Can. J. Math., Vol. XXXIII, No. 2, 1981, pp. 476–484

COUNTING COLOURED GRAPHS OF HIGH CONNECTIVITY

BÉLA BOLLOBÁS

1. Introduction. Find exact or asymptotic formulae for the number of labelled graphs of order n having a certain property. The property we are interested in in this note is that of being k-coloured and having connectivity at least l. Special cases of this problem have been tackled by many authors; in particular Gilbert [**6**], Read [**9**] and Robinson [**11**] found exact formulae, and Read and Wright [**10**], and Wright [**12**], [**13**] found asymptotic expressions (for many other examples see [**7**]). Recently Harary and Robinson [**8**] counted labelled bipartite blocks, that is 2-connected bipartite graphs. (For terms not defined here and general background in graph theory see [**1**].) Our present investigations have been prompted by [**8**]; in particular, as a very special case of our results, we shall prove the conjecture published in [**8**].

The exact formulae appearing in the enumeration of labelled graphs in general, and in the enumeration of k-coloured labelled graphs in particular, tend to be very pleasing, especially because of the functional equations relating them. However, these exact formulae are not very transparent, and they are not readily amenable to investigations of the order of magnitude. The formulae are such that even with the aid of computers it is difficult to calculate the values beyond the first few cases, and though the initial values may enable one to spot valid relations (as in [**8**]), these relations are not easily proved if we want to rely on the exact formulae. The main aim of this note is to advocate the use of random graphs in order to find the asymptotic values we seek without any need for the exact formulae. (Random graphs have been applied with success in many instances, see [**2**, Chapter VII] for several examples.) Our main results, concerning the connectivity of most k-coloured graphs, are essentially best possible.

When looking for the proportion of graphs with certain properties, it is natural to turn the set of all graphs in question into a probability space, so that the proportion we seek appears as the probability of the set of graphs with that property.

All graphs considered in this note are assumed to be labelled. Furthermore, we are mainly interested in the set $(n_1, n_2, \ldots, n_k)$ of all k-coloured graphs with fixed colour classes $V_1, V_2, \ldots, V_k, |V_i| = n_i$, where

Received November 21, 1979.

$n_1 \leqq n_2 \leqq \ldots n_k \leqq cn_1$ for some constant c. (The slightly artificial restriction $n_k \leqq cn_1$ is explained by the fact that the proportion of k-coloured graphs on n vertices that do not satisfy this condition is $O(c^{-cn^2})$, where $c' > 0$ is a constant.) Thus every edge of these graphs joins vertices in distinct classes. Write $n = \sum_1^k n_i$ for the number of vertices and $e = \sum_{i<j} n_i n_j$ for the number of possible edges, that is for the cardinality of the set E of edges of the complete k-partite graph with colour classes $V_1, V_2, \ldots, V_k$. Then clearly

$$|\mathscr{G}(n_1, n_2, \ldots, n_k)| = 2^e,$$

so in order to give all graphs the same probability, the probability of every graph $G \in \mathscr{G}(n_1, n_2, \ldots, n_k)$ is 2^{-e}. Equivalently, two vertices of $V = \bigcup_1^k V_i$ belonging to different classes are joined by an edge with probability $\frac{1}{2}$, and they are not joined if they belong to the same class. In this way $\mathscr{G} = \mathscr{G}(n_1, n_2, \ldots, n_k)$ is turned into a very simple discrete probability space: the probability of a set $\mathscr{H} \subset \mathscr{G}$ is $|\mathscr{H}|/|\mathscr{G}|$.

We are interested mostly in the probability of a property for large values of n. For example, Wright [**12**] and Read and Wright [**10**] showed that if $\mathscr{H} \subset \mathscr{G}$ is the set of connected graphs then $|\mathscr{H}| \sim |\mathscr{G}|$, that is

$$P(\mathscr{H}) = |\mathscr{H}|/|\mathscr{G}| \to 1 \text{ as } n \to \infty.$$

As customary, we express this by saying that almost every (a.e.) graph in $\mathscr{G}$ is connected (see [**2**, Chapter VII]).

