GRAPHS ON UNLABELLED NODES WITH A GIVEN NUMBER OF EDGES

BY

E. M. WRIGHT

University of Aberdeen, Aberdeen, U.K. (1)

1. Introduction

We write T_{nq} for the number of different graphs on n unlabelled nodes with just q edges. We shall find an asymptotic approximation to T_{nq} for large n and determine the exact range for q for which it holds good. In the graphs we consider, every pair of nodes is joined by just one undirected edge or not so joined, though our method can clearly be extended to other types of graph. If the nodes are labelled, there are N possible edges, where $N = n(n-1)/2$, and the number of graphs with just q edges is

$$F_{nq} = \binom{N}{q} = \frac{N!}{q!\,(N-q)!},$$

the number of ways of selecting q objects out of N.

All our statements carry the implied condition "for large enough n". The number q is subject to bounds depending on n. We use C for a positive number, not always the same at each occurrence, independent of n and q. The notations $O(\ )$ and $o(\ )$ refer to the passage of n to infinity and each of the constants implied is a C.

We shall prove

THEOREM 1. *The necessary and sufficient condition that*

$$T_{nq} \sim F_{nq}/n! \tag{1.1}$$

as $n \to \infty$ is that

$$\min(q, N-q)/n - (\log n)/2 \to \infty. \tag{1.2}$$

Pólya [2] proved (1.1) when $|2q-N| = O(n)$, though he appears never to have published his proof. Recently Oberschelp [4] proved (1.1) under the condition that $|2q-N| <$

(1) The research reported herein has been sponsored by the United States Government.

$0.84n^{3/2}$. My contribution here is to prove (1.1) under the wider condition (1.2), which is equivalent to

$$n\{\log n+2\psi(n)\}/2 \leqslant q \leqslant N-n\{\log n+2\psi(n)\}/2$$

where $\psi(n)\to\infty$ as $n\to\infty$, and to prove also that this condition is necessary as well as sufficient.

My proof requires one simple result from complex integration, viz. that, if m is an integer,

$$\int_{-\pi}^{\pi} e^{mit}\,dt=\begin{cases}2\pi & (m=0),\\ 0 & (m\neq 0).\end{cases} \tag{1.3}$$

Otherwise the proof is "elementary".

Most of the complications of my proof of Theorem 1 arise from the "best possible" nature of the result. The following theorem, a little weaker than Theorem 1, but a considerable advance on the previous results, can be proved much more simply. We require only § 2 and a simple variant of § 4 of the present paper.

THEOREM 2. *The necessary condition for* (1.1) *is* (1.2)*; a sufficient condition is that*

$$3n\log n \leqslant q \leqslant N-3n\log n. \tag{1.4}$$

We write S_n for the symmetric group of permutations of degree n, i.e. the group of all permutations ω of the n nodes labelled (say) by the numbers 1, 2, ..., n. The permutation ω has p_1 cycles of unit length, p_2 of length 2 and so on; it induces a permutation of the possible N edges joining each pair of these nodes. The latter permutation belongs to S_N and has P_1 cycles of unit length, P_2 of length 2 and so on. Then

$$p_1+2p_2+3p_3+\dots+np_n=n, \tag{1.5}$$

$$P_1+2P_2+3P_3+\dots+NP_N=N. \tag{1.6}$$

We write
$$G_\omega=G_\omega(X)=\prod_{j=1}^{N}(1+X^j)^{P_j}$$

and use $[G]_q$ to denote the coefficient of X^q in the polynomial $G=G(X)$. There is a famous theorem due to Pólya [5] which tells us [1, 3, 4, 5] that

$$n!\sum_{q=0}^{N}T_{nq}X^q=\sum_{\omega\in S_n}G_\omega(X),$$

so that
$$n!\,T_{nq}=\sum_{\omega\in S_n}[G_\omega]_q. \tag{1.7}$$

R_a is the set of those ω for which $p_1 = n - a$. We write $H_a = \sum [G_\omega]_q$, where the sum is over all ω in R_a, so that (1.7) takes the form

$$n!\, T_{nq} = \sum_{a=0}^{n} H_a. \tag{1.8}$$

If $a = 0$, ω is the identity I and $H_0 = [G_I]_q = F_{nq}$. There is no ω for which $a = 1$ and so the set R_1 is empty. If $a = 2$, ω is one of N permutations for each of which

$$p_1 = n-2,\ p_2 = 1,\ P_1 = N - 2n + 4,\ P_2 = n-2,\ p_3 = p_4 = \ldots = P_3 = P_4 = \ldots = 0$$

and

$$[G_\omega]_q = [(1+X)^{N-2n+4}(1+X^2)^{n-2}]_q = c_2 \tag{1.9}$$

(say). Hence $H_2 = Nc_2$.

