

ENUMERATION OF FORESTS IN A GRAPH

C. J. LIU AND YUTZE CHOW

ABSTRACT. The enumeration of forests of different orders in a graph is carried out by a procedure that involves formal sums and certain annihilation operators on the space of such sums. The results here extend the well-known matrix-tree theorem to the general case of forests.

Introduction. In pursuing the enumeration of connected spanning subgraphs [1], each containing a fixed number of cycles, for a planar graph, we introduced a formal procedure together with “annihilation” operators acting on face matrices. From the viewpoint of graphic duality, the face matrix is just the Kirchhoff matrix of the *dual* graph [2]. In other words, a spanning subgraph with m independent cycles corresponds to a $(m + 1)$ -forest in the dual graph. It is therefore obvious that our previous result naturally applies to the counting of spanning *forests* for a planar graph. In this article, by a further generalization, the use of the formal procedure via the annihilation operators¹ enables us to solve the enumeration problems for *nonplanar* graphs. The method provides again, as in the planar case, some slick expressions for the various sums appearing in the enumeration formulae. Especially in the case of spanning forests of all possible component orders, this approach leads to an expression which is simply the Kirchhoff matrix acted upon by an exponentiation of the annihilation operators we introduced.

It is a pleasure to thank the referee for many valuable suggestions which resulted in a substantial improvement of the manuscript.

Main theorems. The graph G to be considered will be connected and labelled. Denote by A the adjacency matrix of G . The corresponding Kirchhoff matrix M is obtained from A by replacing in $-A$ each diagonal entry by the degree of its corresponding vertex, i.e. the i th diagonal entry is identified with the degree of the i th vertex v_i . Denote by $M(i)$ the i th principal submatrix, i.e., with the i th row and column deleted from M . Similarly $M(i, j)$ denotes the principal submatrix, obtained from M by deleting both the i th row and column as well as the j th row and column. Let n be the total number of vertices in G . It is convenient to introduce the following conventions for such matrices:

$$(1) \quad M(i_1, \dots, i_n) = 1, \quad \text{if all } i_j \text{ are distinct}$$

Received by the editors July 16, 1980 and, in revised form, February 19, 1981.

1980 *Mathematics Subject Classification.* Primary 05C05, 05C30; Secondary 05A15.

¹The formal procedure used in the present paper differs from that of [1] in that the present method is much more direct. However, the method formulated in [1] has the advantage of being graphically intuitive. In spite of the apparent difference, they are essentially equivalent procedures.

© 1981 American Mathematical Society
0002-9939/81/0000-0550/\$02.00

and

$$(2) \quad \mathbf{M}(i_1, \dots, i_m) = 0, \quad \text{if some } i_j = i_k \text{ for } j \neq k,$$

where 0 and 1 are just numbers. We now define the formal process:

Denote $n^* \equiv \{1, 2, \dots, n - 1\}$. Construct the space V of formal sums generated by the collection of all subsets of n^* , with coefficients in \mathbf{R} . Define a real-valued linear function μ on the space V by

$$(3) \quad \mu: S \mapsto \det \mathbf{M}({}_c S \cup n)$$

where $S \subset n^*$ and ${}_c S$ is the complement of S in n^* . As an example,

$$(4) \quad \{i, j\} \mapsto \det \mathbf{M}(1, \dots, \hat{i}, \dots, \hat{j}, \dots, n), \quad \{i, j\} \subset n^*,$$

where \hat{i} denotes the deletion of argument i . Hereafter we shall say that S generates $\mu(S)$, for convenience. In particular, since $\mu(\phi) = 1$ and $\mu(n^*) = T^{(1)} \equiv$ the total number of spanning trees, by Kirchoff's theorem, we shall say that ϕ generates 1 and n^* generates $T^{(1)}$, etc.

