Math. Proc. Camb. Phil. Soc. (1975), 77, 313
MPCPS 77-27
Printed in Great Britain

On colouring random graphs

BY G. R. GRIMMETT AND C. J. H. McDIARMID
Mathematical Institute, Oxford

(*Received* 3 *May* 1974)

Abstract. Let ω_n denote a random graph with vertex set $\{1, 2, \ldots, n\}$, such that each edge is present with a prescribed probability p, independently of the presence or absence of any other edges. We show that the number of vertices in the largest complete subgraph of ω_n is, with probability one,

$$\frac{2}{\log 1/p} \log n + o(\log n) \quad \text{as} \quad n \to \infty.$$

It follows from this result that the chromatic number of ω_n is, with probability one, at least

$$\tfrac{1}{2} \log 1/q \frac{n}{\log n} + o\left(\frac{n}{\log n}\right) \quad \text{as} \quad n \to \infty;$$

we conjecture that this is also an upper bound. Finally, we study a particular graph colouring algorithm and show that in probability it requires twice this number of colours to colour the graph ω_n.

Introduction. We are concerned with colouring random graphs in which each pair of vertices is joined by an edge with a prescribed probability. So far as we know, no authors have attempted an analytic approach to this problem; but Matula, Marble and Isaacson (5) have discussed the efficiency of certain graph colouring algorithms, and Holgate (3) has established statistical bounds for the chromatic number of random graphs by studying the colouring algorithms of Welsh and Powell (6), Wilf (7) and others. Erdös and Rényi referred briefly to random colouring problems at the end of a very interesting paper (1) in which they describe the probable structure of large random graphs. However, the probabilistic structure of their model differs from that which we study here. They considered graphs on n labelled vertices with $N(n)$ edges, where N is a prescribed function and each N-subset of the $\binom{n}{2}$ possible edges occurs with equal probability. In this paper, we postulate that each edge occurs with a prescribed probability p independently of the presence or absence of any other edges. Thus, the number of edges in a random graph on n vertices is a random variable with expectation $\tfrac{1}{2}pn(n-1)$, which further contrasts with (1) since Erdös and Rényi considered only cases satisfying $N(n) = o(n^2)$. However, they note that for the problems which they treat there is no essential difference between the two approaches when p is set to $p(n) = N(n) \Big/ \binom{n}{2}$.

The general theory of graph colouring is not only of interest to the graph theoretician. It also has applications in such fields as operations research, since many problems which may be formulated in terms of graphs have solutions which are specified by a vertex colouring using as few colours as possible. In the literature (see for example (5)), several algorithms are described which aim to colour an arbitrary graph G with a number $\Gamma(G)$ of colours, so that $\Gamma(G)$ is close to the chromatic number $\chi(G)$ whenever possible. It is important to be able to calculate the average size of $\Gamma(G)-\chi(G)$ over, say, all graphs with n vertices, and it has been customary, though not very satisfactory, to test algorithms on certain standard classes of graph or on a set of pseudo-randomly generated graphs. Of course, for the simple Welsh–Powell algorithm (6) there exist graphs for which $\Gamma-\chi$ is arbitrarily large, and this may be true for other colouring techniques. A major difficulty in investigating the behaviour of $\Gamma-\chi$ empirically is the amount of computer time required to find the chromatic number of a large graph, and little has been done in this direction save direct comparison of the numbers of colours used by the different algorithms. A large part of this paper is devoted to a detailed analysis of the number of colours used by a very simple method of colouring a random graph. We make a conjecture which, if its truth were established, would imply that this algorithm usually uses about twice the minimum number of colours required.

1. *Definitions.* A *graph* $G = (V, E)$ is a non-empty set V of *points* or *vertices*, together with a set E of distinct unordered pairs $\{u, v\}$ with $u, v \in V$, $u \neq v$. Each element $\{u, v\}$ of E is an *edge* and *joins* u to v. The vertices of an edge are called its *end points*. So, we are concerned with undirected graphs with neither loops nor multiple edges. For the general theory of graphs see (2) or (8).

