Can. J. Math., Vol. XXXIII, No. 3, 1981, pp. 618–640

DIAMETERS OF RANDOM GRAPHS

VICTOR KLEE AND DAVID LARMAN

1. Introduction. For two nodes x and y of a graph G, the distance $\delta_G(x, y)$ is the smallest integer k such that k edges form a path from x to y; $\delta_G(x, x) = 0$, and $\delta_G(x, y) = \infty$ when $x \neq y$ and there is no path from x to y. The *diameter* δ_G is the maximum of $\delta_G(x, y)$ as x and y range over the nodes of G. When G is connected, $\delta(G)$ is the smallest integer k such that any two nodes of G can be joined by a path formed from at most k edges. When G is not connected, $\delta(G) = \infty$ and there is interest in $\delta_c(G)$, the maximum of $\delta(C)$ over the components C of G.

For $2 \leqq n < \infty$ and $0 \leqq E \leqq n(n-1)/2$, let $\mathscr{G}(n, E)$ denote the set of all loopless undirected graphs with the node-set $\{1, \ldots, n\}$ and exactly E edges. Each edge is an unordered pair of distinct nodes, and hence

$$|\mathscr{G}(n, E)| = \binom{n(n-1)/2}{E}.$$

For $1 \leqq d \leqq \infty$, let $\mathscr{G}(n, E, <d)$ [resp. $\mathscr{G}(n, E, d), G(n, E, > d)$] denote the set of all $G \in \mathscr{G}(n, E)$ such that $\delta(G) < d$ [resp. $= d$, $> d$]. Let

$$P(n, E, < d) = |\mathscr{G}(n, E, <d)|/|\mathscr{G}(n, E)|,$$

the probability that a random labelled graph with n nodes and E edges is of diameter $<d$. The numbers $P(n, E, d)$ and $P(n, E, > d)$ are similarly defined.

Our main results are as follows:

THEOREM 1. *If the positive integers $d \geqq 2$, $E(1)$, $E(2), \ldots$ are such that*

$$E(n)^{d-1}/n^d \to 0 \text{ as } n \to \infty$$

then $P(n, E(n), <d) \to 0$.

THEOREM 2. *If the positive integers $d \geqq 2$, $E(1)$, $E(2), \ldots$ are such that*

$$(E(n)^d/n^{d+1}) - \log n \to \infty \text{ as } n \to \infty,$$

then $P(n, E(n), >d) \to 0$.

COROLLARY. *If the positive integers $d \geqq 2$, $E(1)$, $E(2), \ldots$ are such that*

$$E(n)^{d-1}/n^d \to 0 \text{ and } (E(n)^d/n^{d+1}) - \log n \to \infty \text{ as } n \to \infty,$$

then $P(n, E(n), d) \to 1$.

Received October 26, 1979. This research was supported in part by the U.S. Office of Naval Research. The authors are indebted to David Anacker for a helpful comment.

Under the conditions of the Corollary, the diameter function δ may range widely on $\mathscr{G}(n, E(n))$, for

$$\min\{\delta(G):G \in \mathscr{G}(n, E)\} \leqq 2 \text{ if and only if } E \geqq n - 1$$

and

$$\max\{\delta(G):G \in \mathscr{G}(n, E)\} = \infty \quad \text{if and only if } E \leqq (n-1)(n-2)/2.$$

However, the Corollary asserts that for large n, "almost all" members of $\mathscr{G}(n, E(n))$ are of diameter d. In addition to its intrinsic interest, this result may be useful in the analysis of graph-theoretic algorithms based on breadth-first search [**1**].

The study of random graphs was initiated in a series of papers by Erdös and Rényi (see [**2**]), who obtained some beautiful and striking results. In our notation, their connectedness theorem [**3**] asserts that if λ is a real constant and

$$E_\lambda(n) = \lfloor \tfrac{1}{2} n \log n + \lambda n \rfloor$$

then

$$P(n, E_\lambda(n), < \infty) \rightarrow \exp(-e^{-2\lambda}) \text{ as } n \rightarrow \infty.$$

This provides a background for Theorem 2, since when $d < \infty$ it is clear that if $P(n, E(n), > d) \rightarrow 0$ then $E(n)$ grows more rapidly than any $E_\lambda(n)$.

In existing work on random graphs, diameters seem to have been neglected aside from [**3**], results of Moon and Moser [**5**] for $\delta = 2$ (see Section 4 below), and results of Koršunov [**4**] on the behavior of δ_C for very sparse graphs (see Section 12). Our Theorems 1 and 2 fill most of the gap between [**3**] and [**5**].

Note added in proof. After the present paper was accepted for publication, we learned of a paper by B. Bollobás that also studies diameters of random graphs. We have not seen his paper, but apparently its methods are quite different from ours and its results are somewhat sharper. It will appear in the Transactions of the American Mathematical Society.

2. Adjacency matrices. For purposes of counting, graphs are represented by their adjacency matrices. When G is a graph with node-set $\{1, \ldots, n\}$ the adjacency matrix $A_G = (a_{ij})$ is an $n \times n$ symmetric matrix of 0's and 1's, with $a_{ij} = 1$ if and only if $\{i, j\}$ is an edge of G. Since G is assumed to be loopless, the main diagonal of A_G is 0. It would be equally reasonable to assume there is a loop at each node, or to make no assumption about loops. The details of the counting would be essentially the same and our theorems would remain valid because the diameter of a graph is not affected by the addition or removal of loops.

Let $B_G = A_G + I$, the adjacency matrix of the graph obtained by adding a loop at each node of G. The following three conditions are equivalent for each positive integer k:

(a) $\delta(G) \leqq k$;

(b) for $1 \leqq i < j \leqq n$, at least one of the matrices $A_G, A_G^2, \ldots, A_G^k$ has a positive entry in position (i, j);

(c) all entries of B_G^k are positive.

In studying the diameters of random graphs, the cases of diameter $\delta \geqq 3$ are much harder to handle than the case $\delta = 2$ because they involve higher powers of the adjacency matrix.

The set $\mathscr{A}(n, E) = \{A_G : G \in \mathscr{G}(n, E)\}$ consists of all $n \times n$ symmetric matrices that have 0's on the main diagonal, $n^2 - n - 2E$ additional 0's, and 2E 1's. Because of symmetry, members of $\mathscr{A}(n, E)$ are determined by their restrictions to the set $\Omega(n)$ of positions covered by the minor diagonals that are parallel to the main diagonal and start in the following positions:

$$(1, 2, (1, 4), \ldots, (1, n-1) \text{ and } (3, 1), (5, 1), \ldots, (n, 1) \quad \text{when } n \text{ is odd};$$

$$(1, 2), (1, 4), \ldots, (1, n) \text{ and } (3, 1), (5, 1), \ldots, (n-1, 1) \quad \text{when } n \text{ is even}.$$

In other words, $\Omega(n)$ is the set

$$\{(1 + k, j + k) : 2 \leqq \text{even } j \leqq n - 1, 0 \leqq k \leqq n - j\} \cup \{(i + k, 1 + k) : 3 \leqq \text{odd } i \leqq n, 0 \leqq k \leqq n - i\}.$$

Note that

(d) for $1 \leqq i < j \leqq n$, $\Omega(n)$ includes exactly one of (i, j) and (j, i);

(e) when n is odd, each row and each column of an $n \times n$ matrix includes exactly $(n - 1)/2$ positions in $\Omega(n)$.

These facts are used frequently without explicit reference. Because of (e), counting is simpler in the case of odd n. In our proofs of the main theorems, the arguments for even n are omitted because they are similar to those for odd n but are technically more complicated in ways that are irritating but not interesting. One way of handling even numbers of nodes is to treat simultaneously the set $\mathscr{G}(n, E)$ for odd n and the set $\mathscr{G}(n - 1, E)$ as naturally embedded in $\mathscr{G}(n, E)$. Under this embedding, the members of $\mathscr{G}(n - 1, E)$ correspond to members of $\mathscr{G}(n, E)$ in which the node n is isolated and to members of $\mathscr{A}(n, E)$ in which the last row and last column are 0.

3. Elementary estimates. This section collects some elementary estimates that are used throughout the paper and are henceforth referred to by number. We use the combinatorial inequality

(1) $$\binom{n}{k} \leqq \frac{n^k}{k!} \quad \text{for } 1 \leqq k \leqq n,$$

the analytic inequality

(2) $$1 + x \leqq e^x \quad \text{for all } x,$$

and the fact that for $0 \leqq |x| < y$,

(3) $$\log\left(1 - \frac{x}{y}\right)^{y-x} = (y - x)\left(-\sum_{k=1}^{\infty} \frac{x^k}{ky^k}\right) = -x + \sum_{k=2}^{\infty} \frac{1}{k(k-1)} \frac{x^k}{y^{k-1}}.$$

When the functions α and β are defined for positive integers,

$\alpha \sim \beta$ means $\lim_{n\to\infty} \alpha(n)/\beta(n) = 1$

and $\alpha \prec \beta$ means $\alpha(n) < \beta(n)$ for all sufficiently large n.