It is useful to introduce the following \mathbf{R} -linear "annihilation" operators, on V , defined by

$$(5) \quad \alpha_i: S \mapsto S - \{i\}, \quad \text{if } i \in S \subset n^*,$$

$$(6) \quad S \mapsto 0, \quad \text{if } i \notin S \subset n^*.$$

Then the following operator plays an essential role in our theorems to follow:

$$(7) \quad \alpha \equiv \sum_{i=1}^{n-1} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n-1} a_{ij} \alpha_i \alpha_j$$

where a_{ij} are the entries of the adjacency matrix \mathbf{A} of the graph G .

THEOREM I. *The formal sum*

$$(8) \quad \mathbf{T}^{(m)} = \alpha^{m-1} n^* / (m - 1)!$$

generates $T^{(m)}$, the total number of m -forests in G .

PROOF. Use mathematical induction on m . For $m = 1$, $\mathbf{T}^{(1)}$ obviously generates $T^{(1)} \equiv \mu(n^*)$, the total number of 1-forests (i.e. spanning trees). For $m > 1$, assume that the theorem is true for $m = k$. Instead of the original graph G , we now, by the induction hypothesis, apply the theorem to the graphs corresponding to terms ${}_c \{i\}$ and ${}_c \{i, j\}$ with $\{i, j\} \subset n^*$. Graphically, we may consider that n^* corresponds to the original graph G . Then ${}_c \{i\}$ corresponds to the modified graph obtained from G by identifying the vertices v_i and v_n (regardless of whether they are adjacent or not). Similar graphic implications obviously apply to ${}_c \{i, j\}$, etc. Hence the formal sum

$$(9) \quad \alpha^{k-1}({}_c \{i\}) / (k - 1)!$$

generates the total number of $(k + 1)$ -forests, each of which does not contain any path joining v_i and v_n . Thus

$$(10) \quad \mu \left\{ \frac{1}{(k - 1)!} \alpha^{k-1} \sum_i {}_c \{i\} \right\} = \sum_{F_{k+1}} \bar{v}_n(F_{k+1})$$

where F_{k+1} runs over all $(k + 1)$ -forests in G in the above sum and $\bar{v}_n(F)$ is the total number of vertices not connected to v_n by any path in F . Similarly, for $i \neq j$, the formal sum

$$(11) \quad a_{ij} \alpha^{k-1} ({}_c\{i, j\}) / (k - 1)!$$

generates the total number of $(k + 1)$ -forests, in each of which v_i and v_j are adjacent but are not path-connected to v_n . Therefore we have

$$(12) \quad \mu \left\{ \frac{1}{(k - 1)!} \alpha^{k-1} \left[\frac{1}{2} \sum_{i,j} a_{ij} \cdot {}_c\{i, j\} \right] \right\} = \sum_{F_{k+1}} \bar{e}_n(F_{k+1})$$

where $\bar{e}_n(F)$ is the total number of edges not connected to v_n by any path in F . Consequently,

$$(13) \quad \mu \left\{ \frac{1}{(k - 1)!} \alpha^{k-1} \left[\sum_i {}_c\{i\} - \frac{1}{2} \sum_{i,j} a_{ij} \cdot {}_c\{i, j\} \right] \right\} \\ = \sum_{F_{k+1}} [\bar{v}_n(F_{k+1}) - \bar{e}_n(F_{k+1})] = kT^{(k+1)},$$

which completes the induction. \square

COROLLARY. *Expressed in the language of the usual generating function, we have (for an indeterminate x):*

$$(14) \quad \mu([\exp(\alpha x)]n^*) = \sum_{m=1}^{\infty} T^{(m)} x^{m-1}.$$

We note that since $\alpha^m n^* = 0$ for $m \geq n$, the sum on the right-hand side of (14) is finite.

