ON CONNECTED GRAPHS, I.

by

A. RÉNYI

Let $G_{n,N}$ denote a connected graph having the n labelled vertices $P_1, P_2, \ldots, P_n$ and N edges. Let $C(n, N)$ denote the number of all possible $G_{n,N}$. Let us call the number $d = N - n + 1$ the „degree of connectivity" of the connected graph $G_{n,N}$. Clearly $d \geqq 0$, and the graph is a tree if and only if $d = 0$. The number of different trees with n labelled vertices is, according to a classical result of CAYLEY [1] equal to n^{n-2}, i. e.

$$C(n, n-1) = n^{n-2} \qquad (n = 1,2,\ldots). \tag{1}$$

A simple explicit formula for $C(n, N)$ is not known, moreover the asymptotic behaviour of $C(n, n + d - 1)$, if d is fixed and $n \to +\infty$ has according to the knowledge of the author, not determined up to now.

One can give some rather complicated explicit formulae resp. recursive relations for $C(n, N)$ (see [2] and [3]). For instance one obtains by the usual sieve-method

$$C(n, N) = \binom{\binom{n}{2}}{N} - \frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}\binom{\binom{k}{2}+\binom{n-k}{2}}{N} + \tag{2}$$

$$+ \frac{1}{3}\sum_{k+l<n}\sum \frac{n!}{k!\,l!\,(n-k-l)!}\binom{\binom{k}{2}+\binom{l}{2}+\binom{n-k-l}{2}}{N} - \ldots$$

which leads to an explicit formula for the generating function

$$\sum_{n=1}^{\infty}\sum_{N=1}^{\infty}\frac{c(n, N)\, x^n y^N}{n!} = \log\left(1 + \sum_{k=1}^{\infty}\frac{(1+y)^{\binom{k}{2}} x^k}{k!}\right). \tag{3}$$

There is also the recursion formula

$$\binom{\binom{n+1}{2}}{N} = \sum_{k=0}^{n}\binom{n}{k}\sum_{m=k}^{\binom{k+1}{2}}\binom{\binom{n-k}{2}}{N-m} C(k+1, m). \tag{4}$$

It seems however that these formulae do not help much if one wants to determine the asymptotic behaviour of $C(n, n + d - 1)$.

The aim of the present note is to discuss the case $d = 1$ that is to give for $C(n, n)$ an asymptotic formula. In a subsequent note we shall discuss the cases $d \geqq 2$. Clearly if a connected graph consists of n vertices and n edges, it contains exactly one circle, some vertices of which are the roots of a tree, these trees being disjoint. (Fig. 1. shows all possible types of connected graphs with 6 vertices and 6 edges.)

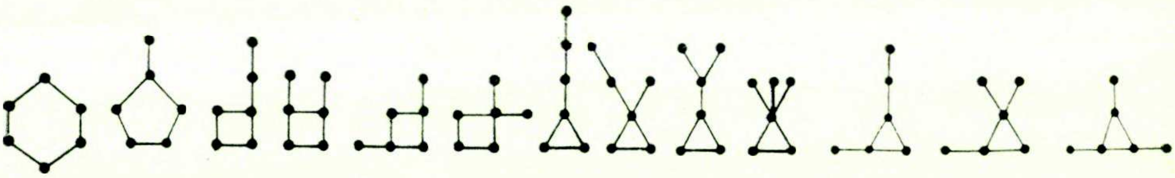

Figure 1.

If the simple circle contained in such a graph consists of k vertices, then $3 \leq k \leq n$ all these values being possible. If the edges of this circle are removed from a graph of the considered type there remains a graph with n vertices consisting of k disjoint trees as subgraphs such that the k vertices of the removed circle belong to different subgraphs. The number $T(n, k)$ of such graphs has been determined already by CAYLEY [3]; he asserted that

$$T(n, k) = kn^{n-k-1}. \tag{4}$$

A proof of this result of Cayley has been given in [4]. This formula is contained as a special case in a more general result of G. W. FORD and G. E. UHLENBECK [5] (see also [6]). Using this result one easily obtains

$$C(n, n) = \sum_{k=3}^{n} T(n, k) \binom{n}{k} \frac{(k-1)!}{2} = \frac{n^{n-1}}{2} \sum_{k=3}^{n} \binom{n}{k} \frac{k!}{n^k}. \tag{5}$$

As a matter of fact the vertices of a circle consisting of k points can be chosen in $\binom{n}{k}$ ways; they can be arranged to a circle in $\frac{(k-1)!}{2}$ different ways; the remaining points can be arranged as mentioned above in $T(n, k)$ different ways; as k may take any value from 3 to n, formula (5) follows immediately from (4). (5) can be written also in the form

$$C(n, n) = \frac{1}{2} n^{n-1} \sum_{k=3}^{n} \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right). \tag{5'}$$

