

THE NUMBER OF LINEAR, DIRECTED, ROOTED, AND CONNECTED GRAPHS

BY

FRANK HARARY⁽¹⁾

1. **Introduction.** Seifert and Threlfall [20, p. 4] have described the principal problem of topology as " . . . zu entscheiden, ob zwei vorgelegte Figuren homöomorph sind und wo möglich alle Klassen nichthomöomorpher Figuren aufzuzählen." Our object is to obtain the number of nonisomorphic linear graphs with p points and k lines, and also to count various kinds of generalizations of graphs. These include directed graphs (digraphs), rooted graphs, multiply rooted graphs, and two other generalizations which will be called graphs of strength s and graphs of type t .

The fundamental theorem used to secure these results is due to Pólya [15] and will be reviewed very briefly in the next section. The author is happy to take this opportunity to thank Professor Pólya for kindly permitting the presentation of his unpublished formula for the number of linear graphs in this paper.

The form of the solution in every case will be the counting polynomial. Thus $g_p(x)$ defined by:

$$g_p(x) = \sum_{k=0}^{p(p-1)/2} g_{pk} x^k$$

where g_{pk} is the number of graphs with p points and k lines, serves to count all graphs of p points.

After counting several generalizations of graphs, we shall derive a formula for the number of connected graphs of any given topological type in terms of the total number of (connected as well as disconnected) graphs of this type. The number of connected graphs in terms of the total number of graphs, which first appeared in Riddell [16] and then in Riddell and Uhlenbeck [18], as well as the number of weakly connected digraphs obtained by Pólya (unpublished) will follow as corollaries. A simple variation of the method enables one to count the rooted connected graphs of any given type in terms of the unrooted connected ones and the total number of such graphs. To illustrate the method, the number of forests and rooted forests will be found in terms of the known number of trees. The final section states some unsolved combinatorial problems.

Presented to the Society, September 2, 1952; received by the editors January 20, 1954.

⁽¹⁾ The preparation of this paper was supported by a grant from the Rockefeller Foundation to the Research Center for Group Dynamics of the University of Michigan. The author wishes to thank R. Z. Norman and G. E. Uhlenbeck for critical readings of the manuscript.

As historical background, we review briefly some previous results on the enumeration of graphs of various kinds which are of interest in several different fields of study. Thus Cayley [3] first found formulas for the number of trees and rooted trees. A more elegant formula for the number of trees was found by Pólya [15] and a still neater formulation was obtained by Otter [14], whose proof was subsequently simplified in [7] and [9] using the methods of [15], i.e., Pólya's Theorem.

Foster [5] and Riordan and Shannon [19] have provided solutions for some counting problems in electrical network and communication theory.

Carnap [2] asked the well-known question as to the number of nonisomorphic binary relations on a field of p elements. This problem was solved by Davis [4] who also found the number of irreflexive and symmetric binary relations on p elements. Using the familiar correspondence between binary relations and digraphs (as in König [11]), Carnap's question can be restated graphically. The counting polynomial

$$\bar{g}_p(x) = \sum_{k=0}^{p(p-1)} \bar{g}_{pk} x^k$$

for digraphs of p points, obtained here, provides additional information to the question on relations. For the number \bar{g}_{pk} of digraphs with p points and k directed lines is the number of nonisomorphic irreflexive relations *with k ordered couples* on a field of p elements.

Cayley [3] recognized the possibility of studying the chemical problem of the number of isomers by making use of the notion of a tree. Lunn and Senior [12] discussed and solved some related questions through the use of the properties of permutation groups. Blair and Henze [1] succeeded in solving several such counting problems for organic chemistry by enumerating the class of trees and rooted trees with points of given maximum degree. In his classic publication Pólya [15] obtained all of the above recursion formulas, and also solved a wealth of other problems connected with chemical isomers.

Motivated by social psychological considerations [6] in which points denote persons, lines stand for relations, and graphs represent "social groups," we were led to the definitions of graphs of strength s , and of type t . These are counted in §6.

The Feynman-Dyson graphs which are appropriate for studying certain problems in quantum electrodynamics and general field theory were first enumerated by Hurst [10], who found some asymptotic numbers using the classical Laguerre and Hermite polynomials. These graphs were later counted more concisely in Riddell [17] by elementary combinatorial considerations.

Finally, the question of the number of Husimi trees was asked by Uhlenbeck [21]. Contributions to the solution of this problem were given in [7] and [9]. The final solution was obtained by Norman [13], using a point of

view mentioned in [8], by an ingenious application of Pólya's Theorem.

2. Pólya's Theorem. We shall state Pólya's Theorem (the Hauptsatz of [15]) in the form which will be used to derive the counting polynomials for the various kinds of graphs. The desired form is a specialization of Pólya's statement to one variable.

Let *figure* be an undefined term. To each figure there is assigned a non-negative integer called its *content*. Let ϕ_k denote the number of different figures of content k . Then the figure counting series $\phi(x)$ is defined by

$$(1) \quad \phi(x) = \sum_{k=0}^{\infty} \phi_k x^k.$$

Let G be a permutation group of degree s and order h . A *configuration* of length s is a sequence of s figures. The content of a configuration is the sum of the contents of its figures. Two configurations are *G-equivalent* if there is a permutation of G sending one into the other. Let F_k denote the number of G -inequivalent configurations of content k . The configuration counting series $F(x)$ is defined by

$$(2) \quad F(x) = \sum_{k=0}^{\infty} F_k x^k.$$

We shall call G the *configuration group* henceforth.