A graph $G' = (V', E')$ is a *subgraph* of a graph $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. The subgraph of $G = (V, E)$ *induced* by the subset V' of V is the graph $G' = (V', E')$ where E' is the set of all elements of E which join pairs of elements of V'. Thus, the subgraph of G induced by V' is the largest subgraph of G with vertex set V'. Unless we specify otherwise, any subgraph referred to in this paper is the subgraph which is induced by its vertex set.

The *complementary graph* $\bar{G}$ of $G = (V, E)$ is the graph $(V, \bar{E})$, where $\{u, v\} \in \bar{E}$ if and only if $\{u, v\} \notin E$, for each $u, v \in V$.

A graph is called *complete* if each and every pair of vertices is joined by an edge. A *null* graph is the complementary graph of a complete graph; it has no edges.

A *colouring* of a graph is an assignment of colours to the vertices such that no two vertices have the same colour if they are joined by an edge. So, a colouring of a graph gives a partition of its vertex set such that each set in the partition induces a null subgraph. A *k-colouring* is a colouring which uses k colours, and a graph is said to be *k-colourable* if it has a k-colouring. The *chromatic number* of a finite graph is the least integer k for which the graph is k-colourable. The *edge chromatic number* or *chromatic index* of a finite graph is the least number of colours required in order to colour each edge so that no two edges with a common end point have the same colour.

The *degree* of a vertex is the number of edges of which it is an end point.

Let Ω be the set of all graphs with vertex set $\mathbb{N} = \{1, 2, 3, \ldots\}$. By the following routine argument, we will define a probability measure on a certain σ-field of subsets of Ω in such a way that the probability that an edge joins any two given vertices of a randomly chosen graph is a fixed number p, independently of any set of information about the presence or absence of other edges.

If $\omega \in \Omega$ and $I \subseteq \mathbb{N}$, we write $\omega(I)$ for the subgraph of ω induced by the vertices I. If $I = \{1, 2, \ldots, n\}$ then we write ω_n for $\omega(I)$.

For each finite subset $I \subseteq \mathbb{N}$ and each graph G with vertex set I, let

$$[G: I] = \{\omega \in \Omega: \omega(I) \text{ is } G\}.$$

That is, $[G: I]$ is the subset of Ω consisting of all members of Ω which have G as their subgraph induced by I. The set of *finite-dimensional cylinders* of Ω is the set of all such $[G: I]$ as G ranges over all graphs on finite subsets I of $\mathbb{N}$. We use $\mathscr{F}$ to denote the smallest σ-field of subsets of Ω which contains all the finite-dimensional cylinders of Ω. Now, we are able to define a probability measure P on $(\Omega, \mathscr{F})$ by specifying its value on each finite-dimensional cylinder as follows:

$$P([G: I]) = p^e q^{\binom{v}{2}-e},$$

where p is a prescribed number such that $0 < p = 1-q < 1$, and $v\,(=|I|)$ and e, respectively, are the numbers of vertices and edges in the finite graph G. We ignore the trivial cases $p = 0, 1$.

For any real x, $[x]$ and $\{x\}$ denote the greatest integer not greater than x and the least integer not less than x respectively.

2. *Largest complete subgraph.* In this section we study the size of the largest subset of $\{1, 2, \ldots, n\}$ which induces a complete subgraph of ω_n.

THEOREM 1. *Let $K_n(\omega)$ denote the size of the largest complete subgraph of ω_n. The sequence $\{K_n\}$ of random variables satisfies*

$$\frac{K_n}{\log n} \to \frac{2}{\log 1/p} \quad \text{as} \quad n \to \infty$$

almost surely and in any mean.

In section 3, we use this theorem to derive information about the chromatic number of ω_n.

We require a lemma in order to prove the theorem. Let $N_n^k(\omega)$ denote the number of complete subgraphs of the graph ω_n $(n \geqslant k)$, which have exactly k vertices, and let $\Omega_k \subseteq \Omega$ be the event that ω_k is a complete graph.

LEMMA.

$$\frac{E(N_n^k)}{E(N_n^k \mid \Omega_k)} \leqslant P(K_n \geqslant k) \leqslant E(N_n^k).$$

This was noted by Matula (4). We give a more direct proof.

