SOME REMARKS ON THE THEORY OF TREES

by

ALFRÉD RÉNYI

Introduction

Let T_n denote the number of trees of order n, i. e. of trees with n labelled vertices $P_1, P_2, \ldots, P_n$. It has been proved by A. CAYLEY [1] that

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

Other proofs of (1) have been given by O. DZIOBEK [2] and G. PÓLYA [3]. The most elegant and simple proof is that given by PRÜFER [4]. The number T_n can be interpreted as the number of railway-nets connecting n given cities so that it should be possible to go from any city to any other along the net, and there are no superfluous connections. (The number of connections, i. e. the number of edges of a tree of order n is clearly $n - 1$). A. CAYLEY has treated also a more general problem. Let $T_{n,k}$ denote the number of graphs with $n + k - 1$ labelled vertices $P_1, P_2, \ldots, P_{n+k-1}$ which consist of k disjoint trees as subgraphs and are such that the points $P_1, P_2, \ldots, P_k$ belong to different subgraphs. CAYLEY gives the following formula for $T_{n,k}$

$$T_{n,k} = k(n + k - 1)^{n-2}. \tag{2}$$

Clearly $T_{n,1} = T_n$, and thus (1) is a special case of (2).

The number $T_{n,k}$ can be interpreted as follows: If there are $n + k - 1$ cities, among which k cities are lying on the bank of the same river (or sea), then $T_{n,k}$ is the number of such minimal railway nets connecting some of the cities, that it is possible to go from any city to any other, by travelling by railway along the net and eventually travelling by ship from one of the cities lying at the river to an other such city. Thus the railway-nets in question will consist of k subnets, such that every city lying at the bank of the river belongs to an other subnet, and if the cities A and B belong to the same subnet, one can go by railway from A to B, but if A and B belong to different subnets, one can travel from A to B by going first by train to the harbour-city belonging to the same subnet as A, then by ship to the harbour-city belonging to the same net as B and finally by train to B.

CAYLEY gave (2) without proof and in the literature on (1) we did not find a proof of (2). In § 1 of the present paper we shall give a proof of (2), which is based on the same idea which has been used, e. g. by DZIOBEK in proving (1).[1] Our chief aim is however not this, but the solution of some other

[1] T. GALLAI kindly called my attention to the fact that (2) can be proved also by a modification of the method of PRÜFER.

related questions. In § 1 we shall determine the number $G_k(n)$ of all graphs with n labelled vertices consisting of k disjoint trees. We shall prove that

$$(3) \qquad G_k(n) = \frac{1}{k!} \sum_{j=0}^{k} \left(-\frac{1}{2}\right)^j \binom{k}{j} \binom{n-1}{k+j-1} n^{n-k-j} \cdot (k+j)!\ .$$

Clearly we obtain from (3)

$$G_1(n) = T_n = n^{n-2}\ .$$

Thus (3) can also be considered as a generalization of (1). We obtain further by substituting $k = 2$ and $k = 3$ into (3)

$$(4) \qquad G_2(n) = \frac{1}{2} n^{n-4} (n-1)(n+6)$$

and

$$(5) \qquad G_3(n) = \frac{1}{8} n^{n-6} (n-1)(n-2)(n^2 + 13n + 60)\,.$$

Similarly we can express $G_k(n)$ for other small values of k. It is easy to prove by means of (3) that for a fixed value of k and for $n \to +\infty$ we have

$$(6) \qquad G_k(n) \sim \frac{n^{n-2}}{2^{k-1} \cdot (k-1)!}\,.$$

The formula (3) solves a problem left open in the paper of J. DÉNES [5], who proved that

$$(7) \qquad G_k(n) = n! \sum_{\substack{\sum_{j=1}^{n} a_j = k \\ \sum_{j=1}^{n} j a_j = n}} \prod_{j=1}^{n} \frac{\left(\frac{j^{j-2}}{j!}\right)^{a_j}}{a_j!}$$