The object of Pólya's Theorem is to express $F(x)$ in terms of $\phi(x)$ and G . This is accomplished using the cycle index of G , defined as follows. Let h_{j_1, j_2, \dots, j_s} denote the number of elements of G of type (j_1, j_2, \dots, j_s) , i.e., having j_k cycles of length k , $k=1, 2, \dots, s$. Thus

$$(3) \quad 1 \cdot j_1 + 2 \cdot j_2 + \dots + s \cdot j_s = s.$$

Let f_1, f_2, \dots, f_s be s indeterminates. Then $Z(G)$, the *cycle index* of G , is defined, as in Pólya [15, p. 159], by:

$$(4) \quad Z(G) = \frac{1}{h} \sum_{(j)} h_{j_1, j_2, \dots, j_s} f_1^{j_1} f_2^{j_2} \dots f_s^{j_s},$$

where the sum is taken over all partitions (j) of s satisfying (3). For any function $f(x)$, let $Z(G, f(x))$ denote, as in [9], the function obtained from $Z(G)$ by replacing each indeterminate f_k by $f(x^k)$. Using these definitions, we are able to give a concise statement of:

PÓLYA'S THEOREM. *The configuration counting series is obtained by substituting the figure counting series into the cycle index of the configuration group. Symbolically,*

$$(5) \quad F(x) = Z(G, \phi(x)).$$

This theorem reduces the problem of finding the configuration counting

series to the determination of the figure counting series and the cycle index of the configuration group.

3. The number of graphs⁽²⁾. A *graph* may be defined as a set of p points together with some subset of the $p(p-1)/2$ lines joining them. Two points of a graph are *adjacent* if there is a line joining them. Two graphs are *isomorphic* if there is a one-to-one correspondence between their point sets which preserves adjacency.

Using the terminology of the preceding section, the counting of graphs is amenable to Pólya's Theorem by the following observation:

A graph is a configuration whose figures are pairs of points.

The content of a figure, i.e., a pair of points, is zero or one according respectively to whether there does not or does exist a line of the graph joining them. Hence the content of a graph is the number of lines it contains.

Let p be any positive integer. Let g_{pk} denote the number of nonisomorphic graphs having p points and k lines. The *counting polynomial* $g_p(x)$ for graphs of p points is defined by:

$$(6) \quad g_p(x) = \sum_{k=0}^{p(p-1)/2} g_{pk} x^k.$$

This is the desired configuration counting series. In order to apply Pólya's Theorem, we need to know both the figure counting series $\phi(x)$, and the cycle index of the configuration group. It follows at once that

$$(7) \quad \phi(x) = 1 + x.$$

The configuration group for graphs is *the full pair group on p points*. (This descriptive nomenclature is due to G. W. Ford.) This group R_p is a permutation group isomorphic as an abstract group to S_p , the symmetric group on p points, but is of degree $p(p-1)/2$. The objects permuted by R_p are pairs of points. The cycle index of R_p may be obtained by a simple calculation from that of S_p .

But even before finding $Z(R_p)$ explicitly, we apply Pólya's Theorem to obtain the form of the result:

$$(8) \quad g_p(x) = Z(R_p, 1 + x).$$

It is well known that $Z(S_p)$ may be obtained from (4) by setting $s = p$, $h = p!$, and

$$(9) \quad h_{i_1, i_2, \dots, i_p} = \frac{p!}{1^{i_1} j_1! 2^{i_2} j_2! \cdots p^{i_p} j_p!}.$$

In order to simplify and clarify the exposition for general p , we first illustrate

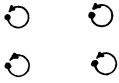

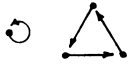
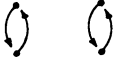
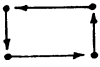
⁽²⁾ The result of this section (equations (8) and (10)) appeared in [16] and [18], also with Pólya's permission.

for $p=4$. Here we have

$$Z(S_4) = \frac{1}{24} (f_1^4 + 6f_1^2f_2 + 8f_1f_3 + 3f_2^2 + 6f_4).$$

Since the objects permuted by S_4 are points, the terms of this cycle index may be pictured as in the second column of Table 1. To avoid possible confusion, let us denote the variables appearing in $Z(R_p)$ by g_i , $i=1, \dots, p(p-1)/2$.

TABLE I

Term of $Z(S_4)$	Diagram	Term of $Z(R_4)$
f_1^4		g_1^6
$f_1^2f_2$		$g_1^2g_2^2$
f_1f_3		g_3^2
f_2^2		$g_1^2g_2^2$
f_4		g_2g_4

Since each point of the diagram for the term f_1^4 lies in S_4 -cycle of length one, it follows that each of the 6 pairs of points lies in an R_4 -cycle of length one. Thus the resulting term of $Z(R_4)$ is g_1^6 . In the diagram for $f_1^2f_2$, the pairs A, B and C, D are left fixed (points C and D are interchanged, but the pair of points is sent into itself). This accounts for the g_1^2 part of the corresponding term of $Z(R_4)$. However, the pair A, D is mapped under the cycles of the term $f_1^2f_2$ into the pair A, C , which is then returned to A, D , forming a cycle of length two in R_4 . Similarly the pairs B, C and B, D form another cycle of length two. Hence $g_1^2g_2^2$ is the term in $Z(R_4)$ corresponding to the term $f_1^2f_2$ in $Z(S_4)$. The remainder of Table 1 may be calculated similarly.

Collecting the results of Table 1, we see that

$$Z(R_4) = \frac{1}{24} (g_1^6 + 9g_1^2 g_2^2 + 8g_3^2 + 6g_2 g_4).$$

Substituting $1+x^i$ for g_i , as indicated in (8), and simplifying, we obtain the counting polynomial for all 4 point graphs:

$$g_4(x) = 1 + x + 2x^2 + 3x^3 + 2x^4 + x^5 + x^6.$$

This polynomial may also be verified pictorially, as in Fig. 1.