If $\omega \in R_a$, the effect of ω is to change just a of the nodes and to leave the remaining $n-a$ unchanged. There are $n!/a!(n-a)!$ ways of choosing these a nodes. The effect of ω on the set of a nodes is isomorphic to one of the permutations of S_a, which has just $a!$ members. Hence the number of members of R_a is at most

$$a!\,(n!/a!\,(n-a)!) = n!/(n-a)! \leqslant n^a.$$

If we write

$$c_a = \max_{\omega \in R_a} [G_\omega]_q,$$

we have $H_a \leqslant n^a c_a$.

We shall prove more than (1.1), namely

THEOREM 3. *If* (1.2) *is true, then*

$$n!\, T_{nq} - F_{nq} \sim H_2 = Nc_2 \sim NF_{nq}\beta^{n-2}e^{-\gamma} = o(F_{nq}), \tag{1.10}$$

where

$$\lambda = q/N, \quad \beta = \lambda^2 + (1-\lambda)^2, \quad \gamma = 4\lambda(1-\lambda)(1-2\lambda)^2\beta^{-2}.$$

To prove the first part of (1.10) it is enough, in view of what we have just said, to prove one or other of

$$\sum_{a=3}^{n} H_a = o(H_2), \quad \sum_{a=3}^{n} n^{a-2}c_a = o(c_2). \tag{1.11}$$

Since there is complete symmetry between q and $N-q$ in all we have said so far, we may, without loss of generality, suppose henceforth that

$$0 < q \leqslant N/2, \tag{1.12}$$

so that $0 < \lambda \leqslant 1/2$, $1/2 \leqslant 1-\lambda < 1$ and (1.2) becomes

$$(q/n) - (\log n)/2 \to \infty \tag{1.13}$$

as $n \to \infty$.

We remark that $1/2 \leqslant \beta \leqslant 1$, $\gamma = O(\lambda) = O(1)$ and $C < e^{-\gamma} < C$. If q satisfies (1.12) and (1.13), we have $\lambda > (\log n)/(n-1)$, and

$$n \log \beta \leqslant -2n\lambda(1-\lambda) \leqslant -n\lambda < -\log n. \tag{1.14}$$

We have $P_1 = \{p_1(p_1-1)/2\} + p_2$ (see, for example, [4]) and, by (1.5), since $p_1 = n-a$, we must have $p_2 \leqslant a/2$. Hence

$$P_1 \geqslant (n-a)(n-a-1)/2, \tag{1.15}$$

$$P_1 \leqslant \{(n-a)(n-a-1)+a\}/2 = N - a(2n-a-2)/2, \tag{1.16}$$

$$N - P_1 \geqslant a(2n-a-2)/2. \tag{1.17}$$

Again, by (1.6) and (1.15),

$$\sum_{j \geqslant 2} jP_j = N - P_1 \leqslant \tfrac{1}{2}\{n^2 - n - (n-a)^2 + (n-a)\} \leqslant 2an,$$

and so

$$P_j \leqslant an \quad (j \geqslant 2). \tag{1.18}$$

A well-known result that we use several times is that

$$F_{nq} \sim LM(2\pi)^{-\frac{1}{2}}, \tag{1.19}$$

where

$$L = \{\lambda^\lambda (1-\lambda)^{1-\lambda}\}^{-N}, \quad M = \{N\lambda(1-\lambda)\}^{-\frac{1}{2}}. \tag{1.20}$$

2. Proof that (1.2) is necessary for (1.1)

Let us write

$$\psi(n) = (q/n) - (\log n)/2 \tag{2.1}$$

and suppose that $\psi(n)$ does not tend to infinity with n. Then there is an infinite sequence of values of n such that $\psi(n) < C$. In this section we suppose n confined to this sequence, so that $q < Cn \log n$.