We use the Stirling-de Moivre estimate,

(4) $$n! \sim \sqrt{2\pi}\, n^{n+1/2} e^{-n}$$

and the following consequences of (3) and (4): if $0 \leqq s \leqq o(N)$ then

(5) $$\frac{N!}{(N-s)!} \sim N^s \exp\left(-\sum_{k=2}^{\infty} \frac{1}{k(k-1)} \frac{s^k}{N^{k-1}}\right) \quad \text{as } N \to \infty;$$

if $0 \leqq w \leqq o(m)$ then

(6) $$\frac{(m-w)!}{m!} \sim m^{-w} \exp\left(\sum_{k=2}^{\infty} \frac{1}{k(k-1)} \frac{w^k}{m^{k-1}}\right) \quad \text{as } m \to \infty.$$

Since

$$\binom{m-w}{N-s} \Big/ \binom{m}{N} = \frac{(m-w)!}{m!} \frac{N!}{(N-s)!} \frac{(m-N)!}{(m-N-w+s)!},$$

it follows from (5) and (6) that if $0 \leqq w \leqq o(m)$, $0 \leqq s \leqq o(N)$ and $0 \leqq w - s \leqq o(m - N)$ then

(7) $$\binom{m-w}{N-s} \Big/ \binom{m}{N} \sim \left(\frac{N}{m}\right)^s \left(1 - \frac{N}{m}\right)^{w-s} \exp\left(\sum_{k=2}^{\infty} \frac{1}{k(k-1)} \times \left(\frac{w^k}{m^{k-1}} - \frac{s^k}{N^{k-1}} - \frac{(w-s)^k}{(m-N)^{k-1}}\right)\right)$$

as $m \to \infty$, $N \to \infty$, $m - N \to \infty$.

It is also useful to know that:

(8) If $E \leqq F$ then $P(n, E, >d) \geqq P(n, F, >d)$ and $P(n, E, <d) \leqq P(n, F, <d)$.

To justify this, note that the diameter of a graph cannot be increased by adding edges or decreased by removing them. Considering the natural subgraph correspondence between members of $\mathscr{G}(n, E)$ and members of

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press

$\mathscr{G}(n, F)$, we see that if $m = n(n-1)/2$ then

$$|\mathscr{G}(n, F, <d)| \geqq |\mathscr{G}(n, E, <d)| \frac{\binom{m-E}{F-E}}{\binom{F}{E}}$$

$$= |\mathscr{G}(n, E, <)| \frac{|\mathscr{G}(n, F)|}{|\mathscr{G}(n, E)|}.$$

This establishes the second inequality of (8) for all d, and the first inequality is a consequence.

4. The case $d = 2$. It was proved by Moon and Moser [**5**] that when E is unrestricted the usual value of δ is 2. More precisely, if $\mathscr{G}(n)$ is the set of all labelled graphs with n nodes and $\mathscr{T}(n)$ consists of those of diameter 2, then $|\mathscr{T}(n)|/|\mathscr{G}(n)| \to 1$ as $n \to \infty$. This is related to the fact that the average number of edges in a member of $\mathscr{G}(n)$ is $n(n-1)/4$, which is close to the maximum $n(n-1)/2$. However, there is interest in the diameters of sparse graphs as well as of dense ones, especially since graphs associated with practical problems are apt to be sparse. That is the motivation for our results.

When $d = 2$, Theorem 1 is obvious and Theorem 2 is due to [**5**] and, independently, to Erdös and Rényi (unpublished). Since the proof in [**5**] is merely sketched, our first lemma treats the case $d = 2$ in detail, aside from ignoring even values of n.

LEMMA 1. *If* $(E^2/n^3) - \log n \to \infty$ *as* $n \to \infty$ *then* $P(n, E, >2) \to 0$.

Proof. For each n, let $r(n) = (n-1)/2$ and $m(n) = nr(n)$. Let $\mathscr{G}_2(n, E)$ denote the set of all $G \in \mathscr{G}(n, E)$ such that each pair of distinct nodes of G is joined by a path of length 2. Lemma 1 is proved by showing, under the stated assumption about the growth of $E(n)$, that

$$|\mathscr{G}_2(n, E)|/|\mathscr{G}(n, E)| \to 1 \text{ as } n \to \infty.$$

It suffices to do this for each subsequence of the n's such that the corresponding subsequence of $m - E$ is constant or tends to ∞ as $n \to \infty$. The first case is obvious and hence we may assume

(9) $\quad m - E \to \infty \quad \text{as} \quad n \to \infty.$

For each $G \in \mathscr{G}(n, E)$ let

$$Z(G) = \{(i, j) \in \Omega(n): \text{the } (i, j) \text{ entry of } A_G^2 \text{ is } 0\}.$$

Since $G \in \mathscr{G}_2(n, E)$ if and only if $|Z(G)| = 0$, it suffices to show $\zeta(n, E) \to 0$ where $\zeta(n, E)$ is the expected value of $|Z(G)|$ for $G \in \mathscr{G}(n, E)$.

For fixed $(s, t) \in \Omega(n)$, let us count the number of $A \in \mathscr{A}(n, E)$ such that the (s, t) entry of A^2 is 0. Suppose that A has precisely k 1's in the

sth row of $\Omega(n)$; say in positions $(s, j_1), \ldots, (s, j_k)$. If $a_{st} = 0$ then the k entries $a_{j_1 t}, \ldots, a_{j_k t}$ of A are 0 and there remain $m - r - k$ places in $\Omega(n)$ to be filled with $E - k$ 1's and the rest 0's. If $a_{st} = 1$ and (say) $j_k = s$ then the $k - 1$ entries $a_{j_1 t}, a_{j_{k-1} t}$ of A are 0 and there remain $m - r - k + 1$ places in $\Omega(n)$ to be filled with $E - k$ 1's and the rest 0's. It follows that the number of A's in question is at most

$$\sum_{k=0}^{r-1} \binom{r-1}{k} \binom{m-r-k}{E-k} + \sum_{k=1}^{r-1} \binom{r-1}{k-1} \binom{m-r-k+1}{E-k} \leqq \sum_{k=0}^{r-1} \binom{r-1}{k} \binom{m-r-k+1}{E-k},$$

whence

$$\binom{m}{E} \zeta(n, E) \leqq m \sum_{k=0}^{r-1} \binom{r-1}{k} \binom{m-r-k+1}{E-k}.$$

Because of (9), (7) is applicable and it follows that $\zeta(n, E)$ is at most

$$m\left(1 - \frac{E}{m}\right) \sum_{k=0}^{r-1} \binom{r-1}{k} \left(\frac{E}{m}\right)^k = m\left(1 - \frac{E}{m}\right)^{r-1} \left(1 + \frac{E}{m}\right)^{r-1} = m\left(1 - \frac{E^2}{m^2}\right)^{r-1},$$

whence

$$\zeta(n, E) \leqq m e^{-E^2(r-1)/m^2}$$

by (2). But then

$$\log \zeta(n, E) \leqq (\log m) \frac{E^2}{m^2}(r - 1) = \log \frac{n(n-1)}{2} - \frac{E^2}{n^2(n-1)^2/4}\left(\frac{1}{2}n - \frac{3}{2}\right) \sim 2\left(\log n - \frac{N^2}{n^3}\right),$$

whence $\log \zeta(n, E) \to -\infty$ by hypothesis and consequently $\zeta(n, E) \to 0$ as $n \to \infty$.

5. Standing hypotheses. In order to avoid repetition, we state here the standing hypotheses that are used in Sections 6-11.

d is an integer $\geqq 3$;
n is an odd integer $\geqq 3$;
$r = (n - 1)/2$; $m = n(n - 1)/2$;
γ is a positive real constant;
E is a nonnegative integer-valued function defined for all n, with $E(n) \leqq m$;
$f(n) = E(N)/n$; $\mu(n) = (2(\gamma + 7) f(n) \log n)^{1/2}$;
$E(n)$, $f(n)$ and $\mu(n)$ are often written as E, f and μ.

The reason for assuming n odd was discussed in Section 2.

6. The essential inequalities. The inequalities of this section are essential in all that follows. Here $\lfloor x \rfloor$ and $\lceil x \rceil$ are respectively the greatest integer $\leqq x$ and the least integer $\geqq x$.