Proof. Clearly

$$\begin{aligned} P(K_n \geqslant k) &= P(\text{some complete subgraph of } \omega_n \text{ has } k \text{ vertices}) \\ &\leqslant E(N_n^k). \end{aligned}$$

Also
$$E(N_n^k) = E(N_n^k|K_n \geqslant k)\,P(K_n \geqslant k)$$
$$\leqslant E(N_n^k|\Omega_k)\,P(K_n \geqslant k).$$

Note that
$$E(N_n^k) = \binom{n}{k} p^{\frac{1}{2}k(k-1)}, \tag{2·1}$$
$$E(N_n^k|\Omega_k) = \sum_{i=m}^{k} \binom{k}{i}\binom{n-k}{k-i} p^{\frac{1}{2}k(k-1)-\frac{1}{2}i(i-1)}, \tag{2·2}$$
where $m = \max(0, \{k-\frac{1}{2}n\})$.

Proof of Theorem 1. First, we prove the almost sure convergence. Let
$$k(n) = \left\{\frac{2}{\log 1/p}\log n\right\}, \quad (n = 1, 2, \ldots).$$
Then, by the lemma and (2·1),
$$P(K_n \geqslant k(n)) \leqslant \binom{n}{k} p^{\frac{1}{2}k(k-1)}$$
$$\leqslant \frac{n^k}{k!} p^{\frac{1}{2}k(k-1)}$$
$$\leqslant \frac{n}{k!}$$
$$= o(n^{-i}) \quad \text{for any integer } i. \tag{2·3}$$

Thus, by the Borel–Cantelli lemma
$$\limsup_{n\to\infty} \frac{K_n}{\log n} \leqslant \frac{2}{\log 1/p}$$
almost surely.

Now, let $0 < \epsilon < 1$,
$$d(n) = \left[\frac{2(1-\epsilon)}{\log 1/p}\log n\right], \quad (n = 1, 2, \ldots),$$
and let $T_i(n)$ be given by
$$T_i(n)\binom{n}{d(n)} = \binom{d(n)}{i}\binom{n-d(n)}{d(n)-i} p^{-\frac{1}{2}i(i-1)}, \quad (n = 1, 2, \ldots; i = 1, 2, \ldots, d(n)).$$

Then, by the lemma, (2·1) and (2·2)
$$P(K_n < d(n)) \leqslant 1 - \left(\sum_{i=0}^{d(n)} T_i(n)\right)^{-1}$$
$$\leqslant \left(\sum_{i=0}^{d(n)} T_i(n)\right) - 1 \tag{2·4}$$
because
$$\sum_{i=0}^{d(n)} T_i(n) \geqslant 1.$$

Now
$$T_0(n) = \left(1-\frac{d}{n}\right)\left(1-\frac{d}{n-1}\right)\cdots\left(1-\frac{d}{n-d+1}\right)$$
$$= 1 - \frac{d^2}{n} + o(n^{-\frac{3}{2}})$$

since $(1-x)^n \geqslant 1-nx$ and $1-x \leqslant e^{-x} \leqslant 1-x+o(x^{\frac{7}{4}})$ for $x > 0$. Also

$$T_1(n) = T_0(n)\left(\frac{d^2}{n}+o(n^{-\frac{3}{2}})\right),$$

and we deduce that

$$T_0(n)+T_1(n)-1 = o(n^{-\frac{3}{2}}). \tag{2·5}$$

Furthermore, for $2 \leqslant j \leqslant d(n)$,

$$\begin{aligned}\frac{T_j(n)}{T_2(n)} &= \frac{(n-2d+2)\,!}{(n-2d+j)\,!}\left(\frac{(d-2)\,!}{(d-j)\,!}\right)^2 \frac{2p^{1-\frac{1}{2}j(j-1)}}{j\,!}\\ &\leqslant \left(\frac{d^2}{n-2d}p^{-\frac{1}{2}(j+1)}\right)^{j-2}.\end{aligned}$$