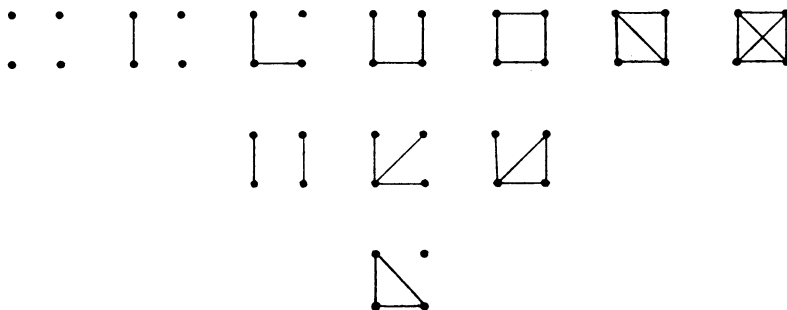


FIG. 1

The computation of $Z(R_p)$ is similar and will be sketched briefly. For convenience later, we shall write $Z_2(S_p) = Z(R_p)$. We shall indicate the correspondence between terms of $Z(S_p)$ and $Z_2(S_p)$ by

$$f_1^4 \rightarrow g_1^6, \quad f_1^2 f_2^2 \rightarrow g_1^2 g_2^2, \quad \text{etc.}$$

Combining (4) and (9), we see that $Z_2(S_p)$ will be determined once the term corresponding to $f_1^{j_1} f_2^{j_2} \cdots f_p^{j_p}$ is found. The contribution to $Z_2(S_p)$ from $f_1^{j_1} f_2^{j_2} \cdots f_p^{j_p}$ can be separated into two independent parts which are then multiplied to yield the result. The first part comes from those points lying on cycles of permutations of S_p of the same length; the second part is obtained from all remaining point pairs.

It is convenient to divide the first part into the contributions from the odd cycles and from the even cycles. The odd cycles yield (where $C_{n,r}$ denotes the binomial coefficient)

$$f_{2n+1}^{j_{2n+1}} \rightarrow g_{2n+1}^{n j_{2n+1} + (2n+1) C_{j_{2n+1}, 2}}.$$

From the even cycles, we obtain

$$f_{2n}^{j_{2n}} \rightarrow (g_n g_{2n})^{n-1} j_{2n}^{n-1} g_{2n}^{2n} C_{j_{2n}, 2}.$$

Finally, the contribution to $Z_2(S_p)$ from $f_q^{j_q} f_r^{j_r}$ not already calculated above is clearly $g_{m(q,r)}^{j_q j_r d(q,r)}$, where $m(q, r)$ and $d(q, r)$ denote least common multiple and greatest common divisor respectively. Collecting these results, we obtain

$$(10) \quad Z(R_p) = Z_2(S_p) = \frac{1}{p!} \sum_{(i)} \frac{p!}{\prod_{i=1}^p i! j_i!} \prod_{n=0}^{[(p-1)/2]} \frac{n! j_{2n+1} + (2n+1) C_{j_{2n+1}, 2}}{g_{2n+1}} \\ \prod_{n=1}^{[p/2]} (g_n g_{2n}^{n-1})^{j_{2n}} g_{2n}^{2n C_{j_{2n}, 2}} \cdot \prod_{1 \leq q < r \leq p} \frac{j_q^{j_r} d(q, r)}{g_m(q, r)}.$$

In view of (8), the substitution of $1+x^i$ for g_i in (10) yields the counting polynomial $g_p(x)$.

The explicit formulas for $g_p(x)$, $p=1$ to 7 are given by:

$$(11) \quad g(x, y) = \sum_{p=1}^{\infty} g_p(x) y^p = \sum_{p=1}^{\infty} \sum_{k=0}^{p(p-1)/2} g_{p,k} x^k y^p. \\ = y + y^2(1+x) + y^3(1+x+x^2+x^3) \\ = y^4(1+x+2x^2+3x^3+2x^4+x^5+x^6) \\ + y^5(1+x+2x^2+4x^3+6x^4+6x^5+6x^6+4x^7+2x^8+x^9+x^{10}) \\ + y^6(1+x+2x^2+5x^3+9x^4+15x^5+21x^6+24x^7+24x^8 \\ + 21x^9+15x^{10}+9x^{11}+5x^{12}+2x^{13}+x^{14}+x^{15}) \\ + y^7(1+x+2x^2+5x^3+10x^4+21x^5+41x^6+65x^7+97x^8 \\ + 131x^9+148x^{10}+148x^{11}+131x^{12}+97x^{13}+65x^{14}+41x^{15}+21x^{16} \\ + 10x^{17}+5x^{18}+2x^{19}+x^{20}+x^{21}) \\ + \dots$$

Let \mathcal{G} be a graph of p points. The *complement* of \mathcal{G} is the graph having the same points as \mathcal{G} , whose lines join exactly those pairs of points not joined in \mathcal{G} . Obviously, the operation of graph complementation is a one-to-one correspondence between the graphs of p points and k lines and the graphs of p points and $p(p-1)/2-k$ lines. Hence $g_{p,k} = g_{p,p(p-1)/2-k}$, accounting for the end-symmetry in the coefficients of $g_p(x)$.

4. The number of digraphs. Analogously, the basic observation necessary to the enumeration of digraphs by means of Pólya's Theorem is as follows: *A digraph is a configuration whose figures are ordered pairs of points.*

In this context, the content of a figure is one or zero in respective accordance with the existence or nonexistence of a directed line from the first member of the figure to its second member. Hence the figure counting series is again $1+x$.

