NOTE ON THE PÓLYA AND OTTER FORMULAS FOR ENUMERATING TREES

Frank Harary

The purpose of this note is to demonstrate that Otter's elegant formulation [2] of the number of trees may be derived directly from Pólya's equations [3] without any additional theory of trees.

1. PÓLYA'S EQUATIONS

A *tree* is a connected graph with no cycles (see König [1, Chap. V]). A *branch* of a tree, determined by a vertex P and a line PQ, consists of all vertices and lines reachable from P by paths starting with the line PQ. The *weight of a branch* of a tree is the number of lines it contains. The *weight of a vertex* P of a tree is the maximum of the weights of the branches at P. The *center of mass* of a tree consists of all vertices of minimum weight. It is well known that the center of mass of a tree consists either of a single vertex or of two adjacent vertices.

Two trees are *isomorphic* if there exists a one-to-one adjacency preserving correspondence between their sets of vertices. A *rooted tree* is a tree in which one vertex (the *root*) is distinguished. Two rooted trees are isomorphic if there exists an isomorphism between them which maps one root onto the other. Let t_n be the number of (nonisomorphic) trees with n lines, or equivalently with $n + 1$ vertices, and let T_n be the corresponding number of rooted trees. Let $T(x) = \Sigma_{n=0}^{\infty} T_n x^n$; then the generating series T(x) satisfies Cayley's functional equation (see [3, p. 149]):

$$(1) \qquad T(x) = \exp \sum_{r=1}^{\infty} \frac{x^r}{r} T(x^r),$$

from which the numbers T_n can be computed recursively. Otter [2, equations (6) and (7)] obtains explicitly the recursion equations implicit in (1).

Let t_n' and t_n'' be the number of trees having n lines, whose center of mass consists of one and two vertices respectively, so that $t_n = t_n' + t_n''$. Pólya [3, pp. 207 and 203] develops formulas (2) and (3) below, which give t_n' and t_n'' in terms of the known numbers T_n.

By $\text{coef}_n\{f(x)\}$ we mean the coefficient of x^n in the power series of f(x). Using this notation, Pólya obtains

$$(2) \qquad t_n' = \text{coef}_n\{(1 - x)^{-T_0}(1 - x^2)^{-T_1} \dots (1 - x^{n+1})^{-T_m}\},$$

where $m = [n/2] - 1$. One sees readily that

$$(3) \qquad t_{2m}'' = 0, \quad t_{2m+1}'' = \binom{T_m + 1}{2}.$$

Received September 27, 1955.

The preparation of this note was supported in part by a grant from the Rockefeller Foundation to the Research Center for Group Dynamics of the University of Michigan. I thank Professor G. E. Uhlenbeck for many helpful comments.

2. OTTER'S EQUATION

An *automorphism* of a graph G is an isomorphism of G with itself. Two vertices of a graph are *similar* if there exists an automorphism of G mapping one onto the other. Similarity of two lines is defined analogously. An *exceptional line* of a graph is one whose vertices are similar. For any tree, let p, k, and k_e be the number of dissimilar vertices, lines, and exceptional lines, respectively. Otter [2] discovered and proved the dissimilarity characteristic equation for trees:

$$p - (k - k_e) = 1; \tag{4}$$

and from this equation, he succeeded in deriving the following concise equation expressing the generating series $t(x) = \Sigma_{n=0}^{\infty} t_n x^n$ in terms of the known function $T(x)$.

$$t(x) = T(x) - \frac{x}{2}[T^2(x) - T(x^2)]. \tag{5}$$

3. DERIVATION OF OTTER'S EQUATION FROM THOSE OF PÓLYA

We now proceed from (2) and (3) directly to (5) without using (4).

Let $T_m(x) = \Sigma_{n=0}^{m} T_n x^n$ be the first $m + 1$ terms of the series $T(x)$. Then (2) can be written in the form:

$$t_n' = \text{coef}_n\left\{\exp \sum_{r=1}^{\infty} \frac{x^r}{r} T_m(x^r)\right\} \quad \left(m = \left[\frac{n}{2}\right] - 1\right). \tag{2'}$$

The equivalence of (2') to (2) is shown as follows:

$$\begin{aligned}\exp \sum_{r=1}^{\infty} \frac{x^r}{r} T_m(x^r) &= \exp \sum_{r=1}^{\infty} \frac{x^r}{r} \sum_{s=0}^{m} T_s x^{rs} \\ &= \exp \sum_{s=0}^{m} T_s \sum_{r=1}^{\infty} \frac{x^{r(s+1)}}{r} \\ &= \exp\left(-\sum_{s=0}^{m} T_s \log(1 - x^{s+1})\right) \\ &= (1 - x)^{-T_0}(1 - x^2)^{-T_1} \dots (1 - x^{m+1})^{-T_m}.\end{aligned}$$

