 \leqslant k \leqslant d - 1$. This does hold for $k = 0$. Assume that $1 \leqslant k < d - 1$ and the inequality holds for smaller values of k. Then

$$1 - P(\Omega_k^*) = 1 - P(\Omega_{k-1}^*) + P(\Omega_{k-1}^*)P\big(\big|\, |\Gamma_k(x)| - (pn)^k\big| > \delta_k (pn)^k|\Omega_{k-1}^*\big).$$

Now if $G \in \Omega^*_{k-1}$ then $a = |\Gamma_{k-1}(x)|$ satisfies $|(pn)^k - apn| < \delta_{k-1}(pn)^k$. Therefore

$$\begin{aligned} P\big(\big|\, |\Gamma_k(x)| - (pn)^k\big| \geq \delta_k(pn)^k|\Omega^*_{k-1}\big) &\leq P(\Omega^*_{k-1})^{-1}P\big(\big|\, |\Gamma_k(x)| - apn\big| \geq (\delta_k - \delta_{k-1})(pn)^k|\Omega_k\big) \\ &\leq P(\Omega^*_{k-1})^{-1}P\big(\big|\, |\Gamma_k(x)| - apn\big| \geq 2(\alpha_k + \beta_k + \gamma_k)(pn)^k|\Omega_k\big) \\ &\leq P(\Omega^*_{k-1})^{-1}P\big(\big|\, |\Gamma_k(x)| - apn\big| \geq (\alpha_k + \beta_k + \gamma_k)apn|\Omega_k\big) \\ &\leq \big(1 - 3(k-1)n^{-K^2/9}\big)^{-1}n^{-K^2/9} < 2n^{-K^2/9}. \end{aligned}$$

The next to last inequality holds because of Lemma 2, and the last inequality holds since $6dn^{-K^2/9} < 1$. Consequently

$$1 - P(\Omega^*_k) \leq 3kn^{-K^2/9},$$

as required. Lemma 3 is an immediate consequence of this inequality. □

Before stating the next lemma we introduce some more notation. Given distinct vertices x and y, and a natural number k, define

$$\Gamma^*_k(x,y) = \{z \in \Gamma_k(x) \cap \Gamma_k(y) \colon \Gamma(z) \cap (\Gamma_{k-1}(x) - \Gamma_{k-1}(y)) \neq \varnothing \text{ and } \Gamma(z) \cap (\Gamma_{k-1}(y) - \Gamma_{k-1}(x)) \neq \varnothing\}.$$

Denote by Δ_k the event that $|\Gamma_{k-1}(x)| < 2(pn)^{k-1}$ and $|\Gamma_{k-1}(y)| < 2(pn)^{k-1}$. In our next lemma we shall give a bound on the probability of $\Gamma^*_k(x,y)$ being rather large, conditional on Δ_k. Pick a constant $K > e^7$. For $1 < k < d/2$ define $c_k = c_k(n,p,K)$ by

$$c_k 4p^{2k}n^{2k-1} = (K+4)\log n,$$

and put

$$m_k = m_k(n,p,K) = \frac{2(K+4)\log n}{\log c_k}.$$

Finally, for $d/2 < k \leq d$ put $m_k = m_k(n,p) = 2p^{2k}n^{2k-1}$.

LEMMA 4. *If n is sufficiently large then for every k, $1 < k < d-1$, we have*

$$P(|\Gamma^*_k(x,y)| \geq m_k|\Delta_k) \leq n^{-K-4}.$$

PROOF. In order to determine $\Gamma_{k-1}(x)$ and $\Gamma_{k-1}(y)$, we have to test which vertices are adjacent to x and y, then which vertices are adjacent to $\Gamma_1(x) \cup \Gamma_1(y)$, and so on, which vertices are adjacent to $\Gamma_{k-2}(x) \cup \Gamma_{k-2}(y)$. Thus we have to test the pairs of vertices at least one of which belongs to $N_{k-1}(x) \cup N_{k-1}(y)$. The choice of these edges determines whether or not our final graph belongs to Δ_k. Suppose it does. The probability of a vertex $z \notin N_{k-1}(x)$ being joined to some vertex in $\Gamma_{k-1}(x) - \Gamma_{k-1}(y)$ is

$$1 - (1-p)^b \leq bp \leq 2p^kn^{k-1}, \quad \text{where } b = |\Gamma_{k-1}(x) - \Gamma_{k-1}(y)|.$$