and asked whether the expression for $G_k(n)$ could be brought to a simpler form. (7) is a consequence of (1) and of the fact, that the partitioning of n points into a_1 sets having each one element, a_2 sets having each 2 elements, . . ., a_j sets of j elements each . . ., etc., where

$$\sum_{j=1}^{n} a_j = k$$

and

$$\sum_{j=1}^{n} j a_j = n$$

can be done in

$$n! \prod_{j=1}^{n} \frac{1}{a_j!\,(j!)^{a_j}}$$

essentially different ways. (Here two partitions are considered identical if there exists a one-to-one correspondence of the subsets of the two partitions so that corresponding subsets contain the same elements.)

In § 2 we consider an other problem, namely the determination of the number $T(n, r)$ of trees with n vertices which have exactly r end points. (A point belonging to a tree is called an end-point of the tree if only one edge of the tree contains the given point.) We shall prove that

$$T(n, r) = \frac{n!}{r!} \mathfrak{S}_{n-2}^{n-r} \tag{8}$$

where $\mathfrak{S}_N^m$ denote the Stirling-numbers of the second kind defined by the identity

$$x^N = \sum_{m=1}^{N} \mathfrak{S}_N^m \cdot x(x-1) \dots (x-m+1) \; . \tag{9}$$

By means of (8) we prove that if ν_n denotes the number of end-points of a randomly chosen tree with n vertices, then the mean value of ν_n is asymptotically n/e for $n \to +\infty$, further that the distribution of the random variable

$$\frac{\nu_n - \frac{n}{e}}{\frac{1}{e}\sqrt{(e-2)\,n}}$$

is tending for $n \to +\infty$ to the normal distribution with mean 0 and variance 1.

The author expresses his thanks to VERA T. SÓS and T. GALLAI for their valuable remarks.

§ 1. On the number of graphs consisting of trees

Our proof of (2) will be based on the identity

$$A_{n,k} = \sum_{\sum_{i=1}^{k} j_i = n} \frac{n!}{j_1!\, j_2! \dots j_k!} j_1^{j_1-1} \cdot j_2^{j_2-1} \dots j_k^{j_k-1} = k n^{n-k} \prod_{h=1}^{k-1} (n-h) \tag{1.1}$$

where for $k = 1$ the empty product means 1.

The identity (1.1) contains as a special case for $k = 2$ the well-known identity

$$A_{n2} = \sum_{j=1}^{n-1} \binom{n}{j} j^{j-1} (n-j)^{n-j-1} = 2n^{n-2}(n-1) \, , \tag{1.2}$$

which has been proved e. g. in [2] and [5].

To prove (1.1) we shall need the identity

$$\left(\sum_{j=1}^{\infty} \frac{j^{j-1} x^j}{j!} \right)^k = \sum_{n=k}^{\infty} \frac{x^n \cdot k n^{n-k} \prod_{h=1}^{k-1} (n-h)}{n!} \tag{1.3}$$

where for $k = 1$ the empty product is to be replaced by 1. We shall prove (1.3) by induction on k. We shall start from the fact that

$$y = \sum_{j=1}^{\infty} \frac{j^{j-1} x^j}{j!} \tag{1.4}$$

is the Bürmann—Lagrange series for the inverse function $y = y(x)$ of the function

$$x = ye^{-y} \tag{1.5}$$

(see e. g. [6]). Now we obtain from (1.5)

$$ky^{k-1} y' = ky^{k-2} y' - \frac{ky^{k-1}}{x} \tag{1.6}$$

and integrating both sides of (1.6) from 0 to x we get

$$y^k = \frac{k}{k-1} y^{k-1} - k \int_0^x \frac{y^{k-1}(t)}{t} dt . \tag{1.7}$$

Now (1.3) is trivial for $k = 1$ and if it holds for $k - 1$ instead of k, then by (1.7) it follows that it holds for k too. Thus (1.3) is proved for any k. (1.3) could be proved also directly by using a well-known generalized form of the Bürmann—Lagrange-series (see [6]).