LEMMA 2. *If*

(10) $\lim_{n\to\infty}(E(n)^{d-1}/n^d) - \log n < \infty$

and

(11) $\lim_{n\to\infty}(E(n)/n \log n) = \infty$

then

$$\sum_{k=0}^{\lceil f-\mu-1\rceil} \frac{1}{k!} f^k \left(1 + \frac{3f^{d-2}}{n}\right)^k \prec \frac{1}{8} n^{-(\gamma+1)} e^f$$

and

$$\sum_{k=\lfloor f+\mu+1\rfloor}^{\infty} \frac{1}{k!} f^k \left(1 + \frac{3f^{d-2}}{n}\right)^k \prec \frac{1}{8} n^{-(\gamma+1)} e^f.$$

Proof. It follows from (10) that $\lim_{n\to\infty}(3f^{d-1}/n) - 3 \log n < \infty$, whence

(12) $3f^{d-1}/n \prec 4 \log n$

and

(13) $$\frac{f^{d-2}}{n} = \left(\frac{f^{d-1/n}}{n^{1/(d-2)}}\right)^{(d-2)/(d-1)} \prec \left(\frac{(4 \log n)/3}{n^{1/(d-2)}}\right)^{(d-2)/(d-1)} \to 0.$$

And (11) implies

(14) $$\frac{f}{\mu} = (2\gamma + 14)^{-1/2}\left(\frac{f}{\log n}\right)^{1/2} = (2\gamma + 14)^{-1/2}\left(\frac{E}{n \log n}\right)^{1/2} \to \infty.$$

For the first conclusion of Lemma 2, verify that

$$\begin{aligned}\sum_{k=0}^{\lceil f-\mu-1\rceil} \frac{1}{k!} f^k \left(1 + \frac{3f^{d-2}}{n}\right)^k &\leqq_{(a)} \lceil f-\mu \rceil \frac{1}{\lfloor f-\mu \rfloor !} f^{\lfloor f-\mu \rfloor}\left(1 + \frac{3f^{d-2}}{n}\right)^f \\ &\prec_{(4,2)} \frac{\lceil f-\mu \rceil}{2\lfloor f-\mu \rfloor^{1/2}} \frac{e^{f-\mu}}{\lfloor f-\mu \rfloor^{\lfloor f-\mu \rfloor}} f^{f-\mu} e^{3f^{d-1}/n} \\ &\prec_{(b)} \tfrac{1}{2} f^{1/2}(f-\mu)^{3/2} \frac{1}{(1-\mu/f)^{f-\mu}} e^{f-\mu} n^4 \\ &\prec_{(3)} \tfrac{1}{2} f^2 e^{\mu-\mu^2/2f} e^{f-\mu} n^4 = \left(\frac{f}{n}\right)^2 n^{-(\gamma+1)} e^f \prec_{(12)} \frac{1}{8} n^{-(\gamma+1)} e^f.\end{aligned}$$

To justify (a), note that if j and k are integers with $0 \leqq j \leqq k \leqq f$ then $f^j/j! \leqq f^k/k!$

For (b), apply (14) to the first factor, (12) to the fourth factor, and use the fact that

$$(f-\mu)^{f-\mu-3/2} \prec \lfloor f-\mu \rfloor^{[f-\mu]}.$$

For the second conclusion of Lemma 2, note that

$$\frac{f}{f+\mu}\left(1+\frac{3f^{d-2}}{n}\right) \prec_{(12)} \frac{f+4\log n}{f+\mu} \prec_{(14)} 1,$$

whence

$$(15)\quad \sum_{k=0}^{\infty}\left(\frac{f}{f+\mu}\right)^{k}\left(1+\frac{3f^{d-2}}{n}\right)^{k} \prec \frac{f+\mu}{\mu-4\log n}.$$

Note also that

$$(16)\quad \left(1+\frac{3f^{d-2}}{n}\right)^{f+\mu+1} \prec_{(2)} e^{(3f^{d-2}/n)(f+\mu+1)} \prec_{(14)} e^{27f^{d-1}/8n} \prec_{(12)} n^{9/2}$$

and

$$(17)\quad \mu+\mu^{2}/2f \prec_{(3,14)} \log\left(1+\frac{\mu}{f}\right)^{f+\mu}.$$

But then

$$\sum_{k=\lfloor f+\mu+1\rfloor}^{\infty}\frac{1}{k!}f^{k}\left(1+\frac{3f^{d-2}}{n}\right)^{k}$$
$$\prec \frac{f^{f+\mu+1}}{\lfloor f+\mu+1\rfloor!}\left(1+\frac{3f^{d-2}}{n}\right)^{f+\mu+1}\sum_{k=0}^{\infty}\left(\frac{f}{f+\mu}\right)^{k}\left(1+\frac{3f^{d-2}}{n}\right)^{k}$$
$$\prec_{(4,3,15)} \frac{f}{(f+\mu)^{1/2}}\frac{1}{\left(1+\frac{\mu}{f}\right)^{f+\mu}}e^{f+\mu+1}n^{9/2}\frac{f+\mu}{\mu-4\log n}$$
$$\prec_{(14,17)} ef^{3/2}n^{-(\gamma+7)}e^{f}n^{9/2} = e\left(\frac{f}{n}\right)^{3/2}n^{-(\gamma+1)}e^{f} \prec_{(13)} \frac{1}{8}n^{-(\gamma+1)}e^{f}.$$

Let $\mathscr{A}'(n, E)$ denote the set of all matrices $A \in \mathscr{A}(n, E)$ such that A has at least $f-\mu$ and at most $f+\mu$ 1's in each row and in each column of $\Omega(n)$.

LEMMA 3 *If*

$$\lim_{n\to\infty}(E^{d-1}/n^{d}) - \log n < \infty = \lim_{n\to\infty}E/(n\log n)$$

then for all sufficiently large n the probability is $< n^{-\gamma}$ that a random member of $\mathscr{A}(n, E)$ is not in $\mathscr{A}'(n, E)$.

Proof. Let ξ_i [resp. η_i] denote the probability that a random member of $\mathscr{A}(n, E)$ has fewer than $f-\mu$ [resp. more than $f+\mu$] 1's in the ith row of Ω. Then

$$\xi_i \leqq \sum_{k=0}^{[f-\mu-1]} \binom{r}{k}\binom{m-r}{E-k} \Big/ \binom{m}{E} \prec_{(1,7)} e^{1/2} \sum_{k=0}^{[f-\mu-1]} \frac{r^k}{k!}\left(\frac{E}{m}\right)^k$$
$$\times \left(1 - \frac{E}{m}\right)^{r-k} \prec_{(a)} 2e^{-f} \sum_{k=0}^{[f-\mu-1]} \frac{1}{k!} f^k\left(1 - \frac{f}{r}\right)^{-k}$$
$$\prec_{(b)} 2e^{-f} \sum_{k=0}^{[f-\mu-1]} \frac{1}{k} f^k\left(1 + \frac{2f}{r}\right)^k \prec \frac{1}{4} n^{-(\gamma+1)}.$$

For (a), note that $E/m = f/r$ and apply (2) to obtain e^{-f}. For (b), note that $(1 - x)^{-1} < 1 + 2x$ for $0 < x < \frac{1}{2}$. For the final inequality, use the first conclusion of Lemma 2.

In a similar way, it follows from the second conclusion of Lemma 2 that $\eta_i \prec n^{-(\gamma+1)}/4$, and of course the same inequalities apply to columns of Ω. To complete the proof of Lemma 3, apply these inequalities to each of the n rows and n columns of Ω.

7. Crosses, cemeteries and catacombs. For $1 \leqq k \leqq n$ let $\Omega_k(n) = \{(i, j) \in \Omega(n) : k \in \{i, j\}\}$. In accordance with our practice of suppressing n, $\Omega(n)$ and $\Omega_k(n)$ are often denoted by Ω and Ω_k respectively.

A *cemetery* is a set of crosses, and a *cross* is a subset X of Ω such that $2 < 2f - 2\mu \leqq |X| \leqq 2f + 2\mu$ and $X \subset \Omega_k$ for some k. The integer k, which is unique because $|X| > 2$, is denoted by σX and called the *station* of the cross X. For each $x \in X$, $\eta_X(x)$ denotes the coordinate of x that is different from σX.

In discussing paths that join node s to node t, the notion of an (h, s, t)-*catacomb* is employed. For an integer $h \geqq 1$, and distinct integers s and t between 1 and n, this is a sequence $\mathbf{C} = (\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}, Z)$ that satisfies the following four conditions:

(a) Z is a cross of station t;

(b) when $h \geqq 2$, $\mathscr{L}_1$ is a cemetery consisting of a single cross X_1 of station s;

(c) for $2 \leqq k \leqq h - 1$, the cemetery $\mathscr{L}_k$ is of the form

$$\{Y_{x,X} : x \in X \in \mathscr{L}_{k-1}\}$$

where, for each $X \in \mathscr{L}_{k-1}$ and $x \in X$, $Y_{x,X}$ is a cross of station $\eta_X(x)$;

(d) for each cross $X \in V\mathbf{C} = \left(\bigcup_{k=1}^{h-1} \mathscr{L}_k\right) \cup \{Z\}$, $(\bigcup V\mathbf{C}) \cap \Omega_{\sigma X} = X$.

The cemeteries $\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}$, and $\{Z\}$ may be regarded as the successive *layers* of the catacomb $\mathbf{C}$, and h as the *depth* of $\mathbf{C}$. The cemetery $V\mathbf{C}$ consists of all the crosses associated with $\mathbf{C}$, and $\bigcup V\mathbf{C}$ is the set of all positions in Ω covered by those crosses.

For each pair (i, j) of distinct integers between 1 and n, let $[i, j]$ denote whichever of (i, j) and (j, i) belongs to $\Omega(n)$. Let $\mathscr{A}(n, E, \mathbf{C})$

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press

denote the set of all matrices $A = (a_{ij}) \in \mathscr{A}(n, E)$ such that for each cross $X \in V\mathbf{C}$ and each $(i, j) \in \Omega_{\sigma X}$, $a_{ij} = 1$ if and only if $[i, j] \in X$.

