THE NUMBER OF SPARSELY EDGED LABELLED HAMILTONIAN GRAPHS

by E. M. WRIGHT†

(Received 20 August, 1981)

An (n, q) graph is a graph on n labelled points and q lines, no loops and no multiple lines. We write $N = \frac{1}{2}n(n-1)$, $B(a, b) = a!/\{b!(a-b)!\}$ and $B(a, 0) = 1$, so that there are just $B(N, q)$ different (n, q) graphs. Again $h(n, q)$ is the number of Hamiltonian (n, q) graphs. Much attention has been devoted to the problem of determining for which $q = q(n)$ "almost all" (n, q) graphs are Hamiltonian, i.e. for which q we have

$$h(n, q)/B(N, q) \to 1$$

as $n \to \infty$. I proved [**8**, Theorem 4] that $qn^{-3/2} \to \infty$ is a sufficient condition by showing that, for such q, almost all (n, q) graphs have about the average number of Hamiltonian circuits (H.c.s). My calculations also showed that this last result was false if $qn^{-3/2} \to 0$ and so that this method would not take us much further. But, by other methods, the sufficient condition has been successively improved by Komlós and Szemerédi to

$$q > Cn \exp(\sqrt{\log n}),$$

by Pósa to

$$q > Cn \log n$$

and again by Komlós and Szemerédi to

$$q > (\tfrac{1}{2} + \varepsilon)n \log n.$$

Finally Korsonov [**5**] announced a proof that

$$\Omega(n, q) = (q/n) - \tfrac{1}{2}\log n - \tfrac{1}{2}\log\log n \to \infty$$

is a sufficient condition. Since this is also a necessary condition (a trivial deduction from [**3**, Theorem 2]), this settles the matter, except for the possibility of a "threshold" result (in the language of [**2**]) when $\Omega(n, q)$ tends to a finite limit as $n \to \infty$.

There remains the problem of finding a formula, exact or asymptotic, for $h(n, q)$ when $\Omega \to -\infty$ as $n \to \infty$. It is trivial that

$$h(n, n+k) = 0 \quad (k < 0), \qquad h(n, n) = \tfrac{1}{2}\{(n-1)!\}. \tag{1}$$

Here I give exact formulae for $h(n, n+k)$ for $k = 1, 2, 3$; the work is a little cumbrous for $k = 3$ but could, with sufficient labour, be extended to $k = 4$. Beyond that, the method seems impracticable. But we can prove two much more extensive asymptotic results fairly simply. We write

$$M = \tfrac{1}{2}\{(n-1)!\}B(N-n, k).$$

† The research reported herein was supported by the European Research Office of the United States Army.

THEOREM 1. *If* $k/n \to 0$ *as* $n \to \infty$, *then* $h(n, n+k)/M \to 1$.

We write

$$\lambda(\theta) = -\log(1-\theta) - \theta = \sum_{t \geq 2} \theta^t/t$$

and ω for the number for which $0 < 4\omega < 1$ and $\lambda(4\omega) = 4$. A routine calculation shows that $\omega = 0{\cdot}248304\ldots$. We write ε for any fixed positive number independent of n and k.

THEOREM 2. *If* $0 < k < (\omega - \varepsilon)n$, *then*

$$\log h(n, n+k) = \log M + O(1)$$

as $n \to \infty$.

It is of some interest to contrast our state of knowledge about $f(n, n+k)$, the number of connected (n, q) graphs, with that about $h(n, n+k)$. Trivially $f(n, n+k) = 0$ when $k < -1$ and (not trivially) $f(n, n-1) = n^{n-2}$, a result due to Cayley [**1**] (see also [**6**]). Again Rényi [**7**] found a formula for $f(n, n)$, Bagaev one for $f(n, n+1)$ and I [**9**] found a recurrence method, well adapted to machine computation, to calculate an exact formula for $f(n, n+k)$ for successive k and general n. The result becomes cumbrous and uninformative as k gets larger, but the calculation was taken as far as $k = 24$, when it was halted by the limits on the memory of the machine and not by any complication of method [**4**]. On the other hand, I have only succeeded in finding an asymptotic approximation to $f(n, n+k)$ for large n when $k = o(n^{1/3})$, a result greatly inferior to Theorem 1 above, and that at the cost of a much more elaborate proof [**11**]. When $k = O(1)$, a much simpler deduction from the recurrence method is sufficient.

Similar results are true for $u(n, n+k)$, the number of non-separable $(n, n+k)$ graphs, and for $v(n, n+k)$, the number of smooth graphs, i.e. connected graphs without end points (see [**10**]). For each of these functions, the asymptotic result is valid for $k = o(n^{1/2})$.