Now evidently

$$\sum_{n=1}^{\infty} \frac{A_{n,k} x^n}{n!} = \left(\sum_{j=1}^{\infty} \frac{j^{j-1} x^j}{j!} \right)^k \tag{1.8}$$

and thus (1.3) and (1.8) imply (1.1).

Formula (1) follows easily by induction from (1.2), as was shown in [2]. For the sake of completeness we reproduce the proof.

Let us select for any $i = 1, 2, \ldots, n - 1$ an arbitrary subset having i elements, of the set of vertices $P_1, \ldots, P_n$; take an arbitrary tree connecting these points, and an arbitrary tree connecting the remaining $n - i$ points. If an arbitrary point of the first tree is connected with an arbitrary point of the second tree (which can be done in $i(n - i)$ different ways) we obtain a tree of order n. Evidently if these operations are effected in every possible way, every tree with vertices $P_1, \ldots, P_n$ is obtained $2(n - 1)$ times. Thus we have

$$2(n-1) T_n = \sum_{i=1}^{n-1} \binom{n}{i} i(n-i) T_i T_{n-i} . \tag{1.9}$$

Thus if $T_i = i^{i-2}$ holds for $i = 1, 2, \ldots, n - 1$, it holds by virtue of (1.2) for $i = n$ too. As clearly $T_1 = 1 = 1^{1-2}$, (1) follows by induction for all n.

As regards (2) it can be deduced from (1) by means of (1.1) as follows :

We have evidently

$$(1.10)\qquad T_{n,k} = \sum_{\sum_{i=1}^{k} j_i = n-1} \frac{(n-1)!}{j_1!\, j_2! \ldots j_k!} T_{j_1+1} \cdot T_{j_2+1} \ldots T_{j_k+1}$$

and thus

$$(1.11)\qquad \prod_{j=1}^{k} (n+j-1)\, T_{n,k} = \sum_{\sum_{i=1}^{k} h_i = n+k-1} \frac{(n+k-1)!}{h_1!\, h_2! \ldots h_k!} h_1^{h_1-1} h_2^{h_2-1} \ldots h_k^{h_k-1}$$

and therefore by (1.1)

$$(n+k-1)(n+k-2)\ldots n \cdot T_{n,k} = k(n+k-1)^{n-1} \prod_{h=1}^{k-1} (n+k-1-h)$$

which, after dividing both sides by $(n+k-1)(n+k-2)\ldots n$, gives evidently (2).

Now we pass to the proof of (3).

We start instead of (7), from the simpler formula

$$(1.12)\qquad G_k(n) = \frac{1}{k!} \sum_{\sum_{i=1}^{k} j_i = n} \frac{n!}{j_1!\, j_2! \ldots j_k!} j_1^{j_1-2} j_2^{j_2-2} \ldots j_k^{j_k-2}$$

(1.12) follows from (1) and the remark that all graphs formed with the vertices $P_1, \ldots, P_n$ and consisting of k trees can be obtained by forming all possible partitions into k subsets of the points $P_1, \ldots, P_n$ (two partitions being taken as identical if any two points belonging to the same subset in the first partition belong also to the same subset of the second partition) and by forming for each subset, independently of each other, all possible trees connecting the points of the subset with each other.

It follows from (1.12) [or from (7)] that

$$(1.13)\qquad \sum_{n=1}^{\infty} \frac{G_k(n)\, x^n}{n!} = \frac{Y^k}{k!}$$

where Y is defined by the power series

$$(1.14)\qquad Y = \sum_{j=1}^{\infty} \frac{j^{j-2} x^j}{j!}.$$

Now we have clearly

$$(1.15)\qquad Y = \int_0^x \frac{y(t)}{t}\, dt$$

where y is the power series defined by (1.4) ; thus in view of (1.7) (for $k = 2$) we obtain

$$Y = y - \frac{y^2}{2} . \tag{1.16}$$

It follows that

$$\sum_{n=1}^{\infty} \frac{G_k(n)\, x^n}{n!} = \frac{1}{k!}\left(y - \frac{y^2}{2}\right)^k = \sum_{j=0}^{k} \frac{1}{k!}\binom{k}{j}\left(-\frac{1}{2}\right)^j y^{k+j} . \tag{1.17}$$