The probability of z being joined to some vertex in $\Gamma_{k-1}(y) - \Gamma_{k-1}(z)$ is also at most $2p^k n^{k-1}$. Since $\Gamma_{k-1}(x) - \Gamma_{k-1}(y)$ and $\Gamma_{k-1}(y) - \Gamma_{k-1}(x)$ are disjoint, the probability that z belongs to $\Gamma_k^*(x, y)$ is at most $(2p^k n^{k-1})^2$. Hence, conditional on the choice of the edges joining vertices in $N_{k-1}(x) \cup N_{k-1}(y)$, with $|\Gamma_{k-1}(x)| \leqslant 2(pn)^{k-1}$ and $|\Gamma_{k-1}(y)| \leqslant 2(pn)^{k-1}$, the probability of $|\Gamma_k^*(x, y)| \geqslant m_k$ is at most $P(S_n^* \geqslant m_k)$, where S_n^* has binomial distribution with parameters n and $p_k^* = 4p^{2k} n^{2k-2}$. Consequently

$$P(|\Gamma_k^*(x, y)| \geqslant m_k | \Delta_k) \leqslant P(S_n^* \geqslant m_k).$$

Now if $n \geqslant 3$ is sufficiently large,

$$p_1^* n \leqslant p_2^* n \leqslant \cdots \leqslant p^*_{\lfloor d/2 \rfloor} n \leqslant 4p^d n^{d-1} = 4 \log(n^2/c) < e^{-4} K \log n,$$

so $c_1 \geqslant c_2 \geqslant \cdots \geqslant c_{\lfloor d/2 \rfloor} > e^4$. Consequently Lemma 1(iii) can be applied with $v = c_k$, so for every k, $1 \leqslant k \leqslant d/2$, we have

$$P(|\Gamma_k^*(x, y)| \geqslant m_k | \Delta_k) \leqslant e^{-(K+4) \log n} = n^{-K}.$$

Furthermore, if n is sufficiently large, we have

$$p_1^* \leqslant p_2^* \leqslant \cdots \leqslant p_{d-1}^* = 4p^{2d-2} n^{2d-4} = 4(\log(n^2/c))^2 (pn)^{-2} < \tfrac{1}{3}$$

and

$$p^*_{\lceil (d+1)/2 \rceil} n \geqslant 4p^{d+1} n^d = 4(pn) \log(n^2/c) > 10^{10} (\log n)^2.$$

Therefore by applying Lemma 1(ii) with $\varepsilon = \frac{1}{25}$ we see that if n is sufficiently large then for every k, $d/2 < k \leqslant d - 1$, we have

$$P(|\Gamma_k^*(x, y)| \geqslant m_k | \Delta_k) \leqslant e^{-(\log n)^2} = n^{-\log n} < n^{-K-4},$$

completing the proof of the lemma. □

LEMMA 5. *Let $K > e^7$ be an arbitrary constant. Then if n is sufficiently large, with probability at least $1 - n^{-K}$ the following assertions hold.*

(i) *For every vertex x*

$$|N_{d-2}(x)| < 2(pn)^{d-2} \quad \textit{and} \quad \big| |\Gamma_{d-1}(x)| - (pn)^{d-1} \big| \leqslant \delta_{d-1} (pn)^{d-1},$$

where δ_{d-1} has the value defined in Lemma 3.

(ii) *For every two vertices x and y*

$$|N_{d-1}(x) \cap N_{d-1}(y)| \leqslant 8p^{2d-2} n^{2d-3}$$

and

$$|\Gamma(N_{d-1}(x) \cap N_{d-1}(y))| \leqslant 16p^{2d-1} n^{2d-2}.$$

PROOF. Since $\delta_1 \leqslant \delta_2 \leqslant \cdots \leqslant \delta_{d-1} \to 0$ and $\sum_{i=0}^{d-2} (pn)^i < \frac{1}{2}(pn)^{d-1}$ if n is sufficiently large, Lemma 3 implies that assertion (i) holds with probability at least $1 - n^{-K-2}$.