We have now by (1.8) and (1.9)

$$n!\,T_{nq} - F_{nq} \geqslant H_2 = Nc_2 = N[(1+X)^{N-2n+4}(1+X^2)^{n-2}]_q \geqslant N[(1+X)^{N-2n+4}]_q = \eta F_{nq},$$

where

$$\eta = N\binom{N-2n+4}{q} \Big/ \binom{N}{q} = N \prod_{s=0}^{q-1} \left(\frac{N-2n+4-s}{N-s}\right)$$

and so

$$\begin{aligned} \log \eta &= \log N + \sum_{s=0}^{q-1} \log\left(1 - \frac{2n-4+s}{N}\right) - \sum_{s=0}^{q-1} \log\left(1 - \frac{s}{N}\right) \\ &= \log N - 2qN^{-1}(n-2) + O(qN^{-2}\{q^2+n^2\}) \\ &= -4\psi(n) + O(1) = O(1), \end{aligned}$$

by (2.1), so that $\eta > C$. Hence, for this sequence of n, we have $n!\,T_{nq} > (1+C)F_{nq}$, and (1.1) is false.

3. Approximation to c_2

LEMMA 1. *If $q \to \infty$ as $n \to \infty$, then $c_2 \sim F_{nq}\beta^{n-2}e^{-\gamma}$.*

We write

$$\varepsilon = (n-2)/N, \quad \Phi = 2(1-2\lambda)/\beta, \quad X_0 = \lambda(1+\varepsilon\Phi)/(1-\lambda).$$

If we put $X = X_0 e^{it}$ and write $T = q^{-2/5}$ and

$$\chi = \chi(t) = (N-2n+4)\log(1+X) + (n-2)\log(1+X^2) - q\log X,$$

we have, by (1.9) and (1.3),

$$2\pi c_2 = \int_{-\pi}^{\pi} e^{\chi(t)}dt = J_1 + J_2, \quad J_1 = \int_{-T}^{T} e^{\chi(t)}dt.$$

We consider first J_1, so that $-T \leqslant t \leqslant T$ and $t = O(T) = o(1)$. We have

$$(1-\lambda)X = (1-\lambda)X_0 e^{it} = \lambda(1+\varepsilon\Phi)e^{it} = \lambda(1+\alpha_1+\alpha_2),$$

where $\alpha = \varepsilon + |t| = o(1)$, $\alpha_1 = \varepsilon\Phi + it = O(\alpha)$ and $\alpha_2 = O(\alpha^2)$. Hence $(1-\lambda)(1+X) = 1+\lambda\alpha_1+\lambda\alpha_2$ and $(1-\lambda)^2(1+X^2) = \beta + 2\lambda^2\alpha_1 + O(\lambda^2\alpha^2)$. We have then

$$N^{-1}\chi = (1-2\varepsilon)\log(1+\lambda\alpha_1+\lambda\alpha_2) + \varepsilon\log(\beta+2\lambda^2\alpha_1+O(\lambda^2\alpha^2)) \\ -\lambda\log\lambda - (1-\lambda)\log(1-\lambda) - \lambda\log(1+\alpha_1+\alpha_2),$$

$$N^{-1}(\chi - \log L) - \varepsilon\log\beta = (1-2\varepsilon)(\lambda\alpha_1 + \lambda\alpha_2 - \tfrac{1}{2}\lambda^2\alpha_1^2) + (2\lambda^2\varepsilon\alpha_1/\beta) - \lambda\alpha_1 - \lambda\alpha_2 + \tfrac{1}{2}\lambda\alpha_1^2 + O(\lambda\alpha^3) \\ = 2\varepsilon\lambda\alpha_1(\lambda-\beta)/\beta + \tfrac{1}{2}\lambda\alpha_1^2(1-\lambda) + O(\lambda\alpha^3).$$

Now $\beta - \lambda = \Phi\beta(1-\lambda)/2$ and so

$$N^{-1}(\chi - \log L) - \varepsilon\log\beta = \tfrac{1}{2}\alpha_1\lambda(1-\lambda)(\alpha_1 - 2\varepsilon\Phi) + O(\lambda\alpha^3) = -\tfrac{1}{2}\lambda(1-\lambda)(\varepsilon^2\Phi^2+t^2) + O(\lambda\alpha^3).$$