But $p^{-\frac{1}{2}(j+1)} \leqslant p^{-\frac{1}{2}(d+1)} = o(n^{1-\frac{1}{2}\epsilon})$, and so for sufficiently large n we have that

$$\frac{T_j(n)}{T_2(n)} \leqslant 1, \quad (j = 1, 2, \ldots, d(n)). \tag{2·6}$$

It is easy to check that $T_2(n) = o(n^{-\frac{7}{4}})$, and therefore

$$\begin{aligned}P(K_n < d(n)) &\leqslant T_0(n)+T_1(n)-1+d(n)\,T_2(n)\\ &= o(n^{-\frac{3}{2}}),\end{aligned} \tag{2·7}$$

by (2·4), (2·5) and (2·6). By the Borel–Cantelli lemma

$$\liminf_{n\to\infty}\frac{K_n}{\log n} \geqslant \frac{2}{\log 1/p}$$

almost surely, and we have shown the almost sure convergence of Theorem 1. Convergence in any mean now follows easily from (2·3) and (2·7), since $0 \leqslant K_n \leqslant n$.

3. *Chromatic Number*. Several authors (see (1), (3), (5)) have investigated properties of the chromatic number of a random graph. In this section we present results which partially describe its asymptotic limit for large graphs.

We deduce the following from Theorem 1.

COROLLARY 2. *Let $I_n(\omega)$ denote the size of the largest null subgraph of the graph ω_n. Then the sequence $\{I_n\}$ satisfies*

$$\frac{I_n}{\log n} \to \frac{2}{\log 1/q} \quad \text{as} \quad n \to \infty$$

almost surely and in any mean.

Proof. Since each edge of $\omega \in \Omega$ is absent with probability q, Theorem 1 holds with K_n and p replaced by I_n and q.

COROLLARY 3. *Let $\chi_n(\omega)$ denote the chromatic number of the graph ω_n. Then the sequence $\{\chi_n\}$ satisfies*

$$\liminf_{n\to\infty}\left(\chi_n\frac{\log n}{n}\right) \geqslant \tfrac{1}{2}\log 1/q$$

almost surely.

Proof. Since each colour set in a colouring of ω_n contains at most $I_n(\omega)$ vertices, we have that $\chi_n(\omega)\, I_n(\omega) \geqslant n$. Hence

$$P\left(\liminf_{n\to\infty}\left(\chi_n \frac{\log n}{n}\right) < \tfrac{1}{2}\log 1/q\right) \leqslant P\left(\limsup_{n\to\infty}\frac{I_n}{\log n} > \frac{2}{\log 1/q}\right) = 0$$

by Corollary 2.

Some authors (for example (5)) have used the size of the largest complete subgraph of a graph as a lower bound for its chromatic number. It is obvious from Corollary 3 and Theorem 1 that, for random graphs, $\chi_n I_n \geqslant n$ provides a much better lower bound than $\chi_n \geqslant K_n$, when n is large.

In section 4, we show that

$$P\left(\chi_n \frac{\log n}{n} \leqslant (1+\epsilon)\log 1/q\right) \to 1 \quad \text{as} \quad n\to\infty \tag{3·1}$$

for any $\epsilon > 0$, but we are unable to improve this upper bound for the limit of the sequence $\{\chi_n(\log n/n)\}$. Indeed we cannot show that the sequence converges at all. However, if E is the event that it converges, then the probability of E is either zero or one. For, E is in the σ-field of events generated by the sequence $(\chi_1, \chi_2, \ldots)$ and is clearly independent of any finite subsequence. The assertion follows by a zero-one law. We make the following conjecture about the sequence.

Conjecture. The sequence $\{\chi_n\}$ *satisfies*

$$\chi_n \frac{\log n}{n} \to \tfrac{1}{2}\log 1/q \quad \text{as} \quad n\to\infty.$$

We have made little progress towards establishing the truth of this conjecture, but we see the next two theorems as evidence to support it.

THEOREM 4. *Let* $C_n^k(\omega)$ *denote the number of different k-colourings of the graph* ω_n. *Then, setting* $k(n) = [\alpha \log 1/q\,(n/\log n)]$, *we have that, as* $n\to\infty$,

$$E(C_n^k) \to \begin{cases} 0 & \text{if} \quad \alpha \leqslant \tfrac{1}{2}, \\ \infty & \text{if} \quad \alpha > \tfrac{1}{2}. \end{cases}$$

THEOREM 5. *For any positive integer r let* $I_n^r(\omega)$ *denote the number of vertices in the largest r-colourable subgraph of* ω_n. *Then the sequence* $\{I_n^r\}$ *satisfies*

$$\frac{I_n^r}{\log n} \to \frac{2r}{\log 1/q} \quad \text{as} \quad n\to\infty$$

almost surely and in any mean.

