Europ. J. Combinatorics (1980) **1**, 311–316

A Probabilistic Proof of an Asymptotic Formula for the Number of Labelled Regular Graphs

BÉLA BOLLOBÁS

Let Δ and n be natural numbers such that $\Delta n = 2m$ is even and $\Delta \leqslant (2 \log n)^{\frac{1}{2}} - 1$. Then as $n \to \infty$, the number of labelled Δ-regular graphs on n vertices is asymptotic to

$$e^{-\lambda-\lambda^2} \frac{(2m)!}{m!2^m(\Delta!)^m},$$

where $\lambda = (\Delta - 1)/2$. As a consequence of the method we determine the asymptotic distribution of the number of short cycles in graphs with a given degree sequence, and give analogous formulae for hypergraphs.

In 1959 Read [6] determined an exact formula for the number of labelled Δ–regular graphs on n vertices. This formula, whose proof is based on Pólya's enumeration theorem [5], is not easily penetrated. In particular, it seems that only for $\Delta \leqslant 3$ can it be used to find its asymptotic value (see [4, p. 175]). Recently Bender and Canfield [1] gave an asymptotic formula for the number of labelled graphs with given degree sequences by enumerating certain classes of involutions. In this note we offer a somewhat different approach, allowing one to obtain a more general asymptotic formula without much effort and without any reference to an exact formula. In particular, our asymptotic formula holds not only for constant Δ but also if Δ increases rather slowly as $n \to \infty$. As a considerable bonus, the model presented here enables one to give asymptotic formulae for various subclasses of labelled regular graphs. We intend to exploit this possibility in the future.

As customary, we use $A \sim B$ to denote the relation $A/B \to 1$ as $n \to \infty$. Furthermore, we write $(a)_b = a(a-1) \cdots (a-b+1)$. Throughout the proof $c_1, c_2, \ldots$ denote positive constants.

THEOREM 1. *Let $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_n$ be natural numbers with $\sum_{i=1}^n d_i = 2m$ even. Suppose*

$$\Delta = d_1 \leqslant (2 \log n)^{\frac{1}{2}} - 1$$

and $m \geqslant \max\{\varepsilon \Delta n, (1+\varepsilon)n\}$ for some fixed $\varepsilon > 0$. Then the number $L(\mathbf{d})$ of labelled graphs with degree sequence $\mathbf{d} = (d_i)_1^n$ satisfies

$$L(\mathbf{d}) \sim e^{-\lambda-\lambda^2}(2m)_m \Big/ \Big\{ 2^m \prod_{i=1}^n d_i! \Big\},$$

where

$$\lambda = \frac{1}{2m} \sum_{i=1}^n \binom{d_i}{2}.$$

PROOF. We shall represent our graphs as images of so called "configurations". Let $W = \bigcup_{j=1}^n W_j$ be a fixed set of $2m = \sum_{j=1}^n d_j$ labelled vertices, where $|W_j| = d_j$. A *configuration* F is a partition of W into m pairs of vertices, called *edges* of F. Clearly there are

$$N = N(m) = \binom{2m}{2}\binom{2m-2}{2} \cdots \binom{2}{2} \Big/ m! = (2m)_m 2^{-m} \tag{1}$$

configurations. Furthermore, if we fix l independent (vertex disjoint) edges then there are

0195-6698/80/040311+06 $01.00/0

exactly

$$N_l(m)=\binom{2m-2l}{2}\binom{2m-2l-2}{2}\cdots\binom{2}{2}\Big/(m-l)!=\frac{N}{(2m-1)(2m-3)\cdots(2m-2l+1)}$$

configurations containing l edges. Consequently if $s, t\le 8\log n$ then

$$1\le (N_{s+2t}/N)(N/N_1)^s(N/N_2)^t\le 1+c_1\frac{(\log n)^2}{n}. \tag{2}$$

A *k-cycle* of a configuration is a set of k edges, say $\{e_1, e_2, \ldots, e_k\}$ such that for some k distinct groups $W_{j_1}, \ldots, W_{j_k}$ the edge e_i joins W_{j_i} to $W_{j_{i+1}}$, where $W_{j_{k+1}}\equiv W_{j_1}$. We shall call a 1-cycle a *loop* and a 2-cycle a *coupling*. For $\sigma\subset V$ put $w(\sigma)=\prod_{i\in\sigma} d_i(d_i-1)$ and set

$$C_k(\mathbf{d})=\tfrac{1}{2}(k-1)!\sum_{|\sigma|=k} w(\sigma). \tag{3}$$

Clearly there are exactly $C_k(\mathbf{d})$ k-sets of pairs of vertices that can be k-cycles of configurations.