As by (1.8) we have

$$y^{k+j} = \sum_{n=1}^{\infty} \frac{A_{n,k+j}\, x^n}{n!} \tag{1.18}$$

it follows, taking (1.1) into account, that

$$G_k(n) = \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j}\left(-\frac{1}{2}\right)^j (k+j)\, n^{n-k-j} \prod_{h=1}^{k+j-1} (n-h) . \tag{1.19}$$

Thus (3) is proved.

Let us calculate the coefficient of the two highest power of n on the right of (1.19). Putting

$$G_k(n) = \sum_{h=1}^{k+j} G_{n,k}^{(h)}\, n^{n-h}$$

we obtain by some calculation

$$G_{n,k}^{(1)} = 0 \qquad \text{and} \qquad G_{n,k}^{(2)} = \frac{1}{2^{k-1} \cdot (k-1)!} .$$

Thus we have for any fixed k

$$\lim_{n \to +\infty} \frac{G_k(n)}{n^{n-2}} = \frac{1}{2^{k-1} \cdot (k-1)!} . \tag{1.20}$$

Of course $G_1(n) = T_n = n^{n-2}$. Other trivial special cases of (1.19) are

$$G_n(n) = 1 ,$$

$$G_{n-1}(n) = \binom{n}{2} ,$$

$$G_{n-2}(n) = 3\binom{n+1}{4} . \tag{1.21}$$

Let us put

$$H(n) = \sum_{k=1}^{n} G_k(n) \ . \tag{1.22}$$

Clearly $H(n)$ is the number of graphs of order n which are the unions of trees, i. e. of all graphs which do not contain circles (closed pathes). From (1.17) we obtain easily, putting $H(0) = 1$,

$$\sum_{n=0}^{\infty} \frac{H(n)\, x^n}{n!} = e^{y - \frac{y^2}{2}} . \tag{1.23}$$

As

$$e^{-xy - \frac{y^2}{2}} = \sum_{r=0}^{\infty} H_r(x)\, y^r$$

where

$$H_r(x) = \frac{1}{r!} e^{\frac{x^2}{2}} \frac{d^r}{dx^r} \left(e^{-\frac{x^2}{2}}\right)$$

is the r-th Hermite polynomial, we obtain

$$\sum_{n=0}^{\infty} \frac{H(n)\, x^n}{n!} = \sum_{r=0}^{\infty} H_r(-1)\, y^r \tag{1.24}$$

and thus, taking again (1.1) into account, we obtain

$$H(n) = \sum_{r=1}^{n} H_r(-1)\, r n^{n-r} \prod_{h=1}^{r-1} (n-h) \ . \tag{1.25}$$

As regards the asymptotic behaviour of $H(n)$ it follows from (1.20) easily that

$$\lim_{n \to +\infty} \frac{H(n)}{n^{n-2}} = \sqrt{e} \ . \tag{1.26}$$

§ 2. On the number of end-points of a random tree

Let us consider a tree whose vertices are the points $P_1, \ldots, P_n$. The point P_i is called an end-point of the tree, if there is only one edge of the tree connecting P_i with some other point P_j. The number of end-points of a tree of order n may have the value $2, 3, \ldots, n-1$. Let $T(n, k)$ denote the number of trees with the (labelled) vertices $P_1, \ldots, P_n$ which have exactly k end-points ($k = 2, 3, \ldots, n-1$). We shall prove first the formula (8).