In what follows we shall assume that n is sufficiently large. Lemma 3 implies that (if n is sufficiently large then) $P(\Delta_k) \geqslant 1 - n^{-K-2}$ for every k, $1 \leqslant k \leqslant d - 1$, so with probability at least $1 - n^{-K-1}$, Lemma 4 gives that every pair of vertices x, y satisfies

$$|\Gamma_k^*(x, y)| \leq m_k. \tag{1}$$

Note that

$$N_{d-1}(x) \cap N_{d-1}(y) \subset N_{d-2}(x) \cup N_{d-2}(y) \subset (\Gamma_{d-1}(x) \cap \Gamma_{d-1}(y)) \tag{2}$$

and

$$\Gamma_{d-1}(x) \cap \Gamma_{d-1}(y) \subset \bigcup_{k=1}^{d-1} \Gamma_{d-1-k}(\Gamma_k^*(x, y)). \tag{3}$$

From Lemma 3 and inequality (3) we find that with probability at least $1 - 2n^{-K-1}$ for every pair of vertices x, y we have

$$|N_{d-2}(x) \cup N_{d-2}(y)| \leq 4(pn)^{d-2}, \tag{4}$$

$$|\Gamma_{d-1}(x) \cap \Gamma_{d-1}(y)| \leq \sum_{k=1}^{d-1} m_k 2(pn)^{d-1-k} < 7p^{2d-2}n^{2d-3} \tag{5}$$

and

$$|\Gamma(N_{d-1}(x) \cap N_{d-1}(y))| \leq 2pn|N_{d-1}(x) \cap N_{d-1}(y)|. \tag{6}$$

To justify (5) note that

$$2\sum_{k=1}^{\lfloor d/2 \rfloor} m_k(pn)^{d-1-k} < 3m_1(pn)^{d-2} < p^{2d-2}n^{2d-3},$$

for

$$p^{2d-2}n^{2d-3}/\left(m_1(pn)^{d-2}\right) = p^d n^{d-1}\frac{\log c_1}{2(K+4)\log n} > \frac{\log c_1}{2K}$$

and $c_1 \to \infty$ as $n \to \infty$. Furthermore,

$$2\sum_{k=\lceil (d+1)/2 \rceil}^{d-1} m_k(pn)^{d-1-k} < 3m_{d-1} = 6p^{2d-2}n^{2d-3},$$

so (5) does hold. Relations (2)–(6) imply

$$|N_{d-1}(x) \cup N_{d-2}(x)| \leq 4(pn)^{d-2} + 7p^{2d-2}n^{2d-3} < 8p^{2d-2}n^{2d-3}$$

and

$$|\Gamma(N_{d-1}(x) \cap N_{d-1}(y))| < 16p^{2d-1}n^{2d-2}.$$

Consequently assertions (i) and (ii) of the lemma hold with probability at least $1 - 4n^{-K-1} > 1 - n^{-K}$. □

Armed with these lemmas, we are ready to prove the main result of the paper.

THEOREM 6. *Let c be a positive constant, $d = d(n) \geq 2$ a natural number, and define $p = p(n, c, d)$, $0 < p < 1$, by*

$$p^d n^{d-1} = \log(n^2/c).$$

Suppose that $(pn)/(\log n)^3 \to \infty$. Then in $\mathcal{G}(P(edge) = p)$ we have

$$\lim_{n\to\infty} P(\operatorname{diam} G = d) = e^{-c/2} \quad \text{and} \quad \lim_{n\to\infty} P(\operatorname{diam} G = d+1) = 1 - e^{-c/2}.$$

PROOF. If for some vertices $x, y \in G$ we have $y \notin N_d(x)$ then we say that x is *remote* from y and (x, y) is a *remote pair*. Let $X = X(G)$ be the number of remote pairs of G and write $X_r = X_r(G) = \binom{X}{r}$ for the number of unordered r-tuples of remote pairs. Our aim is to show that *the distribution of X tends to the Poisson distribution with parameter $c/2$*, so $P(X = k) \sim e^{-c/2}(c/2)^k/k!$. We shall do this by estimating $E(X_r)$ for every $r \geq 1$. Since r disjoint pairs of vertices contain 2^r r-sets of vertices meeting each pair, it is easily seen that

$$E(X_r) = \binom{n}{r} 2^{-r} F_r (1 + o(1)),$$

where F_r is the probability that a fixed r-tuple $\tau = (x_1, \ldots, x_r)$ of vertices consists of vertices remote from some other vertices. Write

$$A_i = \Gamma_{d-1}(x_i) - \bigcup_{j \neq i} N_{d-1}(x_j),$$

$$T = \bigcap_{i \neq j} \left(N_{d-1}(x_i) \cap N_{d-1}(x_j)\right),$$

$$S = V(G) - \bigcup_{i=1} N_{d-1}(x_i),$$

$$S' = S - \Gamma(T),$$

$$a_i = |A_i|, \quad s = |S|, \quad s' = |S'| \quad \text{and} \quad t = |T|.$$

Pick a constant $K > \max\{r + 1, e^7\}$. Then by Lemma 5 with probability at least $1 - n^{-K}$ we have