We shall need that if the sequence $\mathbf{d}$ is decreased a little then $C_k(\mathbf{d})$ does not decrease by much. More precisely, given a non-negative integer q, denote by $\mathbf{d}-q$ the sequence d_{q+1}, $d_{q+2}, \ldots, d_n$ and define $C_k(\mathbf{d}-q)$ by (3). Clearly

$$C_1(\mathbf{d})-C_1(\mathbf{d}-q)\le q\binom{\Delta}{2}$$

and

$$C_2(\mathbf{d})-C_2(\mathbf{d}-q)\le \binom{q}{2}\binom{\Delta}{2}^2+q\Delta^2C_1(\mathbf{d}).$$

From these it follows that if $q\le 8\log n$, say, then

$$(C_1(\mathbf{d}-q)/C_1(\mathbf{d}))^q(C_2(\mathbf{d}-q)/C_2(\mathbf{d}))^q\ge 1-c_2\frac{(\log n)^2}{n}. \tag{4}$$

Finally, we define a *shackle* of a configuration as a pair of loops in the same group W_j or a pair of couplings joining the same two groups. (Note that in the latter case all we need is a set of three edges joining the same two groups.) Let $e_1, e_2, \ldots, e_l$ be independent edges not containing a shackle, where $l\ge 0$. Denote by $N^*(e_1, \ldots, e_l)$ the number of configurations that contain the edges $e_1, \ldots, e_l$ and have a shackle. Then

$$N^*(e_1,\ldots,e_l)\le l(\Delta-1)^2N_{l+1}(m)+n\binom{\Delta}{2}^2N_{l+2}(m)+\binom{n}{2}\binom{\Delta}{3}^2 3!N_{l+3}(m). \tag{5}$$

Indeed, there are at most $l(\Delta-1)^2$ edges that can appear in a shackle containing some edge e_i, $1\le i\le l$, there are at most $n\binom{\Delta}{2}^2$ choices of a new "double loop" and at most $\binom{n}{2}\binom{\Delta}{3}^2 3!$ choices of a new "triple edge". Denote by N_l^* the right-hand side of (5). Note that if $l\le 16\log n$ then, rather crudely,

$$N_l^*/N_l\le c_3\frac{(\log n)^2}{n}. \tag{6}$$

Now we are ready to proceed to the essential part of the proof. Let Φ be the set of all configurations and let $\Phi_0\subset\Phi$ be the *set of configurations without shackles.* Put $M_0=M_0(m)=|\Phi_0|$. Then by (5) and (6) we have

$$N\left(1-c_3\frac{(\log n)^2}{n}\right)\le N-N_0^*\le N-N^*(\varnothing)=M_0\le N. \tag{7}$$

Let us turn Φ_0 into a probability space by giving each configuration $F \in \Phi_0$ the same probability, M_0^{-1}. Given a configuration $F \in \Phi_0$, denote by $X_1(F)$ the *number of loops* of F and by $X_2(F)$ the *number of couplings*. Put $X = X_1 + X_2$. Our aim is to determine the asymptotic value of $P(X=0) = P(X_1+X_2=0)$. We shall do this by estimating $E(s, t)$, the expected number of $(s+t)$-tuples consisting of s loops and t couplings, provided $s+t \leq 8 \log n$.