LEMMA 4 *If* $\mathbf{C}$ *and* $\mathbf{D}$ *are distinct* (h, s, t)*-catacombs the sets of matrices* $\mathscr{A}(n, E, \mathbf{C})$ *and* $\mathscr{A}(n, E, \mathbf{D})$ *are disjoint.*

Proof. Note that a catacomb $\mathbf{C}$ is completely determined when the triple (h, s, t) and the set $\cup V\mathbf{C}$ are specified, for $Z = (\cup V\mathbf{C}) \cap \Omega_t$, $X_1 = (\cup V\mathbf{C}) \cap \Omega_s$, $L_1 = \{X_1\}$, and the successive layers $\mathscr{L}_k$ are then determined by conditions (c) and (d). It follows that if $\mathbf{C}$ and $\mathbf{D}$ are distinct (h, s, t)-catacombs there are distinct crosses $X \in V\mathbf{C}$ and $Y \in V\mathbf{D}$ such that $\sigma X = \sigma Y$. But then there exists $(i', j') \in \Omega_{\sigma X}$ such that $a_{i'j'} \neq b_{i'j'}$ for all $A = (a_{ij}) \in \mathscr{A}(n, E, \mathbf{C})$ and $B = (b_{ij}) \in \mathscr{A}(n, E, \mathbf{D})$.

LEMMA 5. *If* h *is a positive integer,* s *and* t *are integers between* 1 *and* n, *and* $A \in \mathscr{A}'(n, E)$, *then* $A \in \mathscr{A}(n, E, \mathbf{C})$ *for some* (h, s, t)*-catacomb* $\mathbf{C}$.

Proof. Let $Z = \{(i, j) \in \Omega : a_{ij} = 1 \text{ and } t \in \{i, j\}\}$. If $h > 1$ let $X_1 = \{(i, j) \in \Omega : a_{ij} = 1 \text{ and } s \in \{i, j\}\}$, $\mathscr{L}_1 = \{X_1\}$. For $2 \leqq k \leqq h - 1$ let $\mathscr{L}_k = \{Y_{x,X} : x \in X \in \mathscr{L}_{k-1}\}$, where

$$Y_{x,X} = \{(i, j) \in \Omega : a_{ij} = 1 \quad \text{and} \quad \eta_x(x) \in \{i, j\}\}.$$

Then set $\mathbf{C} = (\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}, Z)$.

When A is a matrix (a_{ij}) we write (a_{ij}^p) for the pth power A^p.

LEMMA 6. *If* $\mathbf{C}$ *is an* (h, s, t)*-catacomb* $(\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}, Z)$ *and* $\mathscr{A} = (a_{ij}) \in A(n, E, \mathbf{C})$ *then* $a_{st}^{h+1} \geqq 1$ *if and only if* $a_{ij} = 1$ *for some*

$$[i, j] \in W = \{[\eta_X(x), \eta_Z(z)] : x \in X \in \mathscr{L}_{h-1}, z \in Z\}.$$

Proof. A straightforward induction on p shows that

(18) for $1 \leqq p \leqq h - 1$, $a_{si}^p \geqq 1$ if and only if $i = \eta_X(x)$ for some $x \in X \in \mathscr{L}_p$.

Now if $a_{st}^{h+1} \geqq 1$ there exists j such that $a_{sj}^h \geqq 1 = a_{jt}$ and consequently there exists i such that $a_{si}^{h-1} \geqq 1 = a_{ij}$. Plainly $[j, t] \in Z$ and $j = \eta_Z([j, t])$. And by (18), $i = \eta_X(x)$ for some $x \in X \in \mathscr{L}_{h-1}$.

Suppose, on the other hand, that $a_{st}^{h+1} = 0$, and consider an arbitrary $[i, j] \in W$ with $i = \eta_X(x)$ for some $x \in X \in \mathscr{L}_{h-1}$ and $j = \eta_Z(z)$ for some $z \in Z$. Then $[j, t] \in Z$ so $a_{jt} = 1$, and since $a_{st}^{h+1} = 0$ it follows that $a_{sj}^h = 0$. But $a_{si}^{h-1} \geqq 1$ by (18), and since $a_{sj}^h = 0$ it follows that $a_{ij} = 0$.

8. Proof of theorem 1. For each (h, s, t)-catacomb $\mathbf{C}$ let $\rho(\mathbf{C})$ denote the probability that $a_{st}^{h+1} = 0$ for a random matrix $A = (a_{ij}) \in \mathscr{A}(n, E, \mathbf{C})$.

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press

LEMMA 7. *If* $\lim_{n\to\infty}(E^{d-1}/n^d) - \log n < \infty = \lim_{n\to\infty} E/(n \log n)$ *then it is true for all sufficiently large* n *that* $\rho(\mathbf{C}) > \exp(-2^{h+3}f^{h+1}/r)$ *whenever* $1 \leqq h \leqq d-2$ *and* $\mathbf{C}$ *is an* (h, s, t)*-catacomb.*

Proof. Let $\mathbf{C}$ and W be as in Lemma 6. Let $q = |V\mathbf{C}|$, $u = |\cup V\mathbf{C}|$ and $w = |W|$. Then

$$(19) \quad q \leqq 1 + \sum_{k=1}^{h-1} (2f + 2\mu)^{k-1} < 4(2f + 2\mu)^{h-1},$$

$$(20) \quad u < 4(2f + 2\mu)^{h-1} \quad \text{and} \quad w < 4(2f + 2\mu)^h.$$

For a random member A of $\mathscr{A}(n, E, \mathbf{C})$, entries are fully determined in the $2qr$ positions of Ω that belong to $\cup \{\Omega_{\sigma X} : X \in VC\}$ (and are 1 in u of those positions) but are unrestricted in the remaining $m - 2qr$ positions of Ω. Hence

$$|\mathscr{A}(n, E, C)| = \binom{m - 2qr}{E - u}.$$

By Lemma 6, the condition that $a_{st}{}^{h+1} = 0$ determines the entries of A in an additional w positions of Ω, whence

$$\rho(\mathbf{C}) = \binom{m - 2qr - w}{E - u} \Big/ \binom{m - 2qr}{E - u}$$

$$= \frac{(m - 2qr - w)!}{(m - 2qr - w - E + u)!} \Big/ \frac{(m - 2qr)!}{(m - 2qr - E + u)!},$$

and with $c_k = 1/(k^2 - k)$ it follows from two applications of (5) that

$$(21) \quad \rho(\mathbf{C}) \sim \left(1 - \frac{w}{m - 2qr}\right)^{E-u} \Big/ \exp\left(\sum_{k=2}^{\infty} \left(\frac{c_k(E - u)^k}{(m - 2qr - w)^{k-1}} - \frac{c_k(E - u)^k}{(m - 2qr)^{k-1}}\right)\right).$$

With the aid of (13), (14), (19), and (20) it follows that

$$(22) \quad \frac{qr}{m} \prec \frac{4(3f)^{h-2}}{n} \to 0 \quad \text{and} \quad \frac{w}{m} \prec \frac{4(3f)^h}{n^2} \prec 4^{h+1}\left(\frac{f^{h/2}}{n}\right)^2 \to 0,$$

whence

$$\left(1 - \frac{w}{m - 2qr}\right) \succ \left(1 - \frac{3w}{2m}\right)^E$$

and

$$\log\left(1 - \frac{3w}{2m}\right) = -\sum_{k=1}^{\infty} \frac{1}{k}\left(\frac{3w}{2m}\right)^k \succ \frac{-2w}{m}.$$

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press

Hence

$$(23)\quad \left(1 - \frac{w}{m - 2qr}\right)^{E-u} \succ e^{-2wE/m}.$$

With the aid of (22) and (13) it follows that

$$\sum_{k=5}^{\infty} \left(\frac{c_k(E-u)^k}{(m-2qr-w)^{k-1}} - \frac{c_k(E-u)^k}{(m-2qr)^{k-1}}\right) = O\left(\frac{E^5}{m^4}\right) = O\left(\frac{f^5}{n^3}\right) \to 0$$

and for $2 \leqq k \leqq 4$ that

$$\frac{c_k(E-u)^k}{(m-2qr-w)^{k-1}} - \frac{c_k(E-u)^k}{(m-2qr)^{k-1}} = \frac{c_k(E-u)^k}{(m-2qr-w)^{k-1}}$$
$$\times \left(1 - \left(1 - \frac{w}{m-2qr}\right)^{k-1}\right) \prec 2kc_k \frac{E^k}{m^k} w \to 0.$$

Using these inequalities in conjunction with (21) and (23), we conclude that for all sufficiently large n, $\rho(\mathbf{C}) \succ e^{-2wE/m}$ whenever $1 \leqq h \leqq d - 2$ and $\mathbf{C}$ is an (h, s, t)-catacomb. The proof is completed by noting that $E/m = f/r$, whence

$$e^{-wE/m} \succ_{(20)} \exp(-2^{h-2}(f+\mu)^h f/r) \succ \exp(-2^{h+3}f^{h+1}/r).$$

To justify the last inequality, note that if $0 \leqq p \leqq h$ and $h + p < 2d - 2$ then for appropriate constants ψ and ψ',

$$(24)\quad f^p\mu^{h-p}/r \sim \psi f^{(h+p)/2}(\log n)^{(h+p)/2}/n$$
$$= \psi(f^{d-1}/n)^{(h+p)/(2d-2)}(\log n)^{(h+p)/2}/n^{1-(h+p)/(2d-2)}$$
$$\prec_{(12)} \psi'(\log n)^{(h+p)/(2d-2)+(h+p/2)}/n^{1-(h+p)/(2d-2)} \to 0.$$

For each $k \geqq 1$ let $\beta_k(n, E)$ denote the probability, for a random $A = (a_{ij}) \in A(n, E)$, that $a_{12}{}^p \geqq 1$ for some $p \leqq k$. This is equal, for each pair (s, t) of distinct integers between 1 and n, to the probability that $a_{st}{}^p \geqq 1$ for some $p \leqq k$.