Using (5') it is easy to determine the asymptotic behaviour of $C(n, n)$. As a matter of fact

$$\prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right) \leqq e^{-\frac{1}{n}\binom{k}{2}}$$

and for $k = o(n^{2/3})$

$$\prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right) = e^{-\frac{k^2}{2n}} \left(1 + O\left(\frac{k^3}{n^2}\right)\right). \tag{6}$$

Thus it follows that

$$\sum_{k=3}^{n} \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right) = \sqrt{\frac{n\pi}{2}} + O(1) \tag{7}$$

and therefore

$$C(n, n) \sim \sqrt{\frac{\pi}{8}}\, n^{n-\frac{1}{2}}. \tag{8}$$

It is interesting to determine the distribution of the length of the circle contained in a random $G_{n,n}$.

If γ_n denotes the length of the circle contained in a $G_{n,n}$ chosen at random (so that each of the $C(n, n)$ graphs $G_{n,n}$ has the same probability to be chosen) we obtain

$$\mathbf{P}\left(\frac{\gamma_n}{\sqrt{n}} < x\right) = \frac{1}{2} n^{n-1} \sum_{3 \leq k < x\sqrt{n}} \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right). \tag{9}$$

It follows by some elementary calculation, using (6), that

$$\lim_{n \to +\infty} \mathbf{P}\left(\frac{\gamma_n}{\sqrt{n}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{+x} e^{-\frac{u^2}{2}}\, du. \tag{10}$$

Thus $\frac{\gamma_n}{\sqrt{n}}$ has in the limit for $n \to +\infty$ the same distribution as the absolute value of a random variable having a normal distribution with mean 0 and variance 1. If follows that the mean value of γ_n is asymptotically $\sqrt{\frac{2n}{\pi}}$.

The results of the present note will be used in a forthcoming joint paper of P. Erdős and the author.

(Received October 26, 1959.)

REFERENCES

[1] Cayley, A.: „A theorem on trees." *Quarterly Journal of Pure and Applied Mathematics* **23**(1889) 376—378.

[2] Riddell, R. I.—Uhlenbeck, Jr. G. E.: „On the theory of the virial development of the equation of state of monoatomic gases." *Journal of Chemical Physics* **21** (1953) 2056—2064.

[3] Gilbert, E. N.: „Enumeration of labelled graphs." *Canadian Journal of Mathematics* **8** (1956)405—411.

[4] Rényi, A.: „Some remarks on the theory of trees." *Publications of the Matematical Institute of the Hungarian Academy of Sciences* **4**(1959)73—85.

[5] Ford, G. W.—Uhlenbeck, G. E.: „Combinatorial problems in the theory of graphs, I. "*Proceedings of the National Academy of Sciences of the USA*. **42** (1956) 122—128.

[6] Austin, T. L.—Fagen, R. E.—Penney, W. F.—Riordan, J.: „The number of components in random linear graphs." *Annals of Math. Statistics* **30**(1959)747—754.

Remark added on March 21, 1960.

The following reference should be added: Formula (5) is contained (with an other interpretation) in the paper by L. KATZ: "Probability of indecomposability of a random mapping function" (Annals of Mathematical Statistics 26 (1955) 512—517), where the number $I(n)$ of indecomposable single-valued mappings of a set of n (labelled) elements into itself is determined. It is easy to see that $I(n) = 2C(n, n)$, which establishes the equivalence of the theorem of KATZ with formula (5). The asymptotic formula (8) is also contained in the mentioned paper of KATZ.

ÖSSZEFÜGGŐ GRÁFOKRÓL, I.

RÉNYI A.

Jelölje $C(n, N)$ az n (számozott) pontból és N élből álló összes lehetséges összefüggő gráfok számát. Felhasználva CAYLEY egy képletét (amelynek bizonyítása [4]-ben található meg) a szerző kiszámítja $C(n, n)$-et (azaz azon összefüggő n-csúcspontú gráfok számát, amelyek egyetlen kört tartalmaznak) és kimutatja, hogy $C(n, n) \sim \sqrt{\frac{\pi}{8}}\, n^{n-1/2}$. E dolgozat folytatásában a szerző $C(n, n+d)$ aszimptotikus viselkedését fogja vizsgálni rögzített d mellett $n \to +\infty$ esetében.

О СВЯЗНЫХ ГРАФАХ, I.

А. RÉNYI

Резюме

Пусть $C(n, N)$ означает число всех возможных связных графов с n нумерированными точками и N ребрами. Применяя одну формулу от CAYLEY (доказанную в [4]) автор вычисляет $C(n, n)$ и покажет что $C(n, n) \sim \sqrt{\frac{\pi}{8}}\, n^{n-1/2}$. В продолжении настоящей заметки автор будет заниматься с определением асимптотической поведении от $C(n, n+d)$ для фиксированной $d \geqq 1$ для $n \to +\infty$.