Let \bar{g}_{pk} be the number of nonisomorphic digraphs with p points and k directed lines. The counting polynomial $\bar{g}_p(x)$ for digraphs of p points is given by:

$$(12) \quad \bar{g}_p(x) = \sum_{k=0}^{p(p-1)} \bar{g}_{pk} x^k.$$

The configuration group for digraphs will be called the *ordered pair group on p points*, Q_p , which is a permutation group of degree $p(p-1)$. For convenience we shall write $\bar{Z}_2(S_p) = Z(Q_p)$. Applying Pólya's Theorem, we may even now write the form of the result:

$$(13) \quad \bar{g}_p(x) = Z(Q_p, 1+x) = \bar{Z}_2(S_p, 1+x).$$

The formulation is completed by securing an explicit statement of $Z(Q_p)$. Let the variables appearing in this cycle index be denoted by h_k , $k=1, \dots, p(p-1)$. We briefly illustrate the calculation of $Z(Q_p)$ for $p=3$. Here we have

$$Z(S_3) = (f_1^3 + 3f_1f_2 + 2f_3)/6.$$

Proceeding analogously as for ordinary graphs, but being careful to note that the figures are ordered pairs of points this time, we see that $f_1^3 \rightarrow h_1^6$, $f_1f_2 \rightarrow h_2^3$, and $f_3 \rightarrow h_3^2$. Hence $Z(Q_3) = \bar{Z}_2(S_3) = (h_1^6 + 3h_2^3 + 2h_3^2)/6$. Replacing h_i by $1+x^i$, we obtain $\bar{g}_3(x) = 1+x+4x^2+4x^3+4x^4+x^5+x^6$, which may be easily verified pictorially, as in Fig. 2.

It is again convenient to divide the contribution to $\bar{Z}_2(S_p)$ from $f_1^{j_1}f_2^{j_2} \cdots f_p^{j_p}$ into two parts. The ordered pairs of points lying on S_p -cycles of the same length give $\prod_{n=1}^p h_n^{2nC_{j_n,2}} \prod_{n=2}^p h_n^{(n-1)j_n}$, while the remaining ordered pairs yield $\prod_{1 \leq q < r \leq p} h_{m(q,r)}^{2jq_r d(q,r)}$.

Collecting these results, we have

$$(14) \quad \bar{Z}_2(S_p) = \frac{1}{p!} \sum_{(i)} \frac{p!}{\prod_{i=1}^p i! j_i!} \prod_{n=1}^p h_n^{2nC_{j_n,2}} \prod_{n=2}^p h_n^{(n-1)j_n} \prod_{1 \leq q < r \leq p} h_{m(q,r)}^{2jq_r d(q,r)}.$$

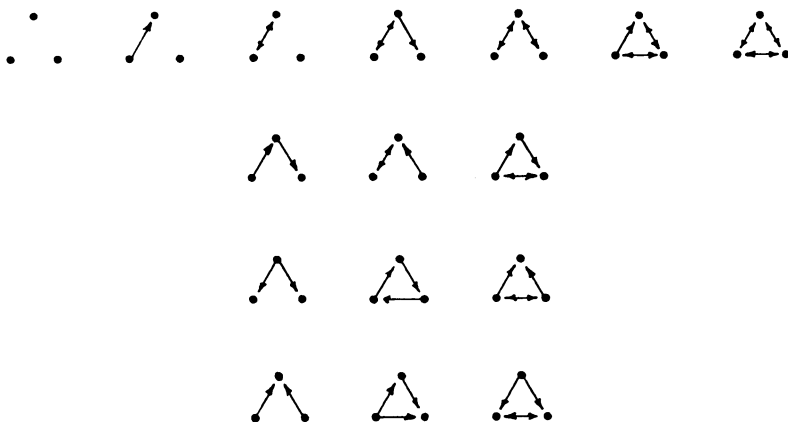


FIG. 2

Substituting $1+x^i$ for h_i in (14) yields $\bar{g}_p(x)$, as indicated in (13). The counting polynomials $\bar{g}_p(x)$ for $p=1$ to 5 are given similarly in the form of the double series $\bar{g}(x, y)$:

$$\begin{aligned}
 \bar{g}(x, y) &= \sum_{p=1}^{\infty} \bar{g}_p(x) y^p \\
 &= y + y^2(1 + x + x^2) \\
 &\quad + y^3(1 + x + 4x^2 + 4x^3 + 4x^4 + x^5 + x^6) \\
 &\quad + y^4(1 + x + 5x^2 + 13x^3 + 27x^4 + 38x^5 + 48x^6 \\
 (15) \quad &\quad + 38x^7 + 27x^8 + 13x^9 + 5x^{10} + x^{11} + x^{12}) \\
 &\quad + y^5(1 + x + 5x^2 + 16x^3 + 61x^4 + 154x^5 + 379x^6 + 707x^7 \\
 &\quad + 1155x^8 + 1490x^9 + 1670x^{10} + 1490x^{11} + 1155x^{12} \\
 &\quad + 707x^{13} + 379x^{14} + 154x^{15} + 61x^{16} + 16x^{17} + 5x^{18} + x^{19} + x^{20}) \\
 &\quad + \dots
 \end{aligned}$$

We conclude this section with some remarks on the correspondence between digraphs and binary relations. Davis [4] succeeded in counting several different kinds of binary relations on p objects. These include, besides all relations, the irreflexive (or reflexive), symmetric, irreflexive symmetric, asymmetric (or antisymmetric), and functional relations. The entries in the table of [4, p. 494] are related to graphs in the following way.

In Davis' notation,

(i) $irs(p)$ = the number of nonisomorphic irreflexive symmetric relations on p objects = the number of ordinary graphs of p points = $\sum_{k=0}^{p(p-1)/2} g_{pk}$.

(ii) $ref(p)$ = the number of nonisomorphic reflexive (or irreflexive) relations on p objects = the number of digraphs of p points = $\sum_{k=0}^{p(p-1)} \bar{g}_{pk}$.