Let $t'(x) = \Sigma_{n=0}^{\infty} t_n' x^n$ and $t''(x) = \Sigma_{n=0}^{\infty} t_n'' x^n$, so that $t(x) = t'(x) + t''(x)$. Then it follows from (3) that

$$t''(x) = \sum_{m=0}^{\infty} \frac{1}{2} T_m(T_m + 1) x^{2m+1},$$

that is,

$$t''(x) = \frac{x}{2} \sum_{m=0}^{\infty} T_m^2 x^{2m} + \frac{x}{2} T(x^2). \tag{6}$$

To derive an expression for $t'(x)$, we note that

$$\exp \sum_{r=1}^{\infty} \frac{x^r}{r} T_m(x^r) = \exp\left\{- \sum_{s=0}^{m} T_s \log(1 - x^{s+1})\right\}$$

$$= T(x) \exp \sum_{s=m+1}^{\infty} T_s \log(1 - x^{s+1}).$$

Therefore, equation (2') with $m = [n/2] - 1$ becomes

$$t'_n = \text{coef}_n\left\{T(x)\left[1 + \sum_{s=m+1}^{\infty} T_s \log(1 - x^{s+1}) + \frac{1}{2!}\left(\sum_{s=m+1}^{\infty} T_s \log(1 - x^{s+1})\right)^2 + \ldots\right]\right\}. \tag{7}$$

But

$$\sum_{s=m+1}^{\infty} T_s \log(1 - x^{s+1}) = - \sum_{s=m+1}^{\infty} \sum_{h=1}^{\infty} T_s \frac{x^{h(s+1)}}{h},$$

so that $m + 2$ is the smallesr degree of x which occurs in this sum. Hence $2m + 4$ is the smallest degree of x in the third term inside the brackets of (7). But $2m + 4 > n$, since $m = [n/2] - 1$, and thus there is no contribution to the coefficient of x^n after the first two terms in the brackets of (7). Therefore we see that

$$t'_n = \text{coef}_n\left\{T(x)\left[1 + \sum_{s=m+1}^{\infty} T_s \log(1 - x^{s+1})\right]\right\}$$

$$= \text{coef}_n\left\{T(x)\left[1 - \sum_{s=m+1}^{\infty} \sum_{h=1}^{\infty} T_s \frac{x^{h(s+1)}}{h}\right]\right\}.$$

Now by exactly the same argument, all terms after $h = 1$ in the last expression can be ignored, since they do not contribute to the coefficient of x^n, and we have

$$t'_n = \text{coef}_n\left\{T(x)\left[1 - \sum_{s=m+1}^{\infty} T_s x^{s+1}\right]\right\}$$

$$= \text{coef}_n\left\{T(x)\left[1 - xT(x) + \sum_{s=0}^{m} T_s x^{s+1}\right]\right\}.$$

From this last equation, one sees at once that

$$t'_n = \sum_{s=0}^{m} T_s T_{n-s-1} + \text{coef}_n\left\{T(x)[1 - xT(x)]\right\},$$

that is,

$$t'(x) = T(x)\,[1 - xT(x)] + \sum_{n=0}^{\infty} x^n \sum_{s=0}^{m} T_s\, T_{n-s-1}.$$

But interchanging the order of summation and using the fact that $m = [n/2] - 1$, one obtains the following simplification of the double series in the preceding equation.

$$\sum_{n=0}^{\infty} x^n \sum_{s=0}^{m} T_s T_{n-s-1} = \sum_{s=0}^{m} \sum_{n=2s+2}^{\infty} T_s T_{n-s-1}\, x^n$$

$$= \sum_{s=0}^{\infty} T_s \sum_{k=s+1}^{\infty} T_k x^{k+s+1}$$

$$= x \sum_{s=0}^{\infty} T_s \sum_{k=s+1}^{\infty} T_k x^{k+s}$$

$$= \frac{x}{2} T^2(x) - \frac{x}{2} \sum_{s=0}^{\infty} T_s^2 x^{2s}.$$

The last equality follows by the use of a conventional summation technique. Collecting these results, we have

$$t'(x) = T(x) - \frac{x}{2} T^2(x) - \frac{x}{2} \sum_{s=0}^{\infty} T_s^2 x^{2s}\,. \tag{8}$$

Finally, addition of equations (6) and (8) gives (5).

REFERENCES

1. D. König, *Theorie der endlichen und unendlichen Graphen,* Leipzig, 1936; reprinted New York, 1950.

2. R. Otter, *The number of trees*, Ann. of Math. (2) 49 (1948), 583-599.

3. G. Pólya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen, und chemische Verbindungen,* Acta Math. 68 (1937), 145-254.

University of Michigan