Our proof utilizes the method by which Prüfer [4] proved formula (1). Prüfer's method consists in that he establishes a one-to-one correspondence between all trees with vertices $P_1, \ldots, P_n$ and all $(n-2)$-tuples of integers

$(s_1, s_2, \ldots, s_{n-2})$ where each s_j can take the values 1, 2, . . ., n. This correspondence is obtained as follows: let us remove from the tree the endpoint P_{i_1} with the least index, and let s_1 denote the index of the (unique) point which is connected by an edge with P_{i_1}. Repeating the same operation with the remaining tree of order $n-1$ we obtain s_2, s_3, etc. until only a single edge remains. PRÜFER has shown that the sequence $(s_1, s_2, \ldots, s_{n-2})$ determines the tree uniquely. Now clearly to the trees with exactly k endpoints there correspond sequences $(s_1, s_2, \ldots, s_{n-2})$ in which exactly $n-k$ of the numbers 1, 2, . . ., n occur at least once. As the failing numbers can be chosen in $\binom{n}{k}$ different ways, and the number of $(n-2)$-tuples formed from $n-k$ symbols which contain each of these $n-k$ symbols at least once is equal to $(n-k)!\,\mathfrak{S}_{n-2}^{n-k}$ (see e. g. [7]), it follows that

$$T(n,k) = \frac{n!}{k!}\,\mathfrak{S}_{n-2}^{n-k}\,. \tag{2.1}$$

According to the well-known recursion formula $\mathfrak{S}_n^k = k\,\mathfrak{S}_{n-1}^k + \mathfrak{S}_{n-1}^{k-1}$ (see e. g. [7], p. 169) it follows from (2.1) that

$$\frac{k}{n}\,T(n,k) = k\,T(n-1,k) + (n-k)\,T(n-1,k-1)\,. \tag{2.2}$$

Conversely (2.1) can be deduced from (2.2). The recursion formula (2.2) can be proved by a direct combinatorial argument, as has been remarked by VERA T. SÓS.

Let us calculate now the mean value and variance of the number of endpoints of a random tree of order n; here and in what follows if we speak about a random tree of order n we mean by this that we select at random a tree of order n with given vertices $P_1, \ldots, P_n$ so that all n^{n-2} possible trees are equiprobable. With this definition the number ν_n of endpoints of a random tree is a random variable. Let us denote by $\mathbf{M}\{\nu_n\}$ resp. $\mathbf{D}^2\{\nu_n\}$ the mean value resp. the variance of ν_n. Now according to the definition of Stirling's numbers of the second kind [see (9)] from (2.1) we have

$$\sum_{k=2}^{n-1} T(n,k)\,\frac{\binom{x}{n-k}}{\binom{n}{n-k}} = x^{n-2}\,. \tag{2.3}$$

The formula (2.3) can be considered as the (factorial) generating function of the sequence $T(n,k)$. Substituting now $x = n-1$ into (2.3) we obtain

$$\sum_{k=2}^{n-1} k\,T(n,k) = n(n-1)^{n-2}$$

and thus

$$\mathbf{M}\{\nu_n\} = n\left(1-\frac{1}{n}\right)^{n-2}. \tag{2.4}$$

Therefore we have

$$\lim_{n\to+\infty} \frac{\mathbf{M}\{\nu_n\}}{n} = \frac{1}{e} . \tag{2.5}$$

This can be expressed by saying that a random tree of order n has in the average approximately n/e endpoints.

To calculate the variance of ν_n let us substitute $x = n - 2$ into (2.2). We obtain

$$\sum_{k=2}^{n-1} k(k-1)\, T(n,k) = n(n-1)\,(n-2)^{n-2} \tag{2.6}$$

and thus

$$\sum_{k=2}^{n-1} k^2\, T(n,k) = n(n-1)\,(n-2)^{n-2} + n(n-1)^{n-2} . \tag{2.7}$$

It follows that

$$\mathbf{D}^2\{\nu_n\} = n(n-1)\left(1-\frac{2}{n}\right)^{n-2} + n\left(1-\frac{1}{n}\right)^{n-2} - n^2\left(1-\frac{1}{n}\right)^{2n-4} . \tag{2.8}$$

This implies

$$\lim_{n\to+\infty} \frac{\mathbf{D}^2\{\nu_n\}}{n} = \frac{e-2}{e^2} . \tag{2.9}$$

Thus the variance of ν_n is asymptotically equal to $n(e-2)/e^2$.