(iii) $sym(p)$ = the number of nonisomorphic symmetric relations = the number of configurations which are modifications of ordinary graphs such that a loop, i.e., an arc joining a point with itself, is permitted at each point. Let the number of these modified graphs having p points and k lines be denoted by λ_{pk} , so that $sym(p) = \sum_{k=0}^{p(p-1)/2} \lambda_{pk}$, and let the counting series for those of p points be $\lambda_p(x)$. Let $Z_2(g_1^{j_1} g_2^{j_2} \cdots g_p^{j_p})$ denote the term of $Z_2(S_p)$ corresponding to the term $f_1^{j_1} f_2^{j_2} \cdots f_p^{j_p}$ of $Z(S_p)$. Then let

$$\Lambda(S_p) = \frac{1}{p!} \sum_{(i)} \frac{p!}{\prod_{i=1}^p i^{i_j} j_i!} g_1^{j_1} g_2^{j_2} \cdots g_p^{j_p} Z_2(g_1^{j_1} g_2^{j_2} \cdots g_p^{j_p}).$$

Then clearly,

$$\lambda_p(x) = \Lambda(S_p, 1 + x).$$

(iv) $st(p)$ = the total number of nonisomorphic relations on p objects = the number of modified digraphs such that loops are permitted.

These modified digraphs with p points and k lines may be counted in an entirely analogous manner to the modified graphs of the preceding paragraph. Namely, the counting polynomial in which the coefficient of x^k is the total number of nonisomorphic binary relations on p objects with k ordered couples is obtained by substituting $\phi(x) = 1 + x$ into the expression formed by multiplying the right-hand member of (14) under the summation sign by $h_1^{j_1} h_2^{j_2} \cdots h_p^{j_p}$.

5. The number of rooted and multiply rooted graphs. A *rooted graph* [11, p. 76] is a graph in which one particular point has been singled out. This distinguished point is called the *root*. Following the convention of [8] and [6], we use capital letters for enumerating rooted graphs. Let G_{pk} be the number of rooted graphs with p points and k lines and

$$(16) \quad G_p(x) = \sum_{k=0}^{p(p-1)/2} G_{pk} x^k.$$

A rooted graph is also a configuration in which the figures are pairs of points. As for graphs, the figure counting series is again $1 + x$. However, the configuration group is different.

The configuration group of rooted graphs of p points is the pair group of $S_1 \times S_{p-1}$, the direct product of S_1 and S_{p-1} , since the root point is in a class by itself, and the remaining $p-1$ points may be interchanged with each other without restriction. Hence an application of Pólya's Theorem yields the counting polynomial for rooted graphs of p points in the form:

$$(17) \quad G_p(x) = Z_2(S_1 \times S_{p-1}, 1 + x).$$

Proceeding exactly as for ordinary graphs, it is a simple matter to compute $Z_2(S_1 \times S_{p-1})$ explicitly. However, it is more convenient to describe the expression for $Z_2(S_1 \times S_p)$. The latter is obtained from the right-hand member of (10) when one replaces j_1 by $j_1 + 1$. The counting polynomials $G_p(x)$ for $p = 1$ to 6 are given by:

$$(18) \quad \begin{aligned} G(x, y) &= \sum_{p=1}^{\infty} G_p(x) y^p \\ &= y + y^2(1+x) + y^3(1+2x+2x^2+x^3) \\ &\quad + y^4(1+2x+4x^2+6x^3+4x^4+2x^5+x^6) \\ &\quad + y^5(1+2x+5x^2+11x^3+17x^4+18x^5+17x^6+11x^7+5x^8+2x^9+x^{10}) \\ &\quad + y^6(1+2x+5x^2+13x^3+29x^4+52x^5+76x^6+94x^7+94x^8 \\ &\quad + 76x^9+52x^{10}+29x^{11}+13x^{12}+5x^{13}+2x^{14}+x^{15}) \\ &\quad + \cdots \end{aligned}$$

In passing, we note that G_{pk} is the number of dissimilar points in all of the g_{pk} ordinary graphs of p points and k lines.

Rooted digraphs are counted similarly, and it is clear that the counting polynomial $\bar{G}_p(x)$ is expressible in the form

$$(19) \quad \bar{G}_p(x) = \bar{Z}_2(S_1 \times S_{p-1}, 1 + x).$$

A *multiply rooted graph* of p points (as defined in [8], for example) is one in which the set Π of points has been partitioned into disjoint distinguished classes Π_i , $i=1$ to q . Let p_i denote the number of points in the i th class, so that $p_1 + p_2 + \cdots + p_q = p$. A rooted graph of p points may be regarded as a multiply rooted graph in which $q=2$, $p_1=1$, and $p_2=p-1$. Similarly an ordinary graph can be considered as a multiply rooted graph in which $q=1$.

Two multiply rooted graphs are isomorphic if there exists a one-to-one adjacency preserving transformation between their point sets which induces a one-to-one correspondence between their respective distinguished classes of points.

Let $g_{p_1, p_2, \dots, p_q}(x)$ denote the counting polynomial in which the coefficient of x^k is the number of multiply rooted graphs of p points and k lines, whose distinguished classes satisfy the partition

$$p_1 + p_2 + \cdots + p_q = p.$$

Then

$$(20) \quad g_{p_1, p_2, \dots, p_q}(x) = Z_2(S_{p_1} \times S_{p_2} \times \cdots \times S_{p_q}, 1 + x).$$

Obviously multiply rooted digraphs are counted analogously.

6. The number of graphs of strength s and of type t . Motivated by psychological considerations, we have introduced, in [6], two notions of generalized graphs. Graphs of a certain strength are used to depict person-to-person relations of different intensity, while graphs of a given type are intended to represent several different interpersonal relationships within the same psychological group. It appears that the structural settings studied in group dynamics pertain primarily to digraphs and their generalizations. For the use of ordinary graphs is equivalent to a restriction to symmetric social relations.

Specifically, a *graph of strength s* consists of p points and not more than s lines joining any pair of points. No distinction is made between any of the lines of such a generalized graph.

On the other hand, a *graph of type t* has p points and not more than t lines joining any pair of points, with the restrictions that there are t different kinds or colors of lines available and no two points are joined by more than one line of a given color.

Of course all sorts of mixtures of ordinary and directed multiply rooted graphs of given type and strength can be precisely defined, and counted.