LEMMA 8. *If* $\lim_{n\to\infty}(E^{d-1})n^d) - \log n < \infty = \lim_{n\to\infty}E(n/\log n)$ *and* $1 \leqq h \leqq d - 2$ *then*

$$\beta_{h+1} \prec \beta_h + n^{-\gamma} + 1 - \exp(-2^{h+4}f^{h+1}/r).$$

Proof. Let β_{h+1}' [resp. β_{h+1}''] denote the probability, for a random $A = (a_{ij}) \in \mathscr{A}(n, E)$, that $a_{12}{}^{h+1} \geqq 1$ and $A \in \mathscr{A}'(n, E)$ [resp. $A \notin \mathscr{A}'(n, E)$]. Then of course

$$\beta_{h+1} \leqq \beta_h + \beta_{h+1}' + \beta_{h+1}''.$$

It follows from Lemmas 3–7 that $\beta_{h+1}'' \prec n^{-\gamma}$ and

$$\beta_{h+1}' \prec 1 - \exp(-2^{h+3}f^{h+1}/r).$$

We are now ready for

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press

THEOREM 1. *If* $E(n)^{d-1}/n^d \to 0$ *as* $n \to \infty$, *then* $P(n, E(n), < d) \to 0$.

Proof. Note that if $E'(n) = n(\log n)^2$ then

$$E'(n)^{d-1} \to 0 \quad \text{and} \quad E'(n)/(n \log n) \to \infty.$$

In view of (8), it suffices to prove Theorem 1 with E replaced by $\max(E, E')$, and then Lemma 8 is applicable.

We claim that $\beta_{h+1}(n, E) \to 0$ for $0 \leqq h \leqq d - 2$. First observe that

$$\beta_1 = \frac{E}{m} \sim 2\left(\frac{E^{d-1}}{n^d}\right)^{1/(d-1)} \frac{1}{n^{2-d/(d-1)}} \to 0.$$

Then proceed by induction, using Lemma 8 and the fact that

$$f^{h+1}/r \leqq f^{d-1}/r \sim 2E^{d-1}/n^d.$$

Now let $\tau(n, E)$ denote the probability, for a random $A \in \mathscr{A}(n, E)$, that all entries of the matrix $A + A^2 + \ldots + A^{d-1}$ in positions that belong to $\Omega(n)$ are nonzero. Then $\tau(n, E)$ is at most the probability β_{d-1} that the entry in position (1, 2) is nonzero. As $\beta_{d-1} \to 0$, the probability tends to 1 that the matrix $A + A^2 + \ldots + A^{d-1}$ has a zero entry in at least one position belonging to $\Omega(n)$. Hence as $n \to \infty$ the probability tends to 1 that a random graph in $\mathscr{G}(n, E)$ is of diameter $\geqq d$.

9. Spiders and their legs. For $1 \leqq k \leqq n$ let

$$R_k(n) = \{(i, j) \in \Omega(n) : i = k\}.$$

For each $x = (i, j) \in \Omega$, let $\eta(x) = j$, the column coordinate of x, and for $X \subset \Omega$ let $\eta X = \{\eta(x) : x \in X\}$.

A *leg* is an ordered pair $L = (L', L'')$ of disjoint subsets of Ω such that $L' \cup L'' \subset R_k$ for some k, $f - \mu \leqq |L'| \leqq f + \mu$ and $|L''| \leqq 2\mu$. The leg L is said to be *of row k*, and R_k to be the *row of* L. The sets L' and L'' are respectively the *large part* and the *small part* of L. A *legset* is a set of legs no two of which are of the same row. When $\mathscr{L}$ is a legset, $\mathscr{L}'$ denotes the set of large parts $\{L' : L \in \mathscr{L}\}$.

In discussing paths that join node s to node t, the notion of an (h, s, t)-*semispider* is employed. For an integer $h \geqq 1$, and distinct integers s and t between 1 and n, this is a sequence $\mathbf{S} = (\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}, Z)$ that satisfies the following four conditions:

(a) Z is a subset of the tth column of Ω, with $f - \mu \leqq |Z| \leqq f + \mu$;

(b) for $1 \leqq k \leqq h - 1$, $\mathscr{L}_k$ is a legset of the form $\{L_p : p \in P_k\}$, where P_k is a subset of the set $C_{k-1} = \eta(\cup \mathscr{L}_{k-1}')$ (except that $C_0 = \{s\}$) and where, for each $p \in P_k$, L_p is a leg of row p;

(c) for $1 \leqq k \leqq h - 2$, $P_k = C_{k-1}$;

(d) with $V\mathbf{S} = (\cup_{k=1}^{h-1} \mathscr{L}_k') \cup \{Z\}$, any two points of the set $\cup V\mathbf{S}$ that have a common coordinate are related in one of the following three ways:

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press

(i) both points belong to Z and hence have the same column coordinate;

(ii) for some $L \in \bigcup_{k=1}^{h-1} \mathscr{L}_k$, both points belong to L' and hence have the same row coordinate;

(iii) for some k with $1 \leqq k \leqq h-2$, the two points are $x \in \bigcup \mathscr{L}_k'$ and $y \in L_x$, so that the row coordinate of y is the column coordinate of x.

The sets $\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}$, and Z may be regarded as the successive *layers* of the semispider **S**, Z as the *body* of S, and h as the *height* of **S**. The set V**S** consists of the body of **S** and the large parts of the legs of **S**, while $\bigcup V$**S** is the set of all positions in Ω that are covered by members of V**S**.

When $h \geqq 2$ and the semispider **S** is as just described, the set $C_{h-2} \sim P_{h-1}$, which indicates the rows of "missing legs" of **S**, is called the *blemish* of **S** and denoted by B**S**. The (h, s, t)-semispider $(\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}, Z)$ is a *spider* if it is unblemished, meaning that B**S** is empty or, equivalently, that condition (c) holds for $k = h - 1$ as well as for $1 \leqq k \leqq h-2$; the extension

$$S^* = (\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}, \emptyset, Z)$$

is then an $(h+1, s, t)$-semispider with $BS^* = \eta(\bigcup \mathscr{L}_{h-1}')$. Note that the $(1, s, t)$-semispiders and $(1, s, t)$-spiders are the same, being of the form (Z) where Z satisfies condition (a). The $(2, s, t)$-semispiders are of the form $(\mathscr{L}_1, Z)$, where $\mathscr{L}_1$ is empty or consists of a single leg of row s; only those of the latter sort are spiders.

Semispiders and spiders are used as aids in counting adjacency matrices. For each (h, s, t)-semispider $\mathbf{S} = (\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}, Z)$, let $\mathscr{A}(n, E, \mathbf{S})$ denote the set of all matrices $A = (a_{ij}) \in \mathscr{A}(n, E)$ such that

(e) $a_{it} = 1$ if and only if $(i, t) \in Z$, and

(f) if **S** has a leg L of row i, then $a_{ij} = 1$ if and only if $(i, j) \in L' \cup L''$.

Two (h, s, t)-spiders

$$(25) \quad \mathbf{S} = (\mathscr{L}_1, \ldots, \mathscr{L}_{h-1}, Z) \quad \text{and} \quad \mathbf{S}^* = (\mathscr{L}_1^*, \ldots, \mathscr{L}_{h-1}^*, Z^*)$$

are said to be *equivalent* if $Z = Z^*$, $\mathscr{L}_k = \mathscr{L}_k^*$ for $1 \leqq k < h - 1$, and

$$(26) \quad \{L' \cup L'' : L \in \mathscr{L}_k\} = \{L^{*\prime} \cup L^{*\prime\prime} : L^* \in \mathscr{L}_k^*\}$$

when $k = h - 1$. Thus the legsets $\mathscr{L}_{h-1}$ and $\mathscr{L}_{h-1}^*$ need not be equal, but they differ only in the ways in which their members are divided into large and small parts.

LEMMA 9. *If S and S^* are both (h, s, t)-spiders then the sets of matrices $\mathscr{A}(n, E, \mathbf{S})$ and $\mathscr{A}(n, E, \mathbf{S}^*)$ are the same if S and S^* are equivalent and disjoint if S and S^* are not equivalent.*

Proof. The first assertion follows immediately from the relevant definitions. For the second, suppose the (h, s, t)-spiders (25) are not equivalent. If $Z \neq Z^*$ there exists i' such that $a_{i't} \neq a_{i't}^*$ whenever

(27) $\quad (a_{ij}) \in \mathscr{A}(n, E, \mathbf{S}) \quad$ and $\quad (a_{ij}^*) \in \mathscr{A}(n, E, \mathbf{S}^*)$.

Suppose, on the other hand, that $Z = Z^*$, and let k be the smallest index for which $\mathscr{L}_k \neq \mathscr{L}_k^*$. If (26) fails then

(28) there exists i' such that $\mathscr{L}_k$ includes a leg L of row i', $\mathscr{L}_k^*$ a leg L^* of row i', and $L' \cup L'' \neq L^{*\prime} \cup L^{*\prime\prime}$.