Throughout the note c denotes a positive constant, different at different occurrences and possibly dependent on other constants chosen earlier but not on n or $G \in \mathscr{G}$. The symbols $O(\ )$, $o(\ )$ and $\sim$ refer to the passage of n to infinity.

We shall make use of the classical De Moivre-Laplace formula (see [**5**, p. 172]). For a fixed p, $0 < p < 1$, and integers $0 \leqq L \leqq n$, denote by

$$S(n, L, p) = \sum_{l \leqq L} \binom{n}{l} p^l (1-p)^{n-l}$$

the probability that there are at most L successes in n Bernoulli trials with probability p of success. Set

$$L = pn - x(p(1-p)n)^{1/2}.$$

If $x = o(n^{1/6})$ and $x \to \infty$ then

$$(1) \quad S(n, L, p) \sim \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-1/2x^2},$$

and if $x \geqq \epsilon n^{1/2}$ then

$$(2) \quad S(n, L, p) \leqq e^{-cn}.$$

2. The connectivity of a.e. k-coloured graphs. Recall that $\mathcal{G} = \mathcal{G}(n_1, n_2, \ldots, n_k)$ is the probability space of all k-partite labelled graphs with vertex classes $V_1, V_2, \ldots, V_k$; $n_i = |V_i|$, $n_1 \leqq n_2 \leqq \ldots \leqq n_k \leqq cn_1$ and $n = \sum_1^k n_i$. To avoid the trivial case we take $k \geqq 2$. For $J \subset \{1, 2, \ldots, k\}$ put

$$V_J = \bigcup_{j \in J} V_j \quad \text{and} \quad n_J = |V_J| = \sum_{j \in J} n_j.$$

Given a vertex $x \in G$, denote by $d_J(x)$ the number of neighbours of x in V_J and set

$$\delta(i, J) = \delta(i, J, G) = \min_{x \in V_i} d_J(x).$$

Finally, given natural numbers a, b and a positive number λ, define

$$f(a, b) = \tfrac{1}{2}b - (\tfrac{1}{2}b \log a)^{1/2} + \tfrac{1}{4}(\log \log a)(b/2 \log a)^{1/2}$$

and

$$L(a, b, \lambda) = \lfloor f(a, b) + c(b/2 \log a)^{1/2} \rfloor,$$

where

$$c = \tfrac{1}{2} \log \lambda + \tfrac{1}{4} \log (4\pi).$$

Our results are modelled on some of the results in [**3**], though here our job is easier.

THEOREM 1. (i) *Given $\lambda > O$, for $i \notin 9$ we have in $\mathcal{G}$*

$$P(\delta(i, J) > L(n_i, n_J, \lambda)) \sim e^{-\lambda}.$$

(ii) *Let $C(n) \to \infty$ arbitrarily slowly. Then a.e. $G \in \mathcal{G}$ is such that*

$$|\delta(i, J) - f(n_i, n_J)| \leqq C(n)(n/\log n)^{1/2}.$$

(iii) *Let r be a natural number and let $\epsilon > 0$. Then a.e. graph $G \in \mathcal{G}$ is such that any r distinct vertices not in V_i are joined to at least $m = \lceil (1 - 2^{-r} - \epsilon) n_i \rceil$ vertices in V_i.*