Let $g_{pk}^{(s)}$ be the number of graphs of strength s having p points and k lines, and let

$$(21) \quad g_p^{(s)}(x) = \sum_{k=0}^{sp(p-1)/2} g_{pk}^{(s)} x^k.$$

Since the figure counting series is clearly

$$(22) \quad \phi(x) = 1 + x + x^2 + \cdots + x^s$$

in this case, an application of Pólya's Theorem shows at once that

$$(23) \quad g_p^{(s)}(x) = Z_2\left(S_p, \sum_{i=0}^s x^i\right).$$

Combining (23) and (10), one can readily compute the polynomials $g_p^{(s)}(x)$.

With the obvious definition of the left-hand member, we see that

$$(24) \quad G_p^{(\infty)}(x) = Z_2\left(S_p, \frac{1}{1-x}\right).$$

For graphs of type t , the figure counting series is

$$(25) \quad \phi(x) = (1+x)^t,$$

since there are $C_{t,r}$ figures of content r .

Hence the counting polynomial for graphs of type t having p points is given by

$$(26) \quad Z_2(S_p, (1+x)^t).$$

Obviously (23) with $s=1$ and (26) with $t=1$ both reduce to (8), for graphs of strength 1 and graphs of type 1 are both equivalent to ordinary graphs.

7. On the number of connected graphs. We begin with Riddell's combinatorial derivation (which is not readily available in the literature) of the number of connected graphs in terms of the total number of graphs. This will be followed by a more general formulation for the number of connected graphs of any given topological type in terms of the total number of such graphs, and then by the number of such rooted connected graphs. As corollaries, we shall obtain the number of weakly connected digraphs, forests, and rooted forests.

Let c_{pk} and d_{pk} denote the number of connected and disconnected graphs of p points and k lines, so that $c_{pk} + d_{pk} = g_{pk}$, and let $c(x, y)$ and $d(x, y)$ be their respective counting series. In any graph of p points and k lines, let m_{ij} be the number of components with i points and j lines. These m_{ij} components can be distributed over all the c_{ij} different connected graphs in $C_{c_{ij}+m_{ij}-1, m_{ij}}$ ways. Therefore

$$(27) \quad g_{pk} = \sum_{(m_{ij})} \prod_{i=1}^p \prod_{j=0}^{i(i-1)/2} C_{c_{ij}+m_{ij}-1, m_{ij}}$$

where the sum is taken over all values of m_{ij} such that $\sum_{i,j} im_{ij} = p$, $\sum_{i,j} jm_{ij} = k$, and not all the m_{ij} are zero. From (27) follows

$$(28) \quad 1 + g(x, y) = \prod_{i,j} \sum_{m_{ij}=0}^{\infty} C_{c_{ij}+m_{ij}-1, m_{ij}} (x^i y^j)^{m_{ij}}.$$

The 1 has to be added to the left-hand member of (28) since the summation on the right-hand member begins with $m_{ij}=0$. Now since

$$\sum_{k=0}^{\infty} C_{c+k-1, k} z^k = \frac{1}{(1-z)^c},$$

one obtains:

$$1 + g(x, y) = \prod_{i,j} \frac{1}{(1 - x^i y^j)^{c_{ij}}},$$

whence

$$(29) \quad \log (1 + g(x, y)) = \sum_{n=1}^{\infty} \frac{1}{n} c(x^n, y^n),$$

or

$$(30) \quad \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} g^n(x, y) = \sum_{n=1}^{\infty} \frac{1}{n} c(x^n, y^n),$$

or

$$(30') \quad \sum_{n=1}^{\infty} \frac{1}{n} d(x^n, y^n) = \sum_{n=1}^{\infty} \frac{1}{n} [g(x^n, y^n) + (-1)^n g^n(x, y)],$$

which is equivalent with (30) since $g(x, y) = c(x, y) + d(x, y)$.

Now let us consider graphs having a given topological property, e.g., linear graphs, trees, Husimi trees, digraphs, graphs with no points of degree 2, graphs with no cycle of length greater than n , Euler graphs, etc. Let the counting series for all of these general graphs, the disconnected ones, and the connected ones be denoted by $\gamma(x, y)$, $\delta(x, y)$, $\kappa(x, y)$ respectively, with $\Gamma(x, y)$, $\Delta(x, y)$, $K(x, y)$ standing for the corresponding series for rooted graphs. Let $\gamma_n(x, y)$ be the counting series for the number of graphs having exactly n connected components. Then by Pólya's Theorem (regarding a graph as a configuration whose figures are connected graphs),

$$(31) \quad \gamma_n(x, y) = Z(S_n, \kappa(x, y)),$$

so that

$$\gamma(x, y) = \sum_{n=1}^{\infty} \gamma_n(x, y) = \sum_{n=1}^{\infty} Z(S_n, \kappa(x, y)).$$

But there is a well known identity for any function $f(x)$:

$$(32) \quad 1 + \sum_{n=1}^{\infty} Z(S_n, f(x)) = \exp \left\{ \sum_{n=1}^{\infty} \frac{1}{n} f(x^n) \right\},$$

which occurs in the counting of rooted trees and can easily be verified using (4), (9) and the familiar exponential power series. Therefore

$$(33) \quad 1 + \gamma(x, y) = \exp \left\{ \sum_{n=1}^{\infty} \frac{1}{n} \kappa(x^n, y^n) \right\},$$

from which the number of connected graphs is expressible in terms of the total number of graphs of general type⁽³⁾.