It would be valuable to know the behaviour of I_n^r when r increases with n. Certainly, a result similar to Theorem 5 holds so long as $\log r = o(\log n)$, but what happens if

$$r = \left[\alpha \log 1/q \frac{n}{\log n}\right]?$$

Proof of Theorem 4. Let $k(n) = [\alpha \log 1/q\,(n/\log n)]$. Given a collection of k colours, there are k^n different ways of assigning a colour to each point of $\{1, 2, \ldots, n\}$. Let $\mathscr{A}$ be such an assignment and let P_i $(1 \leqslant i \leqslant k)$ be the number of points coloured with the ith colour. The probability that $\mathscr{A}$ is a (proper) colouring of ω_n is

$$\begin{aligned}\prod_{i=1}^{k} q^{\frac{1}{2}P_i(P_i-1)} &= q^{\frac{1}{2}(\Sigma P_i^2 - n)} \\ &\leqslant q^{\frac{1}{2}(n^2/k - n)}\end{aligned}$$

by the root mean square inequality. Hence

$$\begin{aligned}E(C_n^k) &\leqslant k^n q^{\frac{1}{2}(n^2/k-n)} \\ &\leqslant \exp\left(n \log k - \frac{n \log n}{2\alpha} + \tfrac{1}{2} n \log 1/q\right) \\ &\to 0 \quad \text{if} \quad \alpha \leqslant \tfrac{1}{2}.\end{aligned}$$

It is easy to construct a second proof of Corollary 3 from this result.

Now let $d(n) = [n/k(n)]$. Then $kd \leqslant n \leqslant k(d+1)$, and $C_n^k(\omega)$ is at least the number of k-colourings of ω_n for which the sets in the corresponding partitions of $\{1, 2, \ldots, n\}$ each have size d or $d+1$. Hence

$$E(C_n^k) \geqslant N(n)\, q^{\frac{1}{2}kd(d+1)},$$

where $N(n)$ is the number of partitions of $\{1, 2, \ldots, n\}$ into k sets each of size d or $d+1$. But then $N(n)$ is at least the number of partitions of $\{1, 2, \ldots, kd\}$ into k sets each of size d, and so

$$N(n) \geqslant \frac{(kd)!}{(d!)^k}.$$

Hence

$$\begin{aligned}E(C_n^k) &\geqslant \frac{(kd)!}{(d!)^k} q^{\frac{1}{2}kd(d+1)} \\ &\to \infty \quad \text{as} \quad n \to \infty \quad \text{if} \quad \alpha > \tfrac{1}{2},\end{aligned}$$

using Stirling's approximation for factorials. This completes the proof of Theorem 4.

Proof of Theorem 5. This follows from Corollary 2 on choosing r disjoint subsets of the vertex set $\{1, 2, \ldots, n\}$, each of cardinality $[n/r]$.

Finally, we consider briefly the chromatic index of a random graph. This is straightforward.

THEOREM 6. *Let $\chi_n'(\omega)$ denote the chromatic index of ω_n. Then the sequence $\{\chi_n'\}$ satisfies*

$$\frac{\chi_n'}{n} \to p \quad \textit{as} \quad n \to \infty$$

almost surely and in any mean.

Proof. Let $\Delta_n(\omega)$ denote the maximum degree of the vertices of ω_n. By Vizing's theorem (see (8) for example), $0 \leqslant \chi_n' - \Delta_n \leqslant 1$ for all n and ω. Thus it is sufficient to show that

$$\frac{\Delta_n}{n} \to p \quad \text{as} \quad n \to \infty$$

almost surely and in any mean. To see this, define $f(x,y)$ by

$$f(x,y) = \left(\frac{x}{y}\right)^{v}\left(\frac{1-x}{1-y}\right)^{1-v}, \quad (0 < x, y < 1).$$