Hence by (f) there exists j' such that $a_{i'j'} \neq a_{i'j'}^*$ whenever (27) holds. If (26) holds then $k < h - 1$ and (28) holds with $L \neq L^*$ but $L' \cup L'' = L^{*\prime} \cup L^{*\prime\prime}$, whence $L' \neq L^{*\prime}$. But if (say) $x \in L' \sim L^{*\prime}$, it can be verified with the aid of (c) and (d) that $\mathscr{L}_{k+1}$ includes a leg of row $\eta(x)$ but $V\mathbf{S}^*$ does not include any leg of row $\eta(x)$. Then there exists j' such that $a_{\eta(x)j'} \neq a_{\eta(x)j'}^*$ whenever (27) holds.

LEMMA 10. *Suppose that* $\mathbf{S}$ *is an* (h, s, t)*-semispider,* l *is the number of legs of* $\mathbf{S}$*, and* q *[resp.* u*] is the number of positions in* Ω *that are covered by the rows of the legs or by the* t*th column of* Ω *[resp. by* $L' \cup L''$ *for some leg* L *of* $\mathbf{S}$ *or by* Z*]. Let* $w = |W\mathbf{S}|$*, where*

$$W\mathbf{S} = \{w : w \in \{i, j\} \text{ for some } (i, j) \in \cup V\mathbf{S}\}.$$

Then

(29) $$|\mathscr{A}(n, E, \mathbf{S})| = \binom{m - q}{E - u}$$

(30) $l \leqq (f + \mu)^{h-2}$ *and* $(l + 1)r - l \leqq q \leqq (l + 1)r$.

If $\lim_{n\to\infty}(E^{d-1}/n^d) - \log n < \infty = \lim_{n\to\infty} E/(n \log n)$ *then*

$$u \prec (f + 4\mu)^{h-1} \quad \textit{and} \quad w \prec 3(f + \mu)^{h-1}.$$

Proof. To prove (29), note that for a random member A of $\mathscr{A}(n, E, \mathbf{S})$, entries are fully determined in q positions of Ω and are 1 in u of those positions, but are unrestricted in the remaining $m - q$ positions of Ω. The inequalities (30) are immediate from the relevant definitions. For (31), note that

$$u \leqq f + \mu + l(f + 3\mu) \leqq_{(30)} (f + \mu)^{h-2}(f + 3\mu + 1) \prec (f + 4\mu)^{h-1}$$

and

$$w \leqq 2(l + 1)(f + \mu) \leqq_{(30)} 2((f + \mu)^{h-2} + 1)(f + \mu) \prec 3(f + \mu)^{h-1}.$$

LEMMA 11. *If* $\lim_{n\to\infty}(E^{d-1}/n^d) - \log n < \infty = \lim_{n\to\infty}E/(n \log n)$ *then the following is true for all sufficiently large* n*: whenever* s *and* t *are distinct integers between* 1 *and* n, S *is a* $(d-1, s, t)$*-spider, and* $\pi(S)$ *is the probability that* $a_{st}{}^d = 0$ *for a random member* $A = (a_{ij}) \in \mathscr{A}(n, E, \mathbf{S})$, *then*

$$\pi(S) \leqq \exp(-(f-\mu)^d/r).$$

Proof. If $A = (a_{ij}) \in A(n, E)$ and $a_{st}{}^d = 0$, then every product of the form

$$a_{sk(1)}a_{k(1)k(2)} \ldots a_{k(d-2)k(d-1)}a_{k(d-1)t}$$

is 0. When $A \in \mathscr{A}(n, E, \mathbf{S})$ it follows from the definitions of spiders and of $\mathscr{A}(n, E, \mathbf{S})$ that there are at least $(f-\mu)^{d-2}$ distinct choices of $k(d-2)$ appearing at the ends of sequences $(s, k(1), \ldots, k(d-2))$ that satisfy the following two conditions:

(a) $(s, k(1)) \in L'$ for some leg $(L', L'') \in \mathscr{L}_1$ and hence $a_{sk(1)} = 1$;
(b) for $2 \leqq i \leqq d-2$, $(k(i-1), k(i)) \in L'$ for some leg $(L', L'') \in \mathscr{L}_i$ and hence $a_{k(i-1)k(i)} = 1$.

Also, there are at least $f - \mu$ choices of $k(d-1)$ for which $(k(d-1), t) \in Z$ and hence $a_{k(d-1)t} = 1$. Hence there are at least $z \geqq (f-\mu)^{d-1}$ pairs $(k(d-2), k(d-1))$ in $\Omega(n)$ for which all $A = (a_{ij}) \in \mathscr{A}(n, E, \mathbf{S})$ have

$$a_{k(d-2)k(d-1)} = 0.$$

Referring to Lemma 10 and its notation, we see that for $A \in \mathscr{A}(n, E, \mathbf{S})$ with $a_{st}{}^d = 0$, entries are fully determined in at least $q + z$ positions of Ω. Hence with $c_k = 1/(k^2 - k)$ we have

$$\begin{aligned}\pi(S) &\leqq \binom{m-q-z}{E-u} \Big/ \binom{m-q}{E-u} \\ &= \frac{(m-q-z)!}{(m-q-z-E+u)!} \Big/ \frac{(m-q)!}{(m-q-E+u)!} \\ &\sim_{(5)} \left(1 - \frac{z}{m-q}\right)^{E-u} \Big/ \exp\left(\sum_{k=2}^{\infty}\left(\frac{c_k(E-u)^k}{(m-q-z)^{k-1}} \right.\right. \\ &\qquad\qquad \left.\left. - \frac{c_k(E-u)^k}{(m-q)^{k-1}}\right)\right).\end{aligned}$$

Estimating the infinite sum as in the proof of Lemma 7, we conclude that

$$\pi(S) \prec \left(1 - \frac{z}{m-q}\right)^{E-u} \leqq_{(2)} e^{-\lambda},$$

where

$$\lambda = \frac{E-u}{m-q} z \succ \frac{E}{m}(f-\mu)^{d-1} \geqq \frac{(f-\mu)^d}{m}.$$

10. Adding legs to spiders. Recall that each $(1, s, t)$-semispider $\mathbf{S}_1$ is in fact a spider and consists merely of a body. That is, $\mathbf{S}_1 = (Z)$ for some subset Z of the tth column of Ω with $f - \mu \leqq |Z| \leqq f + \mu$. The sequence $\mathbf{S}_2{}^0 = (\emptyset, Z)$ is then a $(2, s, t)$-semispider and can perhaps be extended to a $(2, s, t)$-spider $\mathbf{S}_2 = (\mathscr{L}_1, Z)$ by replacing the first "coordinate" $\emptyset$ with a legset $\mathscr{L}_1$ consisting of a single leg of row s. The $(3, s, t)$-semispider $\mathbf{S}_3{}^0 = (\mathscr{L}_1, \emptyset, Z)$ can then perhaps be extended by successive augmentations of its second coordinate, producing a sequence $\mathbf{S}_3{}^0, \mathbf{S}_3{}^1 = (\mathscr{L}_1, \mathscr{L}_2{}^1, Z), \mathbf{S}_3{}^2 = (\mathscr{L}_1, \mathscr{L}_2{}^2, Z), \ldots$, with $\mathbf{L}_2{}^1 \subset \mathbf{L}_2{}^2 \subset \ldots$ until at last a $(3, s, t)$-spider $\mathbf{S}_3 = (\mathscr{L}_1, \mathscr{L}_2, Z)$ is obtained. The $(4, s, t)$-semispider $(\mathscr{L}_1, \mathscr{L}_2, \emptyset, Z)$ can then .... It may be possible, continuing in this manner, to obtain an $(h - 1, s, t)$-spider $\mathbf{S}_{h-1}$ that extends the initial $(1, s, t)$-spider $\mathbf{S}_1$.

In the extension process just described, each augmentation may reduce the set of associated matrices. That is,

$$\mathscr{A}(n, E) \supset \mathscr{A}(n, E, \mathbf{S}_1) = \mathscr{A}(n, E, \mathbf{S}_2{}^0) \supset \mathscr{A}(n, E, \mathbf{S}_2) = \mathscr{A}(n, E, \mathbf{S}_3{}^0)$$

$$\supset A(n, E, \mathbf{S}_3{}^1) \supset A(n, E, \mathbf{S}_3{}^2) \supset \ldots \supset A(n, E, \mathbf{S}_3) \supset \ldots \supset A(n, E, \mathbf{S}_{h-1}),$$

and each of the inclusions $\supset$ may in fact be a strict inclusion. The goal of this section is to prove Lemma 13, asserting that for each pair (s, t) of distinct integers between 1 and n, "almost all" members of $\mathscr{A}(n, E)$ belong to $A(n, E, \mathbf{S})$ for some $(d - 1, s, t)$-spider $\mathbf{S}$. It is proved by successively adding legs, thus reducing and finally removing blemishes and extending semispiders to spiders. The extension is carried out with the aid of Lemma 12 below.

Recall that the family $V\mathbf{S}$ consists of the body of $\mathbf{S}$ and the large parts of the legs of $\mathbf{S}$, while $\cup\ V\mathbf{S}$ is the set of all positions in Ω that are covered by members of $V\mathbf{S}$. Recall that

$$W\mathbf{S} = \{w : w \in \{i, j\} \text{ for some } (i, j) \in \cup\ V\mathbf{S}\}$$

and R_p denotes the pth row of Ω.