The derivation of $K(x, y)$ proceeds similarly. For

$$\begin{aligned} \Gamma(x, y) &= K(x, y) + K(x, y)\gamma_1(x, y) + K(x, y)\gamma_2(x, y) + \cdots \\ &= K(x, y) \left[1 + \sum_{n=1}^{\infty} Z(S_n, \kappa(x, y)) \right] \\ &= K(x, y) \exp \left\{ \sum_{n=1}^{\infty} \frac{1}{n} \kappa(x^n, y^n) \right\}, \end{aligned}$$

and using (33), we get

$$(34) \quad \Gamma(x, y) = K(x, y) [1 + \gamma(x, y)].$$

When (33) is considered as applying to all linear graphs, the equivalent of (29) is obtained. However the expression for linear graphs corresponding to (34) gives the following equation for the number of rooted connected graphs in terms of the known graphs (8) and (10):

$$(35) \quad G(x, y) = C(x, y) [1 + g(x, y)],$$

which can also be written in the form $D(x, y) = C(x, y)g(x, y)$, where $D(x, y)$ is the counting series for the rooted disconnected graphs.

We now turn to "connected" digraphs. A graph is *connected* if there exists a path between any two points. Equivalently, a graph is connected if it is impossible to divide its point set Π into two disjoint classes such that no line has one end point in each class. These two characterizations of connectedness for ordinary graphs lead to two different concepts for digraphs.

⁽³⁾ Of course Riddell's equation (29) is equivalent to (33) when the variables appearing in (29) are regarded as referring to graphs of general type.

A digraph is *strongly connected* if there exists a directed path from each point to every other point. The *symmetrized graph* of a digraph is the ordinary graph obtained by replacing each directed line by an undirected line with the proviso that each pair of directed lines joining the same two points of the digraph becomes exactly one undirected line. A digraph is *weakly connected* if its symmetrized graph is connected. Obviously every strongly connected digraph is weakly connected.

Corresponding to the double series $\bar{g}(x, y)$ for all digraphs, let $\bar{c}(x, y)$ be the counting series for the weakly connected digraphs, and let $\bar{d}(x, y)$ be the series for disconnected digraphs, i.e., those which are not even weakly connected.

Applying (33) to digraphs, we get:

$$(36) \quad \sum_{n=1}^{\infty} \frac{1}{n} \bar{c}(x^n, y^n) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \bar{g}^n(x, y),$$

which counts the weakly connected digraphs (first obtained by Pólya) in terms of the known number of all digraphs (13) and (14). Similarly one sees from (34) that the series $\bar{C}(x, y)$ for the rooted weakly connected digraphs is given by

$$(37) \quad \bar{G}(x, y) = \bar{C}(x, y) [1 + \bar{g}(x, y)],$$

where the rooted digraphs are known from (19).

The enumeration of strongly connected digraphs is an open problem.

From (30') and (36) we obtain the following data for the number of disconnected graphs and digraphs respectively, which can of course be verified pictorially:

$$(38) \quad \begin{aligned} d(x, y) = & y^3(1+x) + y^4(1+x+2x^2+x^3) \\ & + y^5(1+x+2x^2+4x^3+3x^4+x^5+x^6) \\ & + y^6(1+x+2x^2+5x^3+9x^4+9x^5+8x^6+5x^7+2x^8+x^9+x^{10}) \\ & + y^7(1+x+2x^2+5x^3+10x^4+21x^5+30x^6+32x^7+30x^8+24x^9 \\ & + 16x^{10}+10x^{11}+5x^{12}+2x^{13}+x^{14}+x^{15}) \\ & + \dots, \\ \bar{d}(x, y) = & y^2+y^3(1+x+x^2) \\ & + y^4(1+x+5x^2+5x^3+5x^4+x^5+x^6) \\ (39) \quad & + y^5(1+x+5x^2+16x^3+34x^4+46x^5+53x^6+40x^7+28x^8 \\ & + 13x^9+5x^{10}+x^{11}+x^{12}) \\ & + \dots. \end{aligned}$$

We conclude with some formulas for the number of forests in terms of the

known number of trees. By definition, a *tree* is a connected graph without any cycles, and a *forest* is a graph with no cycles, i.e., one whose components are trees. Since a tree is characterizable as a connected graph in which the number of points is one greater than the number of lines, we may omit either one of the parameters p and k in specifying the number of trees with p points and k lines. We shall denote (as in [9]) the number of trees and rooted trees having k lines by t_k and T_k respectively. Let

$$(40) \quad t(x) = \sum_{k=0}^{\infty} t_k x^k,$$

$$(41) \quad T(x) = \sum_{k=0}^{\infty} T_k x^k$$

be the counting series for trees, rooted trees.

It was demonstrated in [15] that

$$(42) \quad T(x) = \exp \left\{ \sum_{r=1}^{\infty} \frac{x^r}{r} \cdot T(x^r) \right\},$$

and in [14] that

$$(43) \quad t(x) = T(x) - (x/2) [T^2(x) - T(x^2)].$$

Hence the numbers t_k and T_k may be regarded as known.

Let f_{pk} denote the number of forests having p points and k lines and

$$(44) \quad f(x, y) = \sum_{p=1}^{\infty} \sum_{k=0}^{p-1} f_{pk} x^k y^p.$$

In order to express the counting series for trees in two variables, let t_{pk} denote the number of trees with p points and k lines, so that $t_{k+1,k} = t_k$ and $t_{pk} = 0$ when $p \neq k+1$. Then

$$t(x, y) = \sum_{p=1}^{\infty} \sum_{k=0}^{p-1} t_{pk} x^k y^p = \sum_{k=0}^{\infty} t_k x^k y^{k+1} = t(xy^{(k+1)/k}).$$

By (33), we have the number of forests in terms of the number of trees in the form:

$$\begin{aligned} 1 + f(x, y) &= \exp \left\{ \sum_{n=1}^{\infty} \frac{1}{n} t(x^n, y^n) \right\} \\ (45) \quad &= \exp \left\{ \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} \frac{t_k}{n} (x^k y^{k+1})^n \right\} \\ &= \prod_{k=0}^{\infty} (1 - x^k y^{k+1})^{-t_k}. \end{aligned}$$