Again $\lambda N\alpha^3 < Cq^{-1/5}$, $N\varepsilon^2 = 2 + O(n^{-1})$, $\lambda(1-\lambda)\Phi^2 = \gamma$ and so

$$\chi(t) = \log L + (n-2)\log\beta - \gamma - \delta^2 t^2 + O(q^{-1/5}), \tag{3.1}$$

where $\delta^2 = \lambda N(1-\lambda)/2$, $Cq < \delta^2 < Cq$ and $2\delta^2 M^2 = 1$ by (1.20). Hence $\chi(t) = \chi(0) - \delta^2 t^2 + O(q^{-1/5})$ and

$$J_1 \sim e^{\chi(0)}\int_{-T}^{T} e^{-\delta^2 t^2}dt = \delta^{-1}e^{\chi(0)}\int_{-\delta T}^{\delta T} e^{-u^2}du \sim \delta^{-1}e^{\chi(0)}\sqrt{\pi},$$

since $\delta^2 T^2 > Cq^{\frac{1}{5}} \to \infty$ as $q \to \infty$. Hence

$$J_1 \sim \sqrt{2\pi}\,Me^{\chi(0)} > Cq^{-\frac{1}{2}}e^{\chi(0)} \tag{3.2}$$

Now
$$J_2 = \int_T^{\pi} + \int_{-\pi}^{-T} e^{\chi(t)}dt \leqslant 2\int_T^{\pi} |e^{\chi(t)}|\,dt.$$

When $T \leqslant t \leqslant \pi$, we have

$$|1+X|^2 = (1+X_0)^2 - 4X_0 \sin^2(t/2) \leqslant (1+X_0)^2 e^{-C\xi T^2}, \tag{3.3}$$

where $\xi = X_0(1+X_0)^{-2} > C\lambda$. Hence

$$|e^{\chi(t)}| = |(1+X)^{N-2n+4}(1+X^2)^{n-2}X^{-q}| \leqslant e^{\chi(0)-C\xi NT^2} \leqslant e^{\chi(0)-CqT^2} = e^{\chi(0)-Cq^{1/5}} = o(J_1)$$

by (3.2). Hence $J_2 = o(J_1)$ and

$$c_2 \sim J_1/(2\pi) \sim (2\pi)^{-\frac{1}{2}} M e^{\chi(0)} \sim F_{nq}\beta^{n-2}e^{-\gamma}$$

by (3.2), (3.1) and (1.19).

4. Proof that (1.4) is sufficient for (1.10)

If $0 < X_1 < 1$ and j is an integer greater than 1, we have $(1+X_1^j)^2 \leqslant (1+X_1^2)^j$. Hence

$$[G_\omega]_q \leqslant X_1^{-q} G_\omega(X_1) = X_1^{-q} \prod_j (1+X_1^j)^{P_j} \leqslant X_1^{-q}(1+X_1)^{P_1}(1+X_1^2)^{(N-P_1)/2}$$

by (1.6). If (1.4) is satisfied we may, by (1.12) suppose that

$$3n \log n < q \leqslant \tfrac{1}{2}N, \; 6(n-1)^{-1} \log n < \lambda \leqslant \tfrac{1}{2}. \tag{4.1}$$

We now choose $X_1 = q/(N-q) = \lambda/(1-\lambda)$, so that we have

$$[G_\omega]_q \leqslant \lambda^{-q}(1-\lambda)^{q-N}\beta^{(N-P_1)/2} = L\beta^{(N-P_1)/2}$$

by (1.20). By Lemma 1, $c_2 > CF_{nq}\beta^{n-2}$ and so

$$c_a/c_2 = \max_{\omega \in R_a} [G_\omega]_q/c_2 \leqslant CL\beta^\mu/F_{nq} \leqslant Cn\beta^\mu$$

by (1.19), where $\mu = \{(N-P_1)/2\} - n + 2$. To prove (1.11), from which (1.10) follows, it is then enough to show that

$$\sum_{a=3}^{n} n^{a-1}\beta^\mu = o(1). \tag{4.2}$$

We have

$$\log \beta = \log(1-2\lambda(1-\lambda)) \leqslant -2\lambda(1-\lambda) \leqslant -\lambda \leqslant -6n^{-1}\log n$$

by (4.1). Again, by (1.17), since $a \leqslant n$,

$$\mu = \{(N-P_1)/2\} - n + 2 \geqslant \{a(2n-a-2)/4\} - n + 2 \geqslant (n-2)(a-4)/4.$$