Now we shall prove that ν_n is in the limit normally distributed ; more exactly we shall prove

$$\lim_{n\to+\infty} \mathbf{P}\left\{\frac{\nu_n - \frac{n}{e}}{\frac{1}{e}\sqrt{n(e-2)}} < x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du \tag{2.10}$$

where $\mathbf{P}\{\ldots\}$ denotes the probability of the event in the brackets.

To prove (2.10), by a well-known theorem of probability theory, it suffices to show that the characteristic function of

$$\frac{\nu_n - \frac{n}{e}}{\frac{1}{e}\sqrt{n(e-2)}}$$

tends to the characteristic function of the normal distribution, that is for every real t we have

$$\lim_{n\to+\infty} \sum_{k=2}^{n-1} \frac{T(n,k)}{n^{n-2}}\, e^{\frac{it(ek-n)}{\sqrt{n(e-2)}}} = e^{-\frac{t^2}{2}} . \tag{2.11}$$

Now (2.11) can be obtained as follows: Let us substitute into (2.2) $x = n - it\sqrt{n}$. We obtain

$$\sum_{k=2}^{n-1} \frac{T(n,k)}{n^{n-2}} \prod_{j=k+1}^{n} \left(1 - \frac{it\sqrt{n}}{j}\right) = \left(1 - \frac{it}{\sqrt{n}}\right)^{n-2}. \tag{2.12}$$

If

$$\left|k - \frac{n}{e}\right| < n^{\beta}$$

where $1/2 < \beta < 1$, we have

$$\log \prod_{j=0}^{n-k-1} \left(1 - \frac{it\sqrt{n}}{n-j}\right) = -it\sqrt{n} + \frac{ite}{\sqrt{n}}\left(k - \frac{n}{e}\right) + \frac{t^2(e-1)}{2} + o(1) . \tag{2.13}$$

By the inequality of Chebysheff

$$\sum_{\left|k-\frac{n}{e}\right|>n^{\beta}} \frac{T(n,k)}{n^{n-2}} = O\left(\frac{1}{n^{2\beta-1}}\right) \tag{2.14}$$

and according to Stirling's formula for complex argument the factor

$$\prod_{j=k+1}^{n} \left(1 - \frac{it\sqrt{n}}{j}\right)$$

remains bounded; thus it follows that

$$\lim_{n\to+\infty} \left|\left(1 - \frac{it}{\sqrt{n}}\right)^{n-2} e^{it\sqrt{n}} - e^{\frac{t^2(e-1)}{2}} \sum_{k=2}^{n-1} \frac{T(n,k)}{n^{n-2}} e^{\frac{ite}{\sqrt{n}}\left(k-\frac{n}{e}\right)}\right| = 0 . \tag{2.15}$$

Now we have

$$\lim_{n\to+\infty} \left(1 - \frac{it}{\sqrt{n}}\right)^{n-2} e^{it\sqrt{n}} = e^{\frac{t^2}{2}}. \tag{2.16}$$

Thus it follows that

$$\lim_{n\to+\infty} \sum_{k=2}^{n} \frac{T(n,k)}{n^{n-2}} e^{\frac{ite}{\sqrt{n}}\left(k-\frac{n}{e}\right)} = e^{-\frac{t^2(\ -2)}{2}} . \tag{2.17}$$

Substituting $\frac{t}{\sqrt{e-2}}$ instead of t, (2.11) and therefore (2.10) follows.

(Received 30 October, 1958.)

Added in proof 5 May, 1959.