Proof. (i) For simplicity put $a = n_i$, $b = n_J$ and

$$L = L(n_i, n_J, \lambda) = L(a, b, \lambda).$$

Then

$$L = \tfrac{1}{2}b - \tfrac{1}{2}b^{1/2}(2 \log a)^{1/2}\left\{1 - \frac{\log \log a}{4 \log a} - \frac{c}{\log a} + o\left(\frac{1}{\log a}\right)\right\},$$

so from (1) we see that

$$\log\{aS(b, L, \tfrac{1}{2})\} = o(1) + \log a - \tfrac{1}{2} \log (2\pi) - \tfrac{1}{2} \log \log a - \tfrac{1}{2} \log 2$$
$$- \log a\left\{1 - \frac{\log \log a}{2 \log a} - \frac{2c}{\log a}\right\} = 2c - \tfrac{1}{2} \log (4\pi) + o(1)$$
$$= \log \lambda + o(1).$$

https://doi.org/10.4153/CJM-1981-041-2 Published online by Cambridge University Press

Consequently

(3) $aS(b, L, \frac{1}{2}) \to \lambda.$

Let $V_i = \{x_1, x_2, \ldots, x_a\}$ and let $X_1, X_2, \ldots, X_a$ be the random variables defined by

$$X_s = \begin{cases} 1 \text{ if } d_J(x_s) \leqq L, \\ 0 \text{ otherwise.} \end{cases}$$

Then $Y = \sum_1^a X_s$ is the number of vertices in V_i joined to at most L vertices in V_J, and so $Y = 0$ if and only if $\delta(i, J) > L$. Clearly

$$p = P(X_s = 1) = \sum_{l \leqq L} \binom{b}{l} 2^{-b} = S(b, L, \tfrac{1}{2})$$

and Y is the number of successes in a Bernoulli trials with probability p of success. Since $ap \to \lambda$, the distribution of Y tends to the Poisson distribution with mean λ (see [**5**, Chapter VI]). In particular

$$P(Y = 0) = P(\delta(i, J) > L) \to e^{-\lambda},$$

as claimed.

(ii) This is an immediate consequence of (i).

(iii) Fix r vertices outside V_i. The probability that a vertex of V_i is joined to none of these r vertices is 2^{-r}. Hence by (2) the probability that at least $n_i - m$ vertices of V_i are joined to none of these r vertices is

$$S(n_i, m, 1 - 2^{-r}) \leqq e^{-cn}$$

where $c = c(r, \epsilon) > 0$. Consequently, the probability that some r vertices outside V_i are joined to less than m vertices in V_i is at most

$$\binom{n}{r} e^{-cn} \to 0.$$

COROLLARY 2. *Let $C(n) \to \infty$ arbitrarily slowly. Then the minimum degree $\delta(G)$ of a.e. $G \in \mathcal{G}$ satisfies*

$$\left| \delta(G) - f\left(n_k, \sum_1^{k-1} n_i\right) \right| \leqq C(n) \left(\frac{n}{\log n} \right)^{1/2}.$$

For convenient reference we note two immediate consequences of Theorem 1 (iii).

COROLLARY 3. *Almost every graph $G \in \mathcal{G}$ is such that*

(i) *for any two vertices, $x, y \in V$ there are at least $\frac{1}{2}\sum_1^{k-1} n_i + \frac{1}{8} n_1$ vertices joined to at least one of them,*

(ii) *whenever $x, y, z \notin V_i$, at most $\frac{1}{7} n_i$ vertices in V_i are joined to none of x, y and z.*

https://doi.org/10.4153/CJM-1981-041-2 Published online by Cambridge University Press

Equipped with these properties of almost all graphs, we can easily prove our main result about connectivity. For simplicity, put

$$m = \sum_{1}^{k-1} n_i.$$

THEOREM 4. *Let* $C(n) \to \infty$ *arbitrarily slowly. Then the connectivity* $\kappa(G)$ *of a.e.* $G \in \mathcal{G}$ *satisfies*

$$\kappa(G) = \delta(G) = \tfrac{1}{2}m - (\tfrac{1}{2}m \log n_k)^{1/2} + \tfrac{1}{4} (\log \log n_k)(m/\log n_k)^{1/2} + O\left(C(n)(n_k/\log n_k)^{1/2}\right).$$