THEOREM II. *The total number of m -forests $T^{(m)}$ in the graph G is given by*

$$(15) \quad T^{(m)} = \sum_{r=0}^{m-1} t_r^{(m)}$$

where

$$(16) \quad t_r^{(m)} = \frac{(-1)^r}{r!} \sum_{i_1 < i_2} \sum_{i_3 < i_4} \dots \sum_{i_{2r-1} < i_{2r}} \sum_{i_{2r+1} < \dots < i_{m+r-1}} a_{i_1 i_2} \\ \dots a_{i_{2r-1} i_{2r}} \det \mathbf{M}(i_1, \dots, i_{m+r-1}, n)$$

with distinct i_k 's ($k = 1, \dots, m + r - 1$) whose summation limits run from 1 to $n - 1$. In particular, for $r = 0$, we have

$$(17) \quad t_0^{(m)} = \sum_{i_1 < \dots < i_{m-1}} \det \mathbf{M}(i_1, \dots, i_{m-1}, n).$$

PROOF. Using the fact $\alpha_i^2 = 0$, the multinomial expansion yields

$$(18) \quad \frac{1}{(m - 1)!} \alpha^{m-1} = \sum_{r=0}^{m-1} \frac{(-1)^r}{r!} \sum_{i_1 < i_2} \dots \sum_{i_{2r-1} < i_{2r}} \sum_{i_{2r+1} < \dots < i_{m+r-1}} a_{i_1 i_2} \\ \dots a_{i_{2r-1} i_{2r}} \alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{m+r-1}}$$

with distinct i_k 's ($k = 1, \dots, m + r - 1$). Then (15) follows from Theorem I. \square

Theorem II can be alternatively stated (as suggested by the referee) in terms of the concept of "matchings" (a *matching* is a collection of independent edges of G). To do this, we first introduce the following notation: given a set $S \equiv \{i_1, \dots, i_r\} \subset n^*$, we write down the corresponding vertex set $v_S \equiv \{v_{i_1}, \dots, v_{i_r}\}$. Denote by $\nu_r(v_S)$ the total number of *matchings* (in G) each consisting of r edges whose endpoints are all in v_S . For convention, we set $\nu_0(v_S) = 1$. Denote by $\tau^{(m+r)}(v_S \cup v_n)$ the total number of $(m+r)$ -forests, in each of which no two different elements of $v_S \cup v_n$ can belong to the same tree. Then we have the following corollary:

COROLLARY.

$$(19) \quad T^{(m)} = \sum_{r=0}^{m-1} (-1)^r \sum_{\substack{S \subset n^* \\ |S|=m+r-1}} \nu_r(v_S) \tau^{(m+r)}(v_S \cup v_n).$$

PROOF. It follows from Theorem II that

$$(20) \quad T^{(m)} = \sum_{r=0}^{m-1} (-1)^r \sum_{\substack{S \subset n^* \\ |S|=m+r-1}} \nu_r(v_S) T^{(1)}(G_S)$$

where G_S is the graph obtained from G by identifying all the vertices in $v_S \cup v_n$. But

$$(21) \quad T^{(1)}(G_S) = \tau^{(m+r)}(v_S \cup v_n).$$

This completes the proof. We note that the proof of (19) can also be carried out by inclusion-exclusion. \square

Further results: Complete graphs. We announce here (to be published elsewhere) the following result concerning *complete* graphs as an application of our present method:

$$(22) \quad T^{(m)} = n^{n-m-1} (n-1)! \sum_{r=0}^{m-1} \left(\frac{-1}{2n}\right)^r \frac{m+r}{r! (m-r-1)! (n-m-r)!}.$$

REFERENCES

1. C. J. Liu and Yutze Chow, *Enumeration of connected spanning subgraphs of a planar theory*, Acta Math. (to appear).
2. F. Harary, *Graph theory*, Addison-Wesley, Reading, Mass., 1969.

DEPARTMENT OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE, MARQUETTE UNIVERSITY, MILWAUKEE, WISCONSIN 53233

DEPARTMENT OF PHYSICS, UNIVERSITY OF WISCONSIN, MILWAUKEE, WISCONSIN 53201