In a recent paper [8] L. E. Clarke gave a new proof of Cayley's formula (1). Clarke proves first that if $C_{n,k}$ denotes the number of trees with n labelled vertices $P_1, P_2, \ldots, P_n$ such that exactly k edges end at the vertex P_n, then we have

$$C_{n,k} = \binom{n-2}{k-1} (n-1)^{n-k-1} \qquad (k = 1,2, \ldots, n-1). \tag{2.18}$$

CLARKE proves (2.18) by giving a recursion formula for $C_{n,k}$, and deduces (1) from (2.18) by remarking that

$$T_n = \sum_{k=1}^{n-1} C_{n,k}. \tag{2.19}$$

It should be mentioned that (2.18) is a simple consequence of CAYLEY's formula (2). As a matter of fact, the k vertices which are connected with P_n can be chosen in $\binom{n-1}{k}$ ways among the vertices $P_1, P_2, \ldots, P_{n-1}$ and after this choice has been made, there remain still $T_{n-k,k}$ possible choices to form the tree in question. Thus we have

$$C_{n,k} = \binom{n-1}{k} T_{n-k,k} \tag{2.20}$$

and therefore by (2)

$$C_{n,k} = \binom{n-1}{k} k(n-1)^{n-k-2} = \binom{n-2}{k-1} (n-1)^{n-k-1}. \tag{2.21}$$

in accordance with (2.18).

REFERENCES

[1] CAYLEY, A. : ,,A theorem on trees". *Quarterly Journal of Pure and Applied Mathematics* **23** (1889) 376—378. See also *The Collected Papers of A. Cayley*, Cambridge, 1897, Vol. 13., pp. 26—28.

[2] DZIOBEK, O. : ,,Eine Formel der Substitutionstheorie", *Sitzungsberichte der Berliner Mathematischen Gesellschaft* **17** (1917) 64—67.

[3] PÓLYA, G. : ,,Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen." *Acta Mathematica (Uppsala)* **68** (1937) 145—255.

[4] PRÜFER, A. : ,,Neuer Beweis eines Satzes über Permutationen" *Archiv für Mathematik und Physik* **27** (1918) 142—144.

[5] DÉNES, J.: ,,The representation of a permutation as the product of a minimal number of transpositions and its connection with the theory of graphs". *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* **4** (1959) 63—71.

[6] PÓLYA, G.—SZEGŐ, G. : *Aufgaben und Lehrsätze aus der Analysis.* Springer, Berlin, 1925, Vol. I., p. 125.

[7] JORDAN, CH. : *Calculus of finite differences.* Budapest, 1939, p. 177.

[8] CLARKE, L. E. : ,,On Cayley's formula for counting trees", *Journal of the London Mathematical Society* **33** (1958) 411—475.

MEGJEGYZÉSEK A ,,FÁK" ELMÉLETÉHEZ

RÉNYI A.

Kivonat

,,Fa" alatt CAYLEY [1] nyomán környélküli összefüggő gráfot értünk. CAYLEY bebizonyította, hogy az n (számozott) szögponttal bíró különböző fák száma

$$T_n = n^{n-2}, \tag{1}$$

míg azon $n + k - 1$ (számozott) súlyponttal bíró gráfok száma, amelyek k idegen fából állnak oly módon, hogy k kijelölt pont különböző fákhoz tartozik,

$$T_{n,k} = k(n+k-1)^{n-2}. \tag{2}$$

Az (1) képletre CAYLEYn kívül O. DZIOBEK [2], PÓLYA GYÖRGY [3] és mások adtak más bizonyításokat. (1) legegyszerűbb és legelegánsabb bizonyítása PRÜFERtől származik [4]. Az 1. §-ban rámutat a szerző, hogy a (2) összefüggés (amely CAYLEYnél bizonyítás nélkül szerepel) egyszerűen bebizonyítható DZIOBEK módszerével.