Then $f(x,x) = 1$ and it is easy to check that $f(x,y) < 1$ for $x \neq y$. Choose ϵ such that $0 < \epsilon < \min\{p,q\}$, and let $m = m(n) = \{n(p+\epsilon)\}$. The degree of each vertex of ω_n is a binomial random variable with expectation less than m. Hence

$$\begin{aligned} P(\Delta_n \geqslant m) &\leqslant n\,P(\text{degree of vertex 1 is at least } m) \\ &\leqslant n^2\binom{n}{m}p^m q^{n-m} \\ &= An^{\frac{3}{2}}(f(p,p+\epsilon))^n\,(1+o(1)), \end{aligned}$$

where A is a positive constant and we have used Stirling's approximation. Similarly, writing $r = r(n) = [n(p-\epsilon)]$,

$$P(\Delta_n \leqslant r) \leqslant Bn^{\frac{3}{2}}(f(p,p-\epsilon))^n\,(1+o(1)),$$

where $B > 0$ is constant; and the almost sure convergence of the theorem follows from the Borel–Cantelli lemma. It is easy to show convergence in any mean since $0 \leqslant \Delta_n < n$ for all n and ω.

4. *A Colouring Algorithm.* Next, we describe a simple graph colouring algorithm and study the number of colours which it uses to colour the vertices of a random graph.

Let G be any graph with vertex set $\{1, 2, \ldots, n\}$ and suppose that the available colours are $\{c_1, c_2, c_3, \ldots\}$. We colour vertex 1 with c_1, and then proceed to colour the remaining vertices in increasing order, using c_j to colour vertex i $(2 \leqslant i \leqslant n)$ where j is the least positive integer such that no vertex already coloured c_j is joined (in G) to i. In the literature (see (5), (6), (7)) several authors describe similar algorithms which colour the vertices of a graph in an order which depends upon prior information about the degrees of the vertices. For example, Matula and others (5) made computer simulations using four different criteria for ordering the vertices of a pseudo-random graph. In general, of course, one would expect these techniques to require fewer colours than the algorithm which we have described above. Nevertheless, we are forced by considerations of independence to use this latter technique.

First, we consider the set of vertices of ω_n which are coloured c_1. Certainly vertex 1 belongs to this set and further vertices are admitted in increasing order, vertex i being included if and only if it is joined to no j $(< i)$ which is already a member. Obviously this set induces a null subgraph of ω_n.

THEOREM 7. *Let $\sigma_n(\omega)$ denote the number of vertices of ω_n which are coloured with the first colour. Then the sequence $\{\sigma_n\}$ satisfies*

$$\frac{\sigma_n}{\log n} \to \frac{1}{\log 1/q} \quad \textit{as} \quad n \to \infty$$

almost surely and in any mean.

This should be compared with Corollary 2. We will use this to prove the main result of this section.

THEOREM 8. *Let* $\Gamma_n(\omega)$ *denote the number of colours required by the algorithm to colour* ω_n. *Then the sequence* $\{\Gamma_n\}$ *satisfies*

$$\Gamma_n \frac{\log n}{n} \to \log 1/q \quad \textit{as} \quad n \to \infty$$

in mean, and hence in probability. Furthermore

$$\liminf_{n\to\infty} \left(\Gamma_n \frac{\log n}{n}\right) = \log 1/q$$

almost surely.

Note that, as with the conjecture in section 3, application of a zero-one law shows that the convergence in Theorem 8 holds with probability zero or one.

Since $\chi_n \leqslant \Gamma_n$, we deduce immediately that (3·1) holds.

Proof of Theorem 7. Let $\omega \in \Omega$, and $\rho_i: \Omega \to \{0, 1, 2, \ldots\}$ $(i = 0, 1, 2, \ldots)$ be given by

$$\rho_0(\omega) = 0, \quad \rho_i(\omega) = \min\{n: \sigma_n(\omega) = i\} \quad \text{for} \quad i \geqslant 1.$$

Then $\delta_i = \rho_{i+1} - \rho_i$ $(i = 0, 1, 2, \ldots)$ defines a sequence of independent random variables with $\delta_0 = 1$ and the δ_i $(i \geqslant 1)$ distributed geometrically as

$$P(\delta_i = j) = q^i(1-q^i)^{j-1}, \quad (j = 1, 2, 3, \ldots). \tag{4·1}$$

Furthermore
$$\rho_j = \sum_{i=0}^{j-1} \delta_i, \quad (j = 1, 2, 3, \ldots). \tag{4·2}$$