$$\left|a_i - (pn)^{d-1}\right| \leq \delta_{d-1}(pn)^{d-1} + 8rp^{2d-2}n^{2d-3}$$

$$= (pn)^{d-1}\left\{\delta_{d-1} + 8r\left(\log(n^2/c)\right)/(pn)\right\} = \delta(pn)^{d-1}, \tag{7}$$

$$n \geq s \geq s' \geq n - 8r^2p^{2d-1}n^{2d-2} = (1 - \varepsilon)n. \tag{8}$$

We claim that

$$\delta \log n \to 0 \quad \text{and} \quad \varepsilon \to 0. \tag{9}$$

Indeed, the first relation holds since if n is large,

$$\delta_{d-1} \leq 3 \sum_{l=1}^{k} (\alpha_l + \beta_l + \gamma_l) \leq 4(\alpha_1 + \beta_{d-1} + \gamma_{d-1})$$

$$= 4\left\{K\left(\frac{\log n}{pn}\right)^{1/2} + p^{d-1}n^{d-2} + 2p^{d-2}n^{d-3}\right\}$$

$$\leq 4\left\{K\left(\frac{\log n}{pn}\right)^{1/2} + \frac{3 \log n}{pn} + \frac{6 \log n}{(pn)^2}\right\}$$

and $(pn)/(\log n)^3 \to \infty$. Furthermore, $\varepsilon \to 0$ since

$$p^{2d-1}n^{2d-3} = (p^d n^{d-1})^2/(pn) < (3 \log n)^2/(pn) \to 0.$$

Denote by $P'(\cdot)$ the probability conditional on a particular choice of the sets A_i, S and S', satisfying (7) and (8). In order to estimate F_r we shall estimate the conditional probability $Q_r = P'$ (τ consists of remote vertices). Put

$$R_r = P'(\exists\, y_i \in S \text{ not joined to } A_i,\ i = 1, \ldots, r)$$

and

$$R_r' = P'(\exists\, y_i \in S' \text{ not joined to } A_i,\ i = 1, \ldots, r).$$

Then clearly $R_r' \leqslant Q_r \leqslant R_r$. Furthermore,

$$R_r = \prod_{i=1}^{r} \left\{1 - (1 - (1 - p)^{a_i})^s\right\},$$

and R_r' is given by an analogous expression.

In order to estimate R_r from above, note that

$$(1 - p)^{a_i} \leqslant e^{-pa_i} \leqslant e^{-p^d n^{d-1}(1-\delta)} = \frac{c}{n^2}(1 + o(1)),$$

$$(1 - (1 - p)^{a_i})^s \geqslant 1 - \frac{sc}{n^2}(1 + o(1)) = 1 - \frac{c}{n} + o(1/n),$$

so $R_r \leqslant (c/n)^r(1 + o(1))$. Similarly R_r' can be estimated from below as follows:

$$(1 - p)^{a_i} \geqslant e^{-pa_i(1+p)} \geqslant e^{-p^d n^{d-1}(1+p)(1+\delta)} = \frac{c}{n^2}(1 + o(1)),$$

$$(1 - (1 - p)^{a_i})^{s'} \leqslant 1 - \frac{s'c}{n^2}(1 + o(1)) = 1 - \frac{c}{n} + o(1/n).$$

Consequently $Q_r = (c/n)^r(1 + o(1))$. Since (7) and (8) hold with probability at least $1 - n^{-K}$,

$$(1 - n^{-K})Q_r \leqslant F_r \leqslant (1 - n^{-K})Q_r + n^{-K},$$

so $F_r = (c/n)^r(1 + o(1))$, giving

$$E(X_r) = \frac{n^r}{r!}\left(\frac{c}{n}\right)^r 2^{-r}(1 + o(1)) = \frac{(c/2)^r}{r!}(1 + o(1)).$$

This relation shows that if r is fixed and $n \to \infty$ then the rth moment of X tends to the rth moment of the Poisson distribution with mean $c/2$. Consequently the distribution of X tends to the Poisson distribution with mean $c/2$ (see Chung [**6**, p. 99]), as claimed. In particular, $P(\text{diam } G \leqslant d) = P(X = 0) \sim e^{-c/2}$.