When $\mathbf{S}$ is a semispider and $B\mathbf{S}$ is its blemish $C_{h-2} \sim P_{h-1}$, indicating the rows of "missing legs" of $\mathbf{S}$, let

$$R_p(\mathbf{S}) = \{(i, j) \in R_p : j \notin W\mathbf{S}\}$$

for each $p \in BS$. Let $\mathscr{A}_p(n, E, S)$ denote the set of all $A \in \mathscr{A}(n, E, S)$ such that A has at least $f - \mu$ and at most $f + \mu$ 1's in the positions of $R_p(\mathbf{S})$ and at most 2μ 1's in the positions of $R_p \sim R_p(\mathbf{S})$.

LEMMA 12. *If* $\lim_{n\to\infty}(E^{d-1}/n^d) - \log n < \infty = \lim_{n\to\infty} E/(n \log n)$ *it is true for all sufficiently large* n *that whenever* $1 \leqq h \leqq d - 2$, $\mathbf{S}$ *is an*

(h, s, t)-semispider, and $p \in B\mathbf{S}$, then the probability is at most $n^{-(\gamma-6)}$ that a random member of $\mathscr{A}(n, E, \mathbf{S})$ does not belong to $\mathscr{A}_p(n, E, \mathbf{S})$.

Proof. The probability in question is at most $\alpha + \beta + \alpha' + \beta'$, where α [resp. α'] is the probability that A has fewer than $f - \mu$ 1's in the positions of $R_p(\mathbf{S})$ [resp. R_p] and β [resp. β'] is the probability that A has more than $f + \mu$ 1's in the positions of $R_p(\mathbf{S})$ [resp. R_p]. We shall establish appropriate bounds for α, β, α' and β'.

To bound α, note that

$$\alpha \leqq \sum_{k=0}^{[f-\mu-1]} \binom{r}{k}\binom{m-q-r+w}{E-u-k} \Big/ \binom{m-q}{E-u},$$

whence by (1) and (7),

$$(32) \quad \alpha \prec \sum_{k=0}^{[f-\mu-1]} \frac{1}{k!}\left(\frac{(E-u)r}{mq}\right)^k \left(1 - \frac{E-u}{m-q}\right)^{r-w-k} e^{T_k}$$

with

$$(33) \quad T_k = \sum_{k=2}^{\infty} \frac{1}{j(j-1)} \frac{(r-w)^j}{(m-q)^{j-1}} - \frac{k^j}{(E-u)^{j-1}} - \frac{(m-w-k)^j}{(m-q-E+u)^{j-1}}.$$

By (2),

$$(34) \quad \left(1 - \frac{E-u}{m-q}\right)^{r-w-k} \leqq e^{-g} \quad \text{with} \quad g = \frac{(r-w-k)(E-u)}{m-q}.$$

But

$$\frac{u}{E} \prec_{(29)} \frac{(f+4\mu)^{h-1}}{E} \sim_{(14)} \frac{f^{h-1}}{E} = \frac{f^{h-2}}{n} \to_{(13)} 0$$

and

$$\frac{q}{m} \leqq_{(28)} \frac{((f+\mu)^{h-2}+1)r}{m} \prec_{(14)} 2\frac{f^{h-2}}{n} \to_{(13)} 0,$$

so

$$(35) \quad g \sim \frac{rE}{m} - \frac{(w+j)E}{m}.$$

Since $rE/m = f$ and

$$\frac{(w+k)E}{m} \leqq_{(31)} \frac{3(f+\mu)^{h-1} + f - \mu}{m} E \prec_{(14)} \frac{4f^{h-1}E}{m} = \frac{4f^j}{n} \prec \frac{5f^{d-1}}{n} \prec_{(12)} 7 \log n,$$

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press

we conclude from (34) and (35) that

$$(36)\quad \left(1-\frac{E-u}{m-q}\right)^{r-w-k} \prec n^7 e^{-f}.$$

From the facts that $m-2q \succ_{(3)} 0$ and $q \leqq_{(29)} (l+1)r$ it follows that

$$\frac{m}{m-q} \prec \frac{m+2r(l+1)}{m},$$

whence

$$(37)\quad \frac{(E-u)r}{m-q} \prec \frac{Er}{m}\left(1+\frac{2r(l+1)}{m}\right) \prec_{(29)} f\left(1+\frac{2(f+\mu)^{h-1}}{n}\right).$$

By reasoning similar to that of Lemma 7, $T_k \to 0$ as $n \to \infty$, the convergence being uniform in k. From this fact, in conjunction with (32)–(37) and Lemma 2, we conclude that

$$(38)\quad \alpha \prec n^7 e^{-f} \sum_{k=0}^{\lceil f-\mu-1\rceil} \frac{1}{k!} f^k\left(1+\frac{3f^{k-3}}{n}\right) \prec \frac{1}{8} n^{-(\gamma+7)}.$$

To bound β, note that

$$\beta \leqq \sum_{k=\lfloor f+\mu+1\rfloor}^{r} \binom{r}{k}\binom{m-q-4+w}{E-u-k} \Big/ \binom{m-q}{E-u},$$

whence by (1) and (7),

$$(39)\quad \beta \prec \sum_{k=\lfloor f+\mu+1\rfloor} \frac{1}{k!}\left(\frac{(E-u)r}{m-q}\right)^k \left(1-\frac{E-u}{m-q}\right)^{r-w-k} e^{T_k},$$

where T_k is given by (33). It follows with the aid of (37) that

$$(40)\quad \left(1-\frac{E-u}{m-q}\right)^{-1} \prec 1+\frac{2f}{r}\left(1+\frac{2(f+4\mu)^{d-3}}{n}\right),$$

so combining (37) and (40) we have

$$(41)\quad \frac{(E-u)r}{m-q}\left(1-\frac{E-u}{m-q}\right)^{-1} \prec f\left(1+\frac{3f^{d-3}}{n}\right).$$

Combining (40) and (41), and using (39),

$$(42)\quad \beta \prec \left(1-\frac{E-u}{m-q}\right)^{r-w} \sum_{k=\lfloor f+\mu+1\rfloor}^{\infty} \frac{1}{k!} f^k\left(1+\frac{3f^{d-3}}{n}\right)$$

$$\prec e^{-b} \sum_{k=\lfloor f+\mu+1\rfloor}^{\infty} \frac{1}{k!} f^k \left(1+\frac{3f^{d-3}}{n}\right)^k$$

where

$$b = \frac{(r-w)(E-u)}{m-q} \succ f.$$

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press

Using (42) and Lemma 2,

$$(43)\quad \beta \leqq e^{-f} \sum_{k=\lfloor f+\mu+1\rfloor}^{\infty} \frac{1}{k!} f^k \left(1 + \frac{3f^{d-3}}{n}\right)^k \prec \frac{1}{4} n^{-(\gamma+1)}.$$

To bound α', note that

$$(44)\quad \alpha' \prec \sum_{k=0}^{\lceil f-\mu-1\rceil} \binom{r}{k} \binom{m-q-r}{E-u-k} \Big/ \binom{m-q}{E-u}$$

$$\prec_{(8)} \sum_{k=0}^{\lceil f-\mu-1\rceil} \frac{1}{k!} \left(\frac{(E-u)r}{m-q}\right)^k \left(1 - \frac{E-u}{m-q}\right)^{r-k}$$

$$\prec n^7 e^{-f} \sum_{k=0}^{\lceil f-\mu-1\rceil} \frac{1}{k!} f^k \left(1 + \frac{2f^{d-3}}{n}\right)^k \prec \frac{1}{8} n^{-(\gamma-6)}.$$

Note finally that

$$(45)\quad \beta' \prec \sum_{k=\lfloor f+\mu+1\rfloor}^{r} \binom{r}{k} \binom{m-q-r+1}{E-u-k} \Big/ \binom{m-q}{E-u} \prec \frac{1}{4} n^{-(\gamma+1)}$$

as in (43). Combining (38), (43), (44) and (45) yields Lemma 12.

LEMMA 13. *If* $\lim_{n\to\infty}(E^{d-1}/n^d) - \log n < \infty = \lim_{n\to\infty} E/(n \log n)$ *it is true for all sufficiently large* n *that whenever* s *and* t *are distinct integers between* 1 *and* n *the probability is at least* $1 - n^{-4}$ *that a random member of* $\mathscr{A}(n, E)$ *belongs to* $\mathscr{A}(n, E, \mathbf{S})$ *for some* $(d-1, s, t)$-*spider* $\mathbf{S}$.

Proof. If the hypotheses of Lemma 12 are satisfied and $\mathbf{S} = (\ldots, \mathscr{L}_{n-1}, Z)$ then all but at most $(1 - n^{-(\gamma-6)})|\mathscr{A}(n, E, \mathbf{S})|$ matrices in $\mathscr{A}(n, E, \mathbf{S})$ are also in $\mathscr{A}(n, E, S^*)$ for some (h, s, t)-semispider $S^* = (\ldots, \mathscr{L}_{h-1}^*, Z)$ that agrees with $\mathbf{S}$ in all coordinates except that the legset $\mathscr{L}_{h-1}^*$ includes a leg of row p and hence $B\mathbf{S}^* = B\mathbf{S} \sim \{p\}$. Applying this successively to each $p \in B\mathbf{S}$, we find in $|B\mathbf{S}| \leqq (f + 4\mu)^{h-1}$ steps that all but at most

$$(1 - n^{-(\gamma-6)})^{|B\mathbf{S}|}$$

matrices in $\mathscr{A}(n, E, \mathbf{S})$ are also in $\mathscr{A}(n, E, \mathbf{S}')$ for some (h, s, t)-spider $\mathbf{S}' = (\ldots, \mathscr{L}_{h-1}', Z)$ that extends $\mathbf{S}$. If $h + 1 \leqq d - 2$ the same extension process can then be applied to the $(h + 1, s, t)$-semispider $(\ldots, L_{h-1}', \emptyset, Z)$.