Note that there are $C_1(\mathbf{d})$ edges that can be loops of configurations and $C_2(\mathbf{d})$ pairs of edges that can form couplings. Since in Φ_0 the s loops and t couplings determine $s+2t$ edges, we find that

$$E(s, t) \leq \binom{C_1(\mathbf{d})}{s}\binom{C_2(\mathbf{d})}{t} N_{s+2t}/M_0.$$

Therefore if we write $\lambda_1 = C_1(\mathbf{d})N_1/N$ and $\lambda_2 = C_2(\mathbf{d})N_2/N$ then (2) and (7) give

$$E(s, t) \leq \frac{\lambda_1^s \lambda_2^t}{s!t!}\left(1 + c_4 \frac{(\log n)^2}{n}\right). \tag{8}$$

To get a lower bound, we count only those $(s+t)$-tuples of loops and couplings whose endvertices belong to distinct groups W_j:

$$E(s, t) \geq \binom{C_1(\mathbf{d}-s-2t)}{s}\binom{C_2(\mathbf{d}-s-2t)}{t}(N_{s+2t} - N^*_{s+2t})/M_0.$$

Hence (4), (6) and (7) give

$$E(s, t) \geq \frac{\lambda_1^s \lambda_2^t}{s!t!}\left(1 - c_5 \frac{(\log n)^2}{n}\right). \tag{9}$$

From (8) and (9) we find that if $s+t \leq 8 \log n$ then

$$\left|E(s, t) - \frac{\lambda_1^s \lambda_2^t}{s!t!}\right| \leq \frac{\lambda_1^s \lambda_2^t}{s!t!} c_6 \frac{(\log n)^2}{n}.$$

This shows that X_1 and X_2 behave like independent Poisson random variables with means λ_1 and λ_2. Straightforward calculations show that

$$\lambda_1 + \lambda_2 = \frac{1}{2(2m-1)}\sum_1^n d_i(d_i-1) + \frac{1}{2(2m-1)(2m-3)}\sum_{i<j} d_i(d_i-1)d_j(d_j-1) < \frac{\Delta^2}{4} \tag{10}$$

and in fact $\lambda_1 - \lambda \to 0$ and $\lambda_2 - \lambda^2 \to 0$.

Put $E_r = E(\binom{X}{r})$, i.e. denote by E_r the expected number of r-tuples of loops and couplings. Clearly

$$E_r = \sum_{s=0}^r E(s, r-s)$$

and

$$\sum_{s=0}^r \frac{\lambda_1^s \lambda_2^{r-s}}{s!(r-s)!} = \frac{(\lambda_1+\lambda_2)^r}{r!}.$$

Consequently for $r \leq 8 \log n$ we have

$$\left|E_r - \frac{(\lambda_1+\lambda_2)^r}{r!}\right| \leq \frac{(\lambda_1+\lambda_2)^r}{r!} c_6 \frac{(\log n)^2}{n}. \tag{11}$$

As is well known, the values E_r are closely related to $P(X=0)$. By the Jordan inequalities (see Comtet [3, p. 195]) if u is a natural number then

$$\sum_{r=0}^{2u+1}(-1)^r E_r \le P(X=0) \le \sum_{r=0}^{2u}(-1)^r E_r. \tag{12}$$

Let us choose u so that $2\log n < u < 3\log n$. Then from (10), (11) and (12) we find that

$$|e^{\lambda_1+\lambda_2}P(X=0)-1| \le e^{\lambda_1+\lambda_2}\left\{\sum_{r=0}^{\infty}\frac{(\lambda_1+\lambda_2)^r}{r!}c_6\frac{(\log n)^2}{n}+2\frac{(\lambda_1+\lambda_2)^{2u}}{(2u)!}\right\}$$

$$\le c_6 e^{\Delta^2/2}\frac{(\log n)^2}{n}+o(1)=o(1).$$

Therefore

$$P(X=0)=P(X_1+X_2=0)\sim e^{-\lambda_1-\lambda_2}\sim e^{-\lambda-\lambda^2}, \tag{13}$$

since $\lambda_1-\lambda\to 0$ and $\lambda_2-\lambda^2\to 0$.

The rest of the proof is straightforward. Put

$$\Omega=\{F\in\Phi_0 : X(F)=0\}.$$

Then (13) states exactly that

$$|\Omega|\sim e^{-\lambda-\lambda^2}|\Phi_0|=e^{-\lambda-\lambda^2}M_0\sim e^{-\lambda-\lambda^2}N,$$