Jelölje $G_k(n)$ azon n (számozott) szögpontú különböző gráfok számát, amelyek k idegen fából állnak, (elejtve a CAYLEY által tett megszorítást, hogy k adott pont különböző fákhoz tartozzék). Az 1 §-ban bebizonyítja a szerző, hogy

$$G_k(n) = \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} (k+j)! \binom{n-1}{k+j-1} \left(-\frac{1}{2}\right)^j n^{n-k-j} . \tag{3}$$

A (3) képlet segítségével $G_k(n)$ értékére, ha k rögzítve van és $n \to +\infty$ aszimptotikus képlet adható meg.

A 2. §. azzal a kérdéssel foglalkozik, hogy hány olyan n szögpontú fa van, amelynek pontosan r végpontja van. Ezt a számot $T(n, r)$-rel jelölve PRÜFER módszerével kimutatható, hogy

$$T(n, r) = \frac{n!}{r!} \mathfrak{S}_{n-2}^{n-r} , \tag{8}$$

ahol $\mathfrak{S}_N^m$ másodfajú Stirling-számokat jelöli, vagyis a $\mathfrak{S}_N^m$ számok az

$$x^N = \sum_{m=1}^{N} \mathfrak{S}_N^m x(x-1) \ldots (x-m+1) \tag{9}$$

összefüggés által vannak definiálva.

A (8) explicit képlet segítségével a szerző kimutatja, hogy egy találomra választott n (számozott) szögpontú fa végpontjainak száma határértékben normális eloszlású n/e várható értékkel és $\sqrt{(e-2)n/e}$ szórással, ha $n \to +\infty$.

ЗАМЕЧАНИЯ К ТЕОРИИ »ДЕРЕВЬЕВ«

A. RÉNYI

Резюме

Следуя CAYLEY [1] связанный граф, не содержащий окружностей называется «деревом». CAYLEY доказал, что число деревьев, имеющих n (нумерированных) угловых точек, равно

$$T_n = n^{n-2} \tag{1}$$

в то время, как число графов с $n+k-1$ (нумерированными) точками состоящих из k деревьев так, что k указанных точек принадлежат различным деревьям, равно

$$T_{n,k} = k(n+k-1)^{n-2} . \tag{2}$$

Для формулы (1) кроме CAYLEY другие доказательства дали также O. DZIOBEK [2] и G. PÓLYA [3]. Самое простое и элегантное доказательство (1)

принадлежит Prüfer-у [4]. В 1. §-е автор показывает, что соотношение (2) (которое у Cayley приведено без доказательства) также может быть доказано, методом Dziobek-а.

Обозначим через $G_k(n)$ число тех различных графов с n (нумерированными) угловыми точками, которые состоят из k деревьев (отказавшись от условия Cayley, согласно которому k данных точек принадлежат различным деревьям). В 1. §-е автор доказывает, что

$$G_k(n) = \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} (k+j)! \binom{n-1}{k+j-1} \left(-\frac{1}{2}\right)^j n^{n-k-j} \quad . \tag{3}$$

С помощью формулы (3) для значения $G_k(n)$ при фиксированном k и $n \to \infty$ может быть дана асимптотическая формула.

§ 2. занимается следующим вопросом: сколько существует таких деревьев с n угловыми точками, которые имеют точно r конечных точек. Обозначая это число через $T(n, r)$, методом Prüfer-а можно доказать, что

$$T_{n,r} = \frac{n!}{r!} \mathfrak{S}_{n-2}^{n-r} \tag{8}$$

где $\mathfrak{S}_N^m$ обозначают числа Stirling-а второго рода, т. е. определяются соотношением

$$x^N = \sum_{m=I}^{N} \mathfrak{S}_N^m x(x-1) \ldots (x-m+1) \quad . \tag{9}$$

С помощью явной формулы (8) автор доказывает, что число конечных точек случайно выбранного дерева с n (нумерированными) угловыми точками в пределе имеет нормальное распределение с математическим ожиданием n/e и дисперсией $\frac{(e-2)\,n}{e^2}$ если $n \to \infty$.