Hence

$$\log(n^{a-1}\beta^\mu)/\log n \leqslant a - 1 - 3\{(n-2)(a-4)/(2n)\} \leqslant -(a/2) + 8$$

and so

$$\sum_{a=18}^{n} n^{a-1}\beta^\mu \leqslant \sum_{a=18}^{n} n^{8-(a/2)} \leqslant 1/n = o(1).$$

If $3 \leqslant a \leqslant 17$, we have $\mu \geqslant \{(a-2)\, n/2\} - C$,

$$\log\, (n^{a-1}\beta^{\mu})/\log n \leqslant -2a+5+o(1) \leqslant -1+o(1)$$

and so

$$\sum_{a=1}^{17} n^{a-1}\beta^{\mu} < C/n = o(1).$$

Hence (4.2) and so (1.11).

5. Proof that (1.2) is sufficient for (1.10)

We now turn our attention to those q which satisfy (1.2) but not (1.4), i.e. those q for which

$$n\{\log n + \psi(n)\}/2 \leqslant q \leqslant 3n \log n, \tag{5.1}$$

where $\psi(n) \to \infty$ as $n \to \infty$. We may suppose that $\psi(n) = o(\log n)$. We have $q < Cn \log n$, $\lambda < Cn^{-1} \log n$ and so

$$-(n-2) \log \beta = 2\lambda(n-2) + O(\lambda^2 n) = 4(q/n) + o(1). \tag{5.2}$$

We write $A = n - n^{3/4} (\log n)^{1/2}$ and consider first those ω for which

$$2 \leqslant a \leqslant A. \tag{5.3}$$

We have

$$[Q_\omega]_q = \frac{1}{2\pi} \int_{-\pi}^{\pi} Q_\omega(X_2 e^{it})\, X_2^{-q} e^{-iqt}\, dt \leqslant C J_3 J_4,$$

where $X_2 = q/(P_1 - q)$ and

$$J_3 = \prod_{j \geqslant 2} (1 + X_2^j)^{P_j}, \quad J_4 = X_2^{-q} \int_{-\pi}^{\pi} |1 + X_2 e^{it}|^{P_1}\, dt.$$

By (1.15) and (5.3), $P_1 \geqslant C(n-A)^2 > Cn^{\frac{3}{2}} \log n$. Again $q < Cn \log n$, so that $X_2 < Cn^{-\frac{1}{2}}$ and, by (1.18),

$$\log J_3 \leqslant \sum_{j \geqslant 2} P_j X_2^j \leqslant Can \sum_{j \geqslant 2} n^{-j/2} \leqslant Ca.$$

By an argument similar to that of (3.3), we have

$$|1 + X_2 e^{it}|^2 \leqslant e^{-C\xi t^2} (1 + X_2)^2,$$

where $\xi = X_2(1 + X_2)^{-2} = qP_1^{-2}(P_1 - q)$. Hence

$$J_4 \leqslant X_2^{-q} (1 + X_2)^{P_1} \int_{-\pi}^{\pi} e^{-CP_1 \xi t^2}\, dt \leqslant C X_2^{-q} (1 + X_2)^{P_1} (P_1 \xi)^{-\frac{1}{2}}$$

$$= CP_1^{P_1 + \frac{1}{2}} q^{-q-\frac{1}{2}} (P_1 - q)^{-P_1 + q - \frac{1}{2}} \sim C \binom{P_1}{q}$$

by a result similar to (1.19). It follows that

$$[Q_\omega]_q \leqslant Ce^{Ca}\binom{P_1}{q}.$$

Hence, by Lemma 1,

$$[Q_\omega]_q/c_2 \leqslant Ce^{Ca}\beta^{2-n}\binom{P_1}{q}\Big/\binom{N}{q} \leqslant Ce^{Ca}\beta^{2-n}\left(\frac{P_1}{N}\right)^q,$$

that is

$$\log\,(c_a/c_2) < C + Ca - (n-2)\log\beta + q\log\,(P_1/N). \tag{5.4}$$

Now, by (1.16),

$$\frac{P_1}{N} \leqslant \frac{(n-a)(n-a-1)+a}{n(n-1)} = \left(1-\frac{a}{n}\right)^2\left(1+\frac{a^2}{(n-1)(n-a)^2}\right).$$