Proof. Let $\mathcal{H}$ be the set of $G \in \mathcal{G}$ satisfying the conditions in Corollary 2, Corollary 3(i) and Corollary 3(ii). Then a.e. graph belongs to $\mathcal{H}$. In order to prove the theorem we show that every $G \in \mathcal{H}$ satisfies $\kappa(G) = \delta(G)$, provided $n_1 \geqq 16k$ and $\delta(G) < \frac{1}{2}m$. Note that these two inequalities do hold if n is sufficiently large and $G \in \mathcal{H}$. Suppose $G \in \mathcal{H}$, $\kappa(G) < \delta(G) < \frac{1}{2}m$ and $n_1 \geqq 16k$. Let S be a separating set with $\kappa(G)$ vertices. Then V can be partitioned as $V = S \cup W_1 \cup W_2$ such that $2 \leqq |W_1| \leqq |W_2|$ and G contains no $W_1 - W_2$ edge.

Let $x, y \in W_1$. Then by Corollary 3(i) there are at least $\frac{1}{2}m + \frac{1}{8}n_1$ vertices joined to at least one of x and y. Hence

$$|W_1| > \tfrac{1}{2}m + \tfrac{1}{8}n_1 - \delta(G) \geqq \tfrac{1}{8}n_1 \geqq 2k,$$

so there are at least three vertices of W_1 in one of the classes $V_1, \ldots, V_k$, say in V_i. (Of course i depends on G.) By Corollary 3(ii) $W_1 \cup S$ has at least $\frac{6}{7}n_j$ vertices in V_j for each $j \neq i$. One of the classes $V_j, j \neq i$, has to contain at least three vertices of W_1, so, again by Corollary 3(ii), $W_1 \cup S$ has at least $\frac{6}{7}n_i$ vertices in V_i. Consequently

$$|W_1 \cup S| \geqq \frac{6}{7} \sum_{1}^{k} n_i = \frac{6}{7} n.$$

However, this cannot hold since it implies

$$|S| = |W_1 \cup S| + |W_2 \cup S| - n \geqq \frac{5}{7} n > \delta(G) > \kappa(G),$$

contradicting our assumptions.

Let us formulate an immediate consequence of Theorem 4.

COROLLARY 5. *Denote by* $M(n_1, n_2, \ldots, n_k; \kappa \geqq s)$ *the number of s-connected graphs in* $\mathcal{G}(n_1, n_2, \ldots, n_k)$. *Then for every* $\alpha < \frac{1}{2}$

$$M(n_1, \ldots, n_k; \kappa \geqq \lceil \alpha m \rceil) \sim M(n_1, \ldots, n_k; \kappa \geqq 0) = 2^e,$$

where

$$m = \sum_{1}^{k-1} n_i \quad \textit{and} \quad e = \sum_{i<j} n_i n_j.$$

3. Graphs of low connectivity. The results so far tell us what the connectivity of most graphs is. However, they do not give any information about the number of exceptional graphs, that is those with relatively low connectivity. Of course, it would not be difficult to strengthen the proofs so that they give us fairly good bounds, but even more striking results can be obtained if we restrict our attention to finite connectivity (that is not tending to infinity with n). As before, we consider the space

$$\mathscr{G} = \mathscr{G}(n_1, n_2, \ldots, n_k), \quad n_1 \leqq n_2 \leqq \ldots \leqq n_k \leqq cn_1, \quad n = \sum_{1}^{k} n_i.$$

THEOREM 6. *Let s be a fixed natural number. Then the probability that the connectivity $\kappa(G)$ of a graph $G \in \mathscr{G}$ is not greater than s is*

$$P(\kappa(G) \leqq s) = O(e^{-cn}).$$

Proof. In order to simplify the notation, we treat the case $k = 2$, $n_1 = n_2 = m$. It is clear that the arguments can be carried over to the general case; in fact the case $k = 2$ easily implies the assertion for $k > 2$.