Now it is easy to deduce the assertions of the theorem. If $d = 2$ then clearly

$$P(\text{diam } G \leqslant 1) = P(G = K^n) = p^{\binom{n}{2}} \to 0.$$

Suppose now that $d \geqslant 3$. Given $L > 0$, choose p_1 so that

$$p_1^{d-1} n^{d-2} = \log \frac{n^2}{L}.$$

Then $p < p_1$ and

$$P(\text{diam } G \leqslant d - 1) \leqslant P_1(\text{diam } G \leqslant d - 1) \sim e^{-L/2},$$

where P_1 denotes the probability in the space $\mathcal{G}(P(\text{edge}) = p_1)$. Since L was arbitrary, $P(\text{diam } G \leqslant d - 1) \sim 0$. Hence for every $d \geqslant 2$ we have

$$\lim_{n\to\infty} P(\text{diam } G \leqslant d - 1) = 0.$$

An analogous argument implies $\lim_{n\to\infty} P(\text{diam } G \leqslant d+1) = 1$, completing the proof. $\square$

As an immediate consequence of Theorem 6 we find the range of p for which almost every graph in $\mathcal{G}(P(\text{edge}) = p)$ has diameter d, provided d does not increase too fast with n.

COROLLARY 7. (i) *Suppose $p^2n - 2\log n \to \infty$ and $n^2(1-p) \to \infty$. Then a.e. graph in $\mathcal{G}(P(edge) = p)$ has diameter* 2.

(ii) *Suppose the function $m = m(n) < \binom{n}{2}$ satisfies*

$$m^2/n^3 - \tfrac{1}{2}\log n \to \infty.$$

Then a.e. graph in $\mathcal{G}(n, m)$ has diameter 2.

Both assertions are best possible.

COROLLARY 8. (i) *Suppose the functions $d = d(n) \geqslant 3$ and $0 < p = p(n) < 1$ satisfy*

$$(\log n)/d - 3\log\log n \to \infty,$$

$$p^d n^{d-1} - 2\log n \to \infty \quad \text{and} \quad p^{d-1}n^{d-2} - 2\log n \to -\infty.$$

Then a.e. graph in $\mathcal{G}(P(edge) = p)$ has diameter d.

(ii) *Suppose the functions $d = d(n) \geqslant 3$ and $m = m(n)$ satisfy*

$$(\log n)/d - 3\log\log n \to \infty,$$

$$2^{d-1}m^d n^{-d-1} - \log n \to \infty \quad \text{and} \quad 2^{d-2}m^{d-1}n^{-d} - \log n \to -\infty.$$

Then a.e. graph in $\mathcal{G}(n, m)$ has diameter d.

Both assertions are best possible.

PROOFS. The first condition in Corollary 7(i) ensures that $P(\text{diam } G \leqslant 2) \sim 1$. As

$$P(\text{diam } G \leqslant 1) = P(G = K^n) = p^{\binom{n}{2}},$$

$P(\text{diam } G \geqslant 2) \to 1$ iff $n^2\log(1/p) \to \infty$.

Corollary 8(i) is an immediate consequence of Theorem 6 since if $(\log n)/d - 3\log\log n \to \infty$ and $p_1^d n^{d-1} = \log(n^2/c)$ then we have $(p_1 n)/(\log n)^3 \to \infty$. The property of having diameter d is a convex property, so the second assertions follow from Theorem 8(ii) [**3**, p. 133]. $\square$

We conclude the paper by discussing a question concerning a property closely related to the diameter of a graph. In what range of p is it true that almost every graph $G \in \mathcal{G}(P(\text{edge}) = p)$ has diameter d *and* for every vertex x there is a vertex y at distance d from x.

THEOREM 9. (i) *Suppose $0 < q < 1$, $nq - \log n \to \infty$, and $p = 1 - q$. Then a.e. graph in $\mathcal{G}(P(edge) = p)$ is such that no vertex is joined to every other vertex.*

(ii) *Suppose $d = d(n) \geqslant 2$ and $0 < p = p(n) < 1$ satisfy $(pn)/(\log n) \to \infty$ and*

$$(\log n)(p^d n^{d-1} - \log n + \log\log n) \to -\infty.$$

Then a.e. graph in $\mathcal{G}(P(edge) = p)$ is such that $\Gamma_d(x) \neq V(G)$ holds for every vertex x.

PROOF. (i) The expected number of vertices of degree $n-1$ is $np^{n-1} = n(1-q)^{n-1} \sim ne^{-qn} \to 0$. Consequently the assertion follows from Chebyshev's inequality.