Now let $\epsilon > 0$ and r be a positive integer. Choose $m = m(\epsilon)$ such that $\epsilon m(\epsilon) > r$. Then, writing

$$k = k(n) = \left\{\frac{(1+\epsilon)}{\log 1/q} \log n\right\},$$

we have

$$\begin{aligned} P(\sigma_n > k(n)) &\leqslant P(\rho_k < n) \\ &\leqslant P(\delta_i \leqslant n, \quad \text{for each} \quad i = 0, 1, \ldots, k-1) \\ &= \prod_{i=0}^{k-1} (1-(1-q^i)^n) \end{aligned}$$

by (4·1) and (4·2). Now, $k(n) > m$ for all sufficiently large n, and so

$$\begin{aligned} P(\sigma_n > k(n)) &\leqslant \prod_{i=k-m}^{k-1} (1-(1-q^i)^n) \\ &\leqslant (1-(1-q^{k-m})^n)^m \\ &\leqslant (nq^{k-m})^m \\ &\leqslant n^{-\epsilon m} q^{-m^2} \\ &= o(n^{-r}), \end{aligned} \tag{4·3}$$

since $(1-x)^n \geqslant 1-nx$ for $0 \leqslant x \leqslant 1$. By the Borel–Cantelli lemma

$$P\left(\limsup_{n\to\infty} \frac{\sigma_n}{\log n} \leqslant \log 1/q\right) = 1.$$

Next, we show that the reverse inequality holds. Let $\epsilon > 0$ and

$$d = d(n) = \left[\frac{(1-\epsilon)}{\log 1/q}\log n\right].$$

Then

$$\begin{aligned} P(\sigma_n < d(n)) &= P(\rho_d > n) \\ &\leqslant P(\delta_i > n/d, \quad \text{for some} \quad i = 0, 1, \ldots, d-1) \\ &\leqslant \sum_{i=0}^{d-1} P(\delta_i > n/d) \\ &= \sum_{i=0}^{d-1} (1-q^i)^{[n/d]} \\ &\leqslant d(1-q^d)^{[n/d]} \\ &\leqslant d \exp\left(-\frac{n^\epsilon}{d} + o(1)\right), \end{aligned} \tag{4·4}$$

since $1-x \leqslant e^{-x}$. Once again we use the Borel–Cantelli lemma and deduce that

$$P\left(\liminf_{n\to\infty} \frac{\sigma_n}{\log n} \geqslant \log 1/q\right) = 1,$$

which completes the proof of almost sure convergence.

It is easy to show convergence in any mean by noting that $0 < \sigma_n \leqslant n$ and using (4·3).

Proof of Theorem 8. We show first that

$$P\left(\liminf_{n\to\infty}\left(\Gamma_n \frac{\log n}{n}\right) \geqslant \log 1/q\right) = 1. \tag{4·5}$$

Let σ_n^i $(i = 1, 2, \ldots, n)$ be the number of vertices of ω_n which are coloured c_i. Then $\sigma_n^1 = \sigma_n$. Let $\epsilon > 0$ and

$$d = d(n) = \left[(1-\epsilon)\frac{n \log 1/q}{\log n}\right].$$

Then

$$\begin{aligned} P(\Gamma_n \leqslant d(n)) &= P\left(\sum_{i=1}^{d} \sigma_n^i = n\right) \\ &\leqslant P(\sigma_n^i \geqslant n/d, \quad \text{for some} \quad i = 1, 2, \ldots, d) \\ &\leqslant \sum_{i=1}^{d} P(\sigma_n^i \geqslant n/d) \\ &\leqslant dP(\sigma_n \geqslant n/d) \\ &= o(n^{-r}) \end{aligned}$$

for any positive integer r by (4·3), and (4·5) follows from the Borel–Cantelli lemma.

To prove convergence in mean, it is now sufficient to show that

$$\limsup_{n\to\infty}\left(\gamma_n\frac{\log n}{n}\right) \leqslant \log 1/q,$$

where $\gamma_n = E\Gamma_n$, because $\Gamma_n \geqslant 0$ and (4·5) holds.