where the last step follows from (7). Let $\mathcal{G}$ be the set of all graphs with vertex set $V=\{W_1, W_2, \ldots, W_n\}$ in which vertex W_j has degree $|W_j|$. Then by definition $L(\mathbf{d})=|\mathcal{G}|$. Given $F\in\Omega$, define a graph $\phi(F)$ with vertex set V by selecting an edge (W_i, W_j) if and only if the configuration F contains an edge from W_i to W_j. Then $\phi(\Omega)=\mathcal{G}$ and for each $G\in\mathcal{G}$ we have $|\phi^{-1}(G)|=\prod_{i=1}^n d_i!$ since if $\phi(F)=G$ then $\phi^{-1}(G)$ consists of the images of F under permutations of W leaving each W_i invariant. Hence

$$L(\mathbf{d})=|\mathcal{G}|=|\Omega|\Big/\prod_{i=1}^n d_i!\sim e^{-\lambda-\lambda^2}(2m)_m\Big/\left\{2^m\prod_{i=1}^n d_i!\right\},$$

as claimed by the theorem.

On putting $d_1=d_2=\cdots=d_n=\Delta$ we find that $2m=\Delta n$ and

$$\lambda=\frac{1}{2m}n\binom{\Delta}{2}=\frac{\Delta-1}{2},$$

which implies the formula given in the abstract.

It is perhaps worth noting that the proof would have been considerably simpler if we had confined our attention to degree sequences in which $\Delta=d_1$ is bounded. To show this, we use the probabilistic model to determine the asymptotic distribution of the number of short cycles in labelled graphs with a given degree sequence.

THEOREM 2. *Let Δ be a fixed natural number and let $\Delta=d_1\ge d_2\cdots\ge d_n$ be such that $\sum_1^n d_i=2m$ is even and $2m-n\to\infty$ as $n\to\infty$. Consider the probability space $\mathcal{G}$ of all graphs with vertex set $\{1,2,\ldots,n\}$ in which the degree of vertex i is d_i. For $G\in\mathcal{G}$ and $i\ge 3$ denote by $X_i(G)$ the number of i-cycles in G. Then for any fixed k the random variables $X_3, X_4, \ldots, X_k$ are asymptotically independent Poisson random variables with X_i having mean*

$$\lambda_i=\frac{\lambda^i}{2i},\quad \text{where } \lambda=\frac{1}{m}\sum_1^n\binom{d_i}{2}.$$

PROOF. As in the proof of Theorem 1, consider the set Φ of all configurations. Define a *p-shackle* of a configuration $F \in \Phi$ as a set of $l+1 \leq p+1$ edges that join vertices of some l groups W_j. (Thus every shackle is a 2-shackle but not conversely.) Given a set $\{e_1, e_2, \ldots, e_q\}$ of independent edges not containing a p-shackle, denote by $N^*(e_1, e_2, \ldots, e_q)$ the number of configurations containing these edges and at least one p-shackle. Put $N_q^* = \max N^*(e_1, \ldots, e_q)$ where the maximum is over all q-sets of independent edges. Choose $p = p(n) \to \infty$ so that $N_q^* = o(N_q)$ for every fixed q; it is immediate that this can be done. Put $M_q = N_q - N_q^*$. Note that for every fixed i and q we have

$$C_i(\mathbf{d})N_i/N \sim C_i(\mathbf{d})M_i/M_0 \sim C_i(\mathbf{d})(M_1/M_0)^i \sim \lambda_i \tag{14}$$

and

$$C_i(\mathbf{d}-q) \sim C_i(\mathbf{d}). \tag{15}$$

And now for the actual proof. Let $\Phi_0 \subset \Phi$ be the set of configurations without shackles and let Ω be the set of configurations without loops and couplings. As before, we consider Φ, Φ_0 and Ω as probability spaces in which all configurations have the same probability. The map $\phi : \Omega \to \mathcal{G}$ shows that $X_3, X_4, \ldots$ have the same distribution over Ω as over $\mathcal{G}$. Our aim is to determine the asymptotic distribution over Ω. Note that $|\Phi| = N \sim M_0 = |\Phi_0|$. For $F \in \Phi$ let $X_k(F)$ be the number of k-cycles in F, and write $E_{\Phi_0}(t_1, t_2, \ldots, t_k)$ for the expectation (in Φ_0) of the number of t-tuples consisting of t_1 1-cycles, t_2 2-cycles, $\ldots$, t_l l-cycles, where $t = \sum_1^l t_i$. Now if $p = p(n) \geq t$ then these t cycles determine $q = \sum_1^l it_i$ edges so

$$E_{\Phi_0}(t_1, \ldots, t_l) \leq \prod_{i=1}^{l} \binom{C_i(\mathbf{d})}{t_i} N_q/M_0.$$