(ii) Suppose $(p_c n)/\log n \to \infty$ and

$$p_c^d n^{d-1} = \log n - \log\log(n/c),$$

where c is a positive constant. Then by Lemma 3 (more precisely, by a trivial variant of it since p_c and d satisfy slightly different conditions) with probability $1 - n^{-2}$ we have

$$a = |\Gamma_{d-1}(x)| \leqslant (1+\delta)(p_c n)^{d-1}$$

and

$$b = |N_{d-1}(x)| \leqslant 2(p_c n)^{d-1}$$

for every vertex x, where $\delta \to 0$. Given $\Gamma_{d-1}(x)$ and $N_{d-1}(x)$, the probability of $N_d(x) = V(G)$ is

$$(1-(1-p)^a)^{n-b}.$$

Consesquently the expected number of vertices x satisfying $N_d(x) = V(G)$ is asymptotic to

$$n\left(1-(1-p_c)^{(p_c n)^{d-1}}\right)^n \sim n(1-e^{-p_c^d n^{d-1}})^n = n\left(1-\frac{\log(n/c)}{n}\right)^n \sim c.$$

Hence by Chebyshev's inequality if n is sufficiently large, the probability that there is a vertex x with $N_d(x) = V(G)$ is at most $2c$, say. Since

$$\log\log(n/c) = \log\log n - \frac{\log c}{\log n}(1+o(1)),$$

our second condition implies that $p < p_c$ if n is sufficiently large. As c can be chosen arbitrarily small, the assertion follows. □

Putting together Theorems 6 and 9 we obtain the following result concerning graphs of diameter d in which every vertex shows that the diameter is at least d.

COROLLARY 10. *Suppose* $d = d(n) \geqslant 2$ *and* $0 < p = p(n) < 1$ *satisfy* $(\log n)/d - 3\log\log n \to \infty$, $p^d n^{d-1} - 2\log n \to \infty$ *and*

$$(\log n)(p^{d-1}n^{d-2} - \log n + \log\log)n \to -\infty.$$

Then in $\mathcal{G}(P(edge) = p)$ *a.e. graph has diameter* d *and no vertex* x *satisfies* $N_{d-1}(x) = V(G)$. □

REFERENCES

1. B. Bollobás, *Extremal graph theory*, Academic Press, London, New York and San Francisco, 1978.

2. ______, *Chromatic number, girth and maximum degree*, Discrete Math. **24** (1978), 311–314.

3. ______, *Graph theory, an introductory course*, Graduate Texts in Math., Vol. 63, Springer-Verlag, Berlin, Heidelberg and New York, 1979.

4. ______, *Degree sequences of random graphs*, Discrete Math. **33** (1981), 1–19.

5. B. Bollobás and P. Erdős, *Cliques in random graphs*, Math. Proc. Cambridge Philos. Soc. **80** (1976), 419–427.

6. K. L. Chung, *A course in probability theory*, 2nd ed., Academic Press, New York and London, 1974.

7. P. Erdös, *Graph theory and probability*, Canad. J. Math. **11** (1959), 34–38.

8. ______, *Graph theory and probability*. II, Canad. J. Math. **13** (1961), 346–352.

9. P. Erdös and A. Rényi, *On random graphs*. I, Publ. Math. Debrecen **6** (1959), 290–291.

10. ______, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci. **5** (1960), 17–61.

11. P. Erdös and R. J. Wilson, *On the chromatic index of almost all graphs*, J. Combin. Theory Ser. B **23** (1977), 255–257.

12. W. Feller, *An introduction to probability theory and its applications*. Vol. I, 3rd ed., Wiley, New York and London, 1968.

13. G. R. Grimmett and C. J. H. McDiarmid, *On colouring random graphs*, Math. Proc. Cambridge Philos. Soc. **77** (1975), 313–324.

14. V. Klee and D. Larman, *Diameters of random graphs* (to appear).

15. A. D. Korshunov, *On the diameter of graphs*, Soviet Math. Dokl. **12** (1971), 302–305.

16. ______, *Solution of a problem of Erdös and Rényi on Hamiltonian cycles in nonoriented graphs*, Soviet Math. Dokl. **17** (1976), 760–764.

17. D. W. Matula, *On the complete subgraphs of a random graph*, Combinatory Math. and its Applications, Chapel Hill, N. C., 1970, pp. 356–369.

18. J. Moon and L. Moser, *Almost all* (0, 1) *matrices are primitive*, Studia Sci. Math. Hungar. **1** (1966), 153–156.

19. L. Pósa, *Hamiltonian cycles in random graphs*, Discrete Math. **14** (1976), 359–364.

Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge, CB2 1SB, England