By choosing an appropriate body Z consisting of at least $f - \mu$ and at most $f + \mu$ members of the tth column of Ω, and then initiating the above extension process with the $(2, s, t)$-semispider $(\emptyset, Z)$, we find that all but at most

$$(1 - n^{-(\gamma-6)})^a |\mathscr{A}(n, E)|$$

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press

members of $\mathscr{A}(n, E)$ are in $\mathscr{A}(n, E, \mathbf{S})$ for some $(d-1, s, t)$-spider $\mathbf{S}$, where

$$a \leqq \sum_{h=0}^{d-2} (f+4\mu)^{h-1} \prec 2^{d-2} f^{d-3} \prec 2\, n \log n / f^2.$$

Thus

$$(1 - n^{-(\gamma-6)})^a \succ (1 - n^{-(\gamma-6)})^{2n \log n / f^2} \succ 1 - n^{-4}$$

if $\gamma \geqq 12$.

11. Proof of theorem 2. We are now ready for the proof of Theorem 2, which depends on Lemmas 1, 3, 9, 11, and 13.

THEOREM 2. *If* $(E(n)^d/n^{d+1}) - \log n \to \infty$ *as* $n \to \infty$ *then*

$$P(n, E(n), > d) \to 0.$$

Proof. It suffices to consider the case in which there exists (possibly infinite)

$$\lambda(c) = \lim_{n\to\infty}(E(n)^c/n^{c+1}) - \log n$$

for each integer c between 1 and d. Since $\lambda(1) = -\infty$ and $\lambda(d) = \infty$, there exists c with $2 \leqq c \leqq d$ such that $\lambda(c-1) < \infty = \lambda(c)$. The case $c = 2$ is settled by Lemma 1, so we may assume $c \geqq 3$. And it suffices to prove $P(n, E(n), > c) \to 0$, for then surely $P(n, E(n), > d) \to 0$. In other words, we may assume without loss of generality that $d \geqq 3$ and

$$\lim_{n\to\infty}(E(n)^{d-1}/n^d) - \log n < \infty = \lim_{n\to\infty}(E(n)^d/n^{d+1}) - \log n.$$

We may also assume the existence of

$$\tau = \lim_{n\to\infty}(E(n)^{d-1}/n^d)\mu(n).$$

To justify the application of certain lemmas, note that

$$(46) \quad \frac{E}{n \log n} > \frac{E}{n^{(2d+1)/2d} \log n} = \left(\frac{E^d}{n^{d+1}}\right)^{1/d} \frac{n^{1/2d}}{\log n} \to \infty.$$

Now consider an arbitrary pair (s, t) of distinct integers between 1 and n, and let $\{\mathbf{S}_\iota : \iota \in I\}$ be a family of $(d-1, s, t)$-spiders $\mathbf{S}_\iota$ that includes precisely one representative from each equivalence class. By Lemma 9, the sets of matrices $\mathscr{A}(n, E, \mathbf{S}_\iota)$ and $\mathscr{A}(n, E, \mathbf{S}_{\iota'})$ are disjoint whenever $\iota \neq \iota'$, and the union

$$\cup_{\iota\in I} \mathscr{A}(n, E, \mathbf{S}_\iota)$$

is equal to the union of the sets $\mathscr{A}(n, E, \mathbf{S})$ for all $(d-1, s, t)$-spiders $\mathbf{S}$. Lemmas 11 and 13 imply that for all sufficiently large n, both of the following are true for all (s, t):

The probability is at least $1 - n^{-4}$ that a random member of $\mathscr{A}(n, E)$ belongs to $\bigcup_{\iota \in I} \mathscr{A}(n, E, \mathbf{S}_\iota)$;

For each $\iota \in I$, the probability is at most $\exp(-(f - \mu)^d/r)$ that $a_{st}{}^d = 0$ for a random member $A = (a_{ij})$ of $\mathscr{A}(n, E, \mathbf{S}_\iota)$.

It follows that for all sufficiently large n,

(47) $\quad \pi_{st} \leqq n^{-4} + \exp(-(f - \mu)^d/r),$

where π_{st} is the probability that $a_{st}{}^d = 0$ for a random member $\mathscr{A} = (a_{ij})$ of $A(n, E)$.

Let $\theta(n, E)$ denote the expectation, for a random $A \in \mathscr{A}(n, E)$, of the number of zero entries of the matrix A^d in positions that belong to $\Omega(n)$. If $\theta(n, E) \to 0$ it follows from the observations in Section 2 that $P(n, E, >d) \to 0$. By (47),

$$\theta(n, E) = \sum_{(s,t) \in \Omega(n)} \pi_{st} \prec m(n^{-4} + \exp(-(f - \mu)^d/r).$$

Since $m/n^4 \to 0$, $f/\mu \to \infty$ by (14), and $f^p \mu^{d-p}/r \to 0$ by (24) when $0 \leqq p < d - 2$, it remains only to show that

$$m \exp\left(-\frac{f^d}{r} + \frac{df^{d-1}\mu}{r}\right) \to 0.$$

We establish the slightly stronger fact that

$$\frac{f^d}{r} - \frac{f^{d-1}\mu}{r} - 2 \log n \to \infty.$$

Since $f^{d-1}\mu/r \to 2\tau$, it suffices when $\tau < \infty$ to observe that

$$\frac{f^d}{r} - 2 \log n \sim 2\left(\frac{E^d}{n^{d+1}} - \log n\right) \to \infty.$$

For the case in which $\tau = \infty$, note that

$$\frac{f^d}{r}\left(\frac{\log n}{f}\right)^{1/2} \to \infty$$

and hence

$$\frac{f^d}{r} - \frac{f^{d-1}\mu}{r} - 2 \log n = \frac{f^d}{r}\left(1 - \frac{\mu}{f}\right) - 2 \log n$$

$$\succ \left(\frac{f}{\log n}\right)^{1/2}\left(1 - (2\gamma + 14)^{1/2}\left(\frac{\log n}{f}\right)^{1/2}\right) - 2 \log n \to_{(a)} \infty.$$

To justify (a), use the fact that

$$\frac{f^{1/2}}{(\log n)^{3/2}} = \frac{E^{1/2}}{n^{1/2}(\log n)^{3/2}} = \left(\frac{E}{n^{(2d+1)/2d} \log n}\right)^{1/2} \frac{n^{1/4d}}{\log n} \to_{(46)} \infty.$$

12. Directions for further research. It seems likely that Theorem 1 holds under the weaker assumption that

$$(48) \quad \lim_{n\to\infty} \left(\frac{E(n)^{d-1}}{n^d} - \log n\right) < \infty$$

but our argument does not establish this.

It would be interesting to study the gap between our theorems and the connectedness result of [**3**], especially the case in which (48) holds for all finite d but

$$(E(n)/n) - \tfrac{1}{2}\log n \to \infty.$$

Plainly the expected diameter of a connected graph in $\mathscr{G}(n, E)$ tends to ∞ as $n \to \infty$ in this range. Probably for $0 < p < 1$ it is possible to choose E in this range so that the expected diameter in $\mathscr{G}(n, E)$ is of the order of n^p as $n \to \infty$.

Finite-state automata are defined in terms of directed graphs, and for them the parameters δ and δ_c are of fundamental importance. Results of Barzdin' and Korsunov (see [**4**] for references) show that in certain classes of directed graphs associated with automata, the expected values of δ_c are much less than the maximum values. A later result of Korsunov [**4**] is that if λ is a real constant $\geqq 2$, $E_\lambda(n) = \lfloor \lambda n \rfloor$, and $P_\lambda(n)$ is the probability, for a random $G \in \mathscr{G}(n, E_\lambda(n))$, that

$$\tfrac{1}{2}\log_\lambda n < \delta_c(G) < 10 \log_\lambda n,$$

then $P_\lambda(n) \to 1$ as $n \to \infty$. So far as we know, a proof of this interesting result has never been published. It is described in [**5**] as being "not trivial at all."

References

1. N. Deo, *Breadth- and depth-first searches in graph-theoretic algorithms*, Report No. TRCS-74-001 (Computer Science Centre, Indian Institute of Technology, Kanpur, 1974).
2. P. Erdös, *The art of counting: Selected writings* (MIT Press, Cambridge, Mass., 1973).
3. P. Erdös and A. Rényi, *On random graphs I*, Publ. Math. (Debrecen) *6* (1959), 290–297.
4. A. D. Koršunov. *On the diameter of graphs*, Soviet Math. Dokl. *12* (1971), 302–305.
5. J. Moon and L. Moser, *Almost all* (0, 1) *matrices are primitive*, Studia Sci. Math. Hungar. *11* (1966), 153–156.

University of Washington,
Seattle, Washington;
University College,
London, England

https://doi.org/10.4153/CJM-1981-050-1 Published online by Cambridge University Press