Thus, for $h(n, n+k)$, we have asymptotic results valid over a much wider range of k than for $f(n, n+k)$, $u(n, n+k)$ or $v(n, n+k)$, but for these last functions we can calculate exact formulae for small k by a recurrence method and this seems to be impossible for $h(n, n+k)$.

1. Preliminaries. We write $L_s = L_s(k)$ for the number of $(n, n+k)$ graphs containing s Hamiltonian circuits (H.c.s), each graph being counted according to the number of different sets of s H.c.s which it contains. By the Inclusion–Exclusion Theorem, we have then

$$h(n, n+k) = L_1 - L_2 + L_3 - L_4 + \ldots \tag{2}$$

and

$$L_1 - L_2 \leq h(n, n+k) \leq L_1. \tag{3}$$

We label the points of an (n, q) graph by the numbers $1, 2, 3, \ldots, n$ and regard a and $a+n$ as labelling the same point. A G_{st} graph is an $(n, n+t)$ graph which has s H.c.s, one of which is the *basic* circuit $1, 2, 3, \ldots n$, and in which every line occurs in at least one of the s H.c.s. The number of G_{st} is T_{st}, where each G_{st} is counted according to the number of such different sets of s H.c.s it contains. Clearly

$$T_{10} = 1, \quad T_{1t} = T_{s0} = T_{s1} = 0 \quad (t > 0, s > 1). \tag{4}$$

We now find a formula for $L_s(k)$ in terms of the T_{st}. An $(n, n+k)$ graph contributes 1 to $L_s(k)$ for every different set of s H.c.s which it contains. The sub-graph formed by a set of s H.c.s is isomorphic to a G_{st} in s ways, since each of the s H.c.s may be taken as isomorphic to the basic circuit in the G_{st}. The H.c. isomorphic to the basic H.c. may be any of $\frac{1}{2}\{(n-1)!\}$ possible H.c.s and, given a particular set of s H.c.s isomorphic to a G_{st}, the remaining $k-t$ edges in the $(n, n+k)$ graph may be chosen in $B(N-n-t, k-t)$ ways. Hence

$$L_s(k) = \frac{(n-1)!}{2s} \sum_{t=0}^{k} T_{st} B(N-n-t, k-t). \tag{5}$$

Using (4) in this, we obtain

$$L_1(k) = \tfrac{1}{2}\{(n-1)!\} B(N-n, k) = M \tag{6}$$

and $L_s(1) = 0$ if $s \geq 2$. Hence

$$h(n, n+1) = L_1(1) = \tfrac{1}{4}(n-3)(n!). \tag{7}$$

Again, for $s \geq 2$, (5) becomes

$$L_s(k) = \frac{(n-1)!}{2s} \sum_{t=2}^{k} T_{st} B(N-n-t, k-t). \tag{8}$$

2. Proof of Theorems 1 and 2. T_{2t} is the number of ways in which t new lines can be added to the basic H.c. so as to produce a further H.c. to which each of the new lines belong. The t lines of the basic H.c. which do not occur in the second H.c. can be chosen in $B(n, t)$ ways. When these t lines are removed, the basic H.c. is reduced to a graph with t components, each of which is either a path or an isolated point. To construct the new H.c. we join up these t components with t new lines. We can arrange the components in $\frac{1}{2}\{(t-1)!\}$ different orders and we can then choose the sense of each component in the new H.c. in at most 2 ways. Hence we can construct the new H.c. in at most $2^{t-1}\{(t-1)!\}$ ways and so

$$T_{2t} \leq 2^{t-1}\{(t-1)!\} B(n, t).$$

Thus, by (8),

$$L_2(k) \leq (n-1)! \sum_{t=2}^{k} 2^{t-3}(t-1)! B(n, t) B(N-n-t, k-t)$$

and, by (6),

$$L_2(k)/L_1(k) \leq \tfrac{1}{4} \sum_{t=2}^{k} \alpha_t,$$

if we write

$$\alpha_t = 2^t(t-1)!B(n,t)B(N-n-t,k-t)/B(N-n,k)$$
$$= \frac{2^t}{t}\left\{\frac{n\ldots(n-t+1)k\ldots(k-t+1)}{(N-n)\ldots(N-n-t+1)}\right\} \leqslant \frac{\theta^t}{t},$$

where $\theta = 2kn/(N-n) = 4k/(n-3)$. Hence

$$L_2(k)/L_1(k) \leqslant -\tfrac{1}{4}\{\log(1-\theta)+\theta\} = \tfrac{1}{4}\lambda(\theta)$$