As in the proof of Theorem 4, given a graph $G \in \mathscr{G}(m, m)$ with $\kappa(G) \leqq s$, choose a set S of s vertices separating G. Choose one of the smallest components of $G - S$, say H. Suppose H has t_1 vertices in V_1 and t_2 in V_2, and $t_1 \geqq t_2$. Then the t_1 vertices of H in V_1 are joined to at most $\min\{s + t_1, s + m/2\}$ vertices in V_2. Thus if we denote by $P(t, u)$ the probability that some t vertices in V_1 are joined to at most u vertices in V_2 then

$$P(t, u) \leqq \binom{m}{t} \sum_{k \leqq u} \binom{m}{k} (1 - 2^{-t})^k 2^{-t(m-k)}$$

and

$$(4) \quad P(\kappa(G) \leqq s) \leqq 2 \sum_{t=1}^{m} P(t, \min\{t + s, s + m/2\}).$$

Clearly

$$P(1, s + 1) \leqq m2^{-m} \sum_{k \leqq s+1} \binom{m}{k} = \frac{m^{s+2}}{(s + 1)!} 2^{-m}(1 + o(1)),$$

$$P(2, s + 2) \leqq \binom{m}{2} \sum_{k \leqq s+2} \binom{m}{k} 2^{-2(m-k)} = o(2^{-m})$$

and if $3 \leqq t$ and $u \leqq 4m/7$ then

$$P(t, u) \leqq \binom{m}{t} \sum_{k \leqq 4m/7} \binom{m}{k} 2^{-3tm/7} \leqq 2^{-m}$$

if m is greater than an absolute constant. Now if m is sufficiently large, then $s + m/2 < 4m/7$, so from (4) we find that

$$P(\kappa(G) \leqq s) \leqq \frac{m^{s+2}}{(s+1)!} 2^{-m+1}(1 + o(1)).$$

The same method can be applied with fair success to estimate the proportion of graphs of low connectivity even for graphs of low order. For example, for $s = 1$ and $m = 12$ the computer-assisted table found by Harary and Robinson [**8**] tells us that about 16/223 of the 2^{144} graphs are not 2-connected, and a brief calculation on the lines above gives the upper bound 1/12.

4. Varying and interchangeable colour classes. Let us see now how the results about the proportion of graphs with relatively low connectivity can be extended to larger sets of k-coloured graphs. Denote by $\mathscr{G}(n, k\text{-col})$ the set of all k-coloured graphs on n distinguishable vertices. Thus each $G \in \mathscr{G}(n, k\text{-col})$ carries its colouring: a subset V_1 is coloured 1, a subset is coloured 2, etc. In particular, with each $G \in \mathscr{G}(n, k\text{-col})$ there is associated a k-tuple $(n_1, n_2, \ldots, n_k)$, $n_i \geqq O$ and $\sum_1^k n_i = n$. Following [**9**] we write

$$M_n(k) = |\mathscr{G}(n, k\text{-col})|.$$

Clearly (see [**9**])

$$M_n(k) = \sum_{(n)} \frac{n!}{n_1! \ldots n_k!} 2^e,$$

where $\sum_{(n)}$ denotes summation over all sets of non-negative integers n_i such that $\sum_1^k n_i = n_1$ and e denotes the sum $\sum_{i<j} n_i n_j$. The following rather crude estimate will enable us to transfer with great ease our results from $\mathscr{G}(n_1, \ldots, n_k)$ to $\mathscr{G}(n, k\text{-col})$.