On the other hand, if we select the cycles one by one, for each i-cycle we have at least $C_i(\mathbf{d}-q)$ choices, so

$$E_{\Phi_0}(t_1, \ldots, t_l) \geq \prod_{i=1}^{l} \binom{C_i(\mathbf{d}-q)}{t_i} M_q/M_0.$$

Because of (14) and (15) these inequalities give

$$E_{\Phi_0}(t_1, \ldots, t_l) \sim \prod_{i=1}^{l} \lambda_i^{t_i}/t_i!.$$

Since the quantity on the right is the appropriate moment of a joint Poisson distribution, this implies (see e.g. Chung [2]) that for every fixed sequence $t_1, t_2, \ldots, t_l$ we have

$$P_{\Phi_0}(X_i = t_i, i = 1, \ldots, l) \sim e^{-\sum_1^l \lambda_i} \prod_{i=1}^{l} \lambda_i^{t_i} t_i!.$$

As $M_0 \sim N$, the same holds over Φ. Since $\Omega = \{F \in \Phi : X_1(F) = X_2(F) = 0\}$ it gives also

$$P_\Omega(X_i = t_i : i = 3, \ldots, l) \sim e^{-\sum_3^l \lambda_i} \prod_{i=3}^{l} \lambda_i^{t_i}/t_i!,$$

completing the proof.

Note that for regular graphs, that is $d_1 = d_2 = \cdots = d_n = \Delta$ and Δn even, we have $\lambda_i = (\Delta - 1)^2/2i$.

The method is easily adapted to find asymptotic formulae for various classes of hypergraphs. In conclusion we state a simple result of this kind.

THEOREM 3. *Let $r \geq 3$ and $\Delta \geq 2$ be fixed and let $\Delta = d_1 \geq d_2 \geq \cdots \geq d_n$ be such that $m = 1/r \sum_1^n d_i$ is an integer and $rm - n \to \infty$. Then the number of labelled r-graphs with degree sequence $(d_i)_1^n$ is asymptotic to*

$$e^{-\lambda} \frac{(rm)!}{m!(r!)^m} \bigg/ \left(\prod_{i=1}^n d_i! \right),$$

where

$$\lambda = \frac{r-1}{rm} \sum_1^n \binom{d_i}{2}.$$

Furthermore, the number of r-graphs in which no two hyperedges have more than one vertex in common is asymptotic to

$$e^{-\lambda-\lambda^2} \frac{(rm)!}{m!(r!)^m} \bigg/ \left\{ \prod_{i=1}^n d_i! \right\}.$$

We only give a brief sketch. Consider a set $W = \bigcup_1^n W_j$ of rm labelled vertices, where $|W_j| = d_j$, and define a configuration as a partition of W into m sets of r vertices. Let Φ be the probability space of all configurations. Define a loop of $F \in \Phi$ as a pair of vertices of the same group W_j contained in a hyperedge of F. Let $X_1(F)$ be the number of loops of F. Then, as in the proofs of the earlier theorems, one can show that the distribution of X_1 tends to the Poisson distribution with mean λ. Since

$$|\Phi| = \frac{(rm)!}{m!(r!)^m} \bigg/ \left\{ \prod_{i=1}^n d_i! \right\},$$

this gives the first assertion. The second is proved similarly.

REFERENCES

1. E. A. Bender and E. R. Canfield, The asymptotic number of labelled graphs with given degree sequences, *J. Combin. Theory Ser. A* **24** (1978), 296–307.
2. K. L. Chung, *A Course in Probability Theory*, 2nd ed., Academic Press, New York and London, 1974.
3. L. Comtet, *Advanced Combinatorics*, Reidel, Dordrecht and Boston, 1974.
4. F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, New York and London, 1973.
5. G. Pólya, Kombinatorische Anzahlbestimmungen für Gruppen und chemische Verbindungen, *Acta Math.* **68** (1937), 145–254.
6. R. C. Read, The enumeration of locally restricted graphs (I), *J. London Math. Soc.* **34** (1959), 417–436.

Received 18 August 1980

B. BOLLOBÁS
Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, U.K.