and so, by this and (3) and (6), we have

$$1-\tfrac{1}{4}\lambda(\theta) \leqslant h(n, n+k)/M \leqslant 1. \tag{9}$$

If $k/n \to 0$ as $n \to \infty$, we have $\theta \to 0$ and $\lambda(\theta) = O(\theta^2) \to 0$. Theorem 1 follows at once from (9). Next, if $k < (\omega - \varepsilon)n$, we have $k < (\omega - \frac{1}{2}\varepsilon)(n-3)$ for $n > n_0(\varepsilon)$ and so $\lambda(\theta) < 4 - \delta$ for a positive δ depending on ε but not on k and n. Theorem 2 follows from (9).

3. The value of $h(n, n+k)$ for $k \leqslant 3$. To avoid trivialities we take $n \geqslant 7$. In view of (1) and (7), we have only to calculate $h(n, n+2)$ and $h(n, n+3)$, i.e. by (8), to calculate T_{s2} and T_{s3} for $s \geqslant 2$.

We can add a pair of lines, viz. $\{a, b\}$ and $\{a+1, b+1\}$, chosen in just $N-n$ ways, to the basic H.c. to produce a second H.c. Hence

$$T_{22} = N - n = \tfrac{1}{2}n(n-3).$$

We cannot obtain a third H.c. in this way and so $T_{s2} = 0$ for $s \geqslant 3$. Hence, by (8),

$$L_2(2) = \tfrac{1}{4}(n-1)!T_{22} = (n-3)(n!)/8 \qquad L_s(2) = 0 \quad (s \geqslant 3),$$

and so, by (2),

$$h(n, n+2) = L_1(2) - L_2(2) = (n+1)!(n-3)(n-4)/16.$$

The calculation of $h(n, n+3)$ is more tedious and we omit the details. We find that

$$T_{23} = \tfrac{1}{3}n(n-4)(2n-7), \qquad T_{33} = \tfrac{3}{2}n(n-3), \qquad T_{43} = n$$

and $T_{s3} = 0$ for $s \geqslant 5$. Hence, by (4) and (8),

$$L_2(3) = n!(n-4)(3n^2+2n-37)/48,$$
$$L_3(3) = n!(n-3)/4, \qquad L_4(3) = n!/8$$

and $L_s(3) = 0$ for $s \geqslant 5$. Thus (2) gives us

$$h(n, n+3) = n!(n^5 - 9n^4 + 15n^3 + 29n^2 + 68n - 404)/96.$$

With sufficient labour, we could use this method to calculate $h(n, n+4)$, but the prospect is somewhat daunting. Beyond $k = 4$, the method seems impracticable.

REFERENCES

1. A. Cayley, A theorem on trees, *Quart. J. Math.* **23** (1889), 376–378.

2. P. Erdös and A. Rényi, On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17–61; reprinted in P. Erdös, *The art of counting* (M.I.T. Press, 1973), 574–617.

3. P. Erdös and A. Rényi, On the strength of connectedness of a random graph, *Acta Math. Acad. Sci. Hungar.* **12** (1961), 261–267; reprinted in P. Erdös, *The art of counting* (M.I.T. Press, 1973), 618–624.

4. P. M. D. Gray, A. M. Murray and N. A. Young, Wright's formulae for the number of connected sparsely edged graphs, *J. Graph Theory* **1** (1977), 331–334.

5. A. D. Koršunov, Solution of a problem of Erdös and Rényi on Hamiltonian cycles in nonoriented graphs, *Soviet Math. Dokl.* **17** (1976), 760–764; (*Dokl. Akad. Nauk SSSR* **228** (1976), *no. 3*, 529–532).

6. J. W. Moon, Various proofs of Cayley's formula for counting trees, *A seminar on graph theory*, ed. F. Harary (Holt, Rinehart and Winston, 1967), 70–78.

7. A. Rényi, On connected graphs I, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **4** (1959), 385–388.

8. E. M. Wright, For how many edges is a graph almost certainly Hamiltonian?, *J. London Math. Soc.* (2) **8** (1974), 44–48.

9. E. M. Wright, The number of connected sparsely edged graphs, *J. Graph Theory* **1** (1977), 317–330.

10. E. M. Wright, The number of connected sparsely edged graphs. II. Smooth graphs and blocks, *J. Graph Theory* **2** (1978), 299–305.

11. E. M. Wright, The number of connected sparsely edged graphs. III. Asymptotic results, *J. Graph Theory* **4** (1980), 393–407.

UNIVERSITY OF ABERDEEN
SCOTLAND