THEOREM 7. *Given $\epsilon > O$, denote by $N_n(k)$ the number of graphs in $\mathscr{G}(n, k\text{-col})$ having a colour class of size at most $(1 - \epsilon)n/k$. Then*

$$Nn(k)/M_n(k) = O(e^{-cn^2}).$$

Proof. Let $n_1^*, \ldots, n_k^*$ be defined by

$$n_1^* \leqq n_2^* \leqq \ldots \leqq n_k^* \leqq n_1^* + 1 \quad \text{and} \quad \sum_1^k n_i^* = n.$$

Trivially

$$M_n(k) > |\mathscr{G}(n_1^*, \ldots, n_k^*)| = 2^{e^*},$$

where $e^* = e(T_n(k)) = \sum_{i<j} n_i^* n_j^*$ is the number of edges in the Turán graph with k classes (see [**2**, p. 71]). Now if the non-negative integers n_i

are such that $\sum_1^k n_i = n$ and $\min_i n_i \leqq (1 - \epsilon)n/k$ then

$$e = \sum_{i<j} n_i n_j \leqq e^* - c'n^2 \quad \text{for some } c' > 0.$$

Consequently

$$Nn(k)/(M_n(k) \leqq n^{2n}2^{e^*-c'n^2}2^{-e^*} = O(e^{-cn^2}).$$

This simple result says in precise terms that most graphs in $\mathscr{G}(n, k\text{-col})$ have almost equal colour classes: that is our somewhat artificial restriction in §§ 2 and 3 holds for most graphs. Consequently Corollary 5 and Theorem 6 have the following consequences. Denote by $C_n(k, \kappa \leqq s)$ the number of graphs in $\mathscr{G}(n, k\text{-col})$ with connectivity at most s.

COROLLARY 8. *For $\epsilon > 0$*

$$C_n\left(k, \kappa \leqq (1 - \epsilon)\frac{k-1}{2k} n\right)/M_n(k) \to 0.$$

COROLLARY 9. *Let s be a fixed natural number. Then*

$$C_n(k, \kappa \leqq s) = O(e^{-cn}).$$

The special case $k = 2$, $s = 1$ of Corollary 9 is a sharp form of the stronger conjecture in [**8**].

By now it must be clear that Corollaries 8 and 9 hold for k-colourable and k-chromatic graphs as well, that is essentially the same results hold (with very similar proofs) for k-coloured graphs with not only varying but also interchangeable colour classes. For example, as in the proof of Theorem 7, we find that only a proportion of $O(e^{-cn^2})$ of k-colourable graphs are also $(k\text{-}1)$-colourable, in fact, almost all k-colourable graphs are uniquely k-colourable with roughly equal colour classes and about $k - 1/2k\, n^2$ edges (see [**4**] for some sharper results in this direction). These straightforward extensions are left to the reader.

REFERENCES

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

2. B. Bollobás, *Graph Theory—An introductory course*, Grad. Texts in Math. *63* (Springer-Verlag, New York and Heidelberg, 1979).

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

4. B. Bollobás and N. Sauer, *Uniquely colourable graphs wkth large girth*, Can. J. Math. *28* (1976), 1340–1344.

5. W. Feller, *An introduction to probability theory and its applications* (Wiley and Chapman, New York and London, 1957).

6. E. N. Gilbert, *Enumeration of labelled graphs*, Can. J. Math. *8* (1956), 405–411.

7. F. Harary and E. M. Palmer, *Graphical enumeration* (Academic Press, New York, 1973).

8. F. Harary and W. Robinson, *Labelled bipartite blocks*, Can. J. Math. *31* (1979), 60–68.

9. R. C. Read, *The number of coloured graphs on labelled nodes*, Can. J. Math. *12* (1960), 410–414.
10. R. C. Read and E. M. Wright, *Coloured graphs: A correction and extension*, Can. J. Math. *22* (1970), 594–596.
11. R. W. Robinson, *Enumeration of non-separable graphs*, J. Combinatorial Theory *9* (1970), 327–356.
12. E. M. Wright, *Counting coloured graphs*, Can. J. Math. *13* (1961), 683–693.
13. ——— *Counting coloured graphs II*, Can. J. Math. *16* (1964), 128–135.

University of Cambridge,
Cambridge, England