Let $\epsilon > 0$, and let N be a positive integer such that, whenever $n \geqslant N$, we have

$$\beta_n = E\sigma_n \geqslant \frac{1}{(1+\epsilon)\log 1/q}\,\frac{(\log n)^2}{(\log n - 1)}. \tag{4·6}$$

Obviously $$\gamma_n \leqslant (1+\epsilon)\frac{n\log 1/q}{\log n} + N \tag{4·7}$$

for all $n \leqslant N$. Suppose that $m > N$ and (4·7) holds for all $n < m$. Then

$$\begin{aligned}\gamma_m &= 1 + \sum_{i=1}^{m}\gamma_{m-i}P(\sigma_m = i)\\ &\leqslant 1 + N + (1+\epsilon)\log 1/q\frac{(m-\beta_m)}{\log(m-\beta_m)}\end{aligned}$$

by the induction hypothesis and the concavity of $n/\log n$. This latter condition also implies that, for $1 \leqslant x \leqslant n-2$,

$$\frac{n}{\log n} - \frac{(n-x)}{\log(n-x)} \geqslant \frac{x(\log n - 1)}{(\log n)^2}.$$

Hence $$\begin{aligned}\gamma_m &\leqslant (1+\epsilon)\frac{m\log 1/q}{\log m} + N + 1 - \beta_m(1+\epsilon)\frac{(\log m - 1)}{(\log m)^2}\log 1/q\\ &\leqslant (1+\epsilon)\frac{m\log 1/q}{\log m} + N\end{aligned}$$

by (4·6), and the induction step is completed. Therefore (4·7) holds for all n, and

$$\limsup_{n\to\infty}\left(\gamma_n\frac{\log n}{n}\right) \leqslant 1+\epsilon$$

for any $\epsilon > 0$. The result of Theorem 8 follows.

Conclusion. There remain many open problems in this field. First, there are two questions which arise directly from this paper; the more important of these is to establish the truth or falsity of the conjecture in section 3. Also, we have been unable to prove any stronger form of convergence in Theorem 8.

There appears to be interesting work to be done in certain areas related to those which have concerned us. The theorems of this paper are all of an asymptotic nature, and are of limited practical use. For example, in section 3 we assert that n/I_n is a better lower bound than K_n for the chromatic number of ω_n, *when n is large.* Apart from one exceptional graph, this is never true for $n \leqslant 6$, and Table V of (4) suggests that it is false for n less than 100 when $p = \frac{1}{2}$. Our results would be of more use if they included some measure of the rates of convergence of the random variables in question. Finally, to what degree is it possible to describe the limiting distributions of these random variables when they are normalized by suitable functions? In general, this would seem to be very difficult, although we are able to do it (see (9)) for the variables $\{\sigma_n\}$.

REFERENCES

(1) ERDÖS, P. and RÉNYI, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* 5A (1960), 17–61.

(2) HARARY, F. *Graph theory* (1969) Addison–Wesley.

(3) HOLGATE, P. Majorants of the chromatic number of a random graph. *J. Roy. Statist. Soc. Ser. B* **31** (1969), 303–309.

(4) MATULA, D. W. On the complete subgraphs of a random graph. *Proc. 2nd. Conf. Comb. Th. and Appl.* (1970) Chapel Hill, 356–369.

(5) MATULA, D. W., MARBLE, G. and ISAACSON, J. D. Graph colouring algorithms, in *Graph theory and computing* (1972) Academic Press, London, 109–122.

(6) WELSH, D. J. A. and POWELL, M. B. An upper bound for the chromatic number of a graph and its application to timetabling problems. *Comput. J.* **10** (1967), 85–86.

(7) WILF, H. S. The eigenvalues of a graph and its chromatic number. *J. London Math. Soc.* **42** (1967), 330-332.

(8) WILSON, R. J. *Introduction to graph theory* (1972) Oliver and Boyd, Edinburgh.

(9) GRIMMETT, G .R. Random Graph Theorems. *Proc. 7th Prague Conf. on Information Theory and related topics* (1974), in press.