By (5.3), $(n-1)(n-a)^2 > C(n-1)\,n^{\frac{3}{2}}\log n > n^2$ and so

$$\log\,(P_1/N) < -2\,a/n. \tag{5.5}$$

Using (5.2) and (5.5) in (5.4), we have

$$\begin{aligned}\log\,(n^{a-2}c_a/c_2) &\leqslant C + (a-2)\log n + Ca - 2(q/n)(a-2)\\ &= C + (a-2)\{C + \log n - 2(q/n)\} \leqslant C + (a-2)\{C - \psi(n)\}\end{aligned}$$

by (5.1) and so

$$\sum_{3\leqslant a\leqslant A} H_a/H_2 \leqslant \sum_{3\leqslant a\leqslant A} n^{a-2}c_a/c_2 \leqslant Ce^{C-\psi(n)} = o(1). \tag{5.6}$$

Finally let us consider those ω for which $A < a \leqslant n$. For these a we have $P_1 \leqslant Cn^{\frac{3}{2}}\log n$ by (1.16). Also, by (1.6),

$$\sum_{j\geqslant 2} P_j \leqslant (N-P_1)/2 \leqslant N/2.$$

We write

$$X_3 = \{q/(N-q)\}^{\frac{1}{2}} < Cn^{-\frac{1}{2}}(\log n)^{\frac{1}{2}} < 1.$$

We have then

$$[Q_\omega]_q \leqslant X_3^{-q}Q_\omega(X_3) = X_3^{-q}\prod_{j\geqslant 1}(1+X_3^j)^{P_j}.$$

Now

$$\log\,(1+X_3)^{P_1} = P_1\log\,(1+X_3) \leqslant P_1X_3 < Cn\,(\log n)^{\frac{3}{2}}$$

and

$$\prod_{j\geqslant 2}(1+X_3^j)^{P_j} \leqslant \prod_{j\geqslant 2}(1+X_3^2)^{P_j} \leqslant (1+X_3^2)^{N/2}$$

and so

$$\log\,[Q_\omega]_q < Cn\,(\log n)^{\frac{3}{2}} - q\log X_3 + \tfrac{1}{2}N\log\,(1+X_3^2) = Cn\,(\log n)^{\frac{3}{2}} + \tfrac{1}{2}\log L.$$

If we write

$$Z = \sum_{A<a\leqslant n} H_a = \sum_{A<a\leqslant n}\ \sum_{\omega\in H_a}[Q_\omega]_q,$$

there are less than $n!$ terms in the double sum and $\log(n!) < Cn \log n$. Again, by Lemma 1 and (1.19),

$$H_2 = Nc_2 > CN\beta^{n-2}F_{nq} > CN\beta^{n-2}LM.$$

Also $-(n-2)\log\beta = O(\log n)$ by (5.2), $\log M > C\log q > C\log n$ and

$$\log L > q\log(N/q) > Cn(\log n)^2.$$

Hence
$$\log(Z/H_2) \leqslant Cn(\log n)^{\frac{3}{2}} - (\log L)/2 \to -\infty$$

as $n \to \infty$ and so $Z = o(H_2)$. Combining this with (5.6), we have the first part of (1.11), and (1.10) follows.

References

[1]. DE BRUIJN, N. G., Chap. 5, *Applied combinatorial mathematics*. E. F. Beckenback (ed.), New York 1964.

[2]. FORD, G. W. & UHLENBECK, G. E., Combinatorial problems in the theory of graphs. *Proc. Nat. Acad. Sci. U.S.A.*, 43 (1957), 163–167.

[3]. HARARY, F., The number of linear, directed, rooted and connected graphs. *Trans. Amer. Math. Soc.*, 78 (1955), 445–463.

[4]. OBERSCHELP, W., Kombinatorische Anzahlbestimmungen in Relationen. *Math. Ann.* 174 (1967), 53–78.

[5]. POLYA, G., Kombinatorische Anzahlbestimmungen für Gruppen, Graphen, und Chemische Verbindungen. *Acta Math.*, 68 (1937), 145–254.

Received July 31, 1969