Similarly the counting series for rooted forests $F(x, y)$ is given by

$$(46) \quad \begin{aligned} F(x, y) &= T(x, y)[1 + f(x, y)] \\ &= T(xy^{(k+1)/k})[1 + f(x, y)]. \end{aligned}$$

Finally, if $\tau(x, y)$ is the counting series for those multiply rooted trees such that *each component is a rooted tree*, then by (33) and (45), we have

$$(47) \quad 1 + \tau(x, y) = \prod_{k=0}^{\infty} (1 - x^k y^{k+1})^{-T_k}.$$

However, Cayley [3] has already found a recursion equation for $T(x)$, the counting series for rooted trees defined in (41), in the form

$$(48) \quad T(x) = \prod_{k=0}^{\infty} (1 - x^{k+1})^{-T_k}.$$

Using (47), equation (48) can be rewritten:

$$(48') \quad 1 + \tau(x, y) = T(yx^{k/(k+1)}).$$

8. Problems. We conclude with a listing of several interesting counting problems which remain to be solved:

I. The number of strongly connected digraphs is less than the known number of weakly connected ones. However, aside from this trivial inequality, the appropriate method for enumerating the strongly connected digraphs has not yet become apparent. There is an intermediate kind of connectedness for digraphs. A *unilaterally connected digraph* is one in which, for each pair A, B of its points, there is either a directed path from A to B or one from B to A . The counting of unilaterally connected digraphs is also an unsolved problem.

II. The enumeration of planar graphs with a given number of points and lines has been an open problem for some time. It is not inconceivable that a solution of this problem may be a partial contribution to the solution of the four color problem.

III. A *star* has been defined in [21] as a connected graph with no articulation points. The number of stars is an unsolved problem whose solution is of interest for application to the theory of condensation, and is one of the principal questions asked by Uhlenbeck [21].

IV. Two graphs are *homeomorphic* (König [11]) if there exists a third graph which can be obtained (up to isomorphism) from each of them by the insertion of additional vertices in their lines. A *homeomorphically irreducible graph* is an ordinary graph which is not homeomorphic to any linear graph of fewer points. The counting of homeomorphically irreducible graphs is an open problem, suggested in [21]. A variation of this problem results if one considers graphs of arbitrary strength. From this point of view, a homeomorphically irreducible graph of unrestricted strength is one with no points

of degree two.

V. As mentioned by Davis [14], the problem of finding the number of transitive digraphs is open.

VI. Finally, we mention the number of simplicial complexes. A *pure complex* of dimension n is one in which each simplex of dimension $m < n$ is contained in an n -simplex. The counting of pure complexes of dimension n is a straightforward generalization of equation (8), when it is rewritten in the form

$$(8') \quad g_p(x) = Z_2(S_p, 1 + x).$$

If $h_{n,p}(x)$ is the counting polynomial for pure n -complexes with p points, then we may write

$$h_{n,p}(x) = Z_{n+1}(S_p, 1 + x),$$

where $Z_m(S_p)$ denotes the cycle index of the permutation group of degree $C_{p,m}$, which is isomorphic to S_p as an abstract group and is obtained from S_p by taking sets of m objects as the new objects to be permuted.

The solution to the problem of enumerating arbitrary simplicial complexes is in progress, and we hope to present it in a future publication.

Added in proof. The following relevant paper has recently appeared: D. Slepian, *On the number of symmetry types of Boolean functions of n variables*, Canadian Journal of Mathematics vol. 5 (1953) pp. 185–193. Slepian independently develops the same enumeration procedure as Davis [4] and applies it to obtain a generalization of the results in G. Pólya, *Sur les types des propositions composées*, J. Symbolic Logic vol. 5 (1940) pp. 98–103. Slepian then applied his method and independently found Pólya's equations (8) and (10) of this paper and also the formula (23) for the number of graphs of strength s ! Since he was guided in his work by applications to switching theory, he called his result "an algorithm for enumerating the types of linear graphs with n nodes and m branches no more than k of which are in parallel." (We introduced the concept of graphs of strength s from the viewpoint of applications to social psychology.) Slepian's work appears in a Technical Memorandum of Bell Telephone Laboratories entitled *Number of types of linear graphs* under date of June 22, 1953.

BIBLIOGRAPHY

1. C. M. Blair and H. R. Henze, *The number of structural isomers of the more important types of aliphatic compounds*, Journal of the American Chemical Society vol. 56 (1934) p. 157.
2. R. Carnap, *Logical foundations of probability*, Chicago, 1950.
3. A. Cayley, *Collected mathematical papers*, Cambridge, 1889–1898, vol. 3, pp. 242–246 and vol. 11, pp. 365–367.
4. R. L. Davis, *The number of structures of finite relations*, Proc. Amer. Math. Soc. vol. 4 (1953) pp. 486–495.
5. R. M. Foster, *The number of series-parallel networks*, Proceedings of the International Congress of Mathematicians, 1950, vol. 1, 1952, p. 646.

6. F. Harary and R. Z. Norman, *Graph theory as a mathematical model in social science*, Research Center for Group Dynamics, Monograph No. 2, Ann Arbor, 1953.
7. ———, *The dissimilarity characteristic of Husimi trees*, Ann. of Math. vol. 58 (1953) pp. 134–141.
8. ———, *The dissimilarity characteristic of linear graphs*, Proc. Amer. Math. Soc. vol. 5 (1954) pp. 131–135.
9. F. Harary and G. E. Uhlenbeck, *On the number of Husimi trees*, I, Proc. Nat. Acad. Sci. U.S.A. vol. 39 (1953) pp. 315–322.
10. C. A. Hurst, *The enumeration of graphs in the Feynman-Dyson technique*, Proc. Royal Soc. London Ser. A vol. 214 (1952) pp. 44–61.
11. D. König, *Theorie der endlichen und unendlichen graphen*, Leipzig, 1936, reprinted, New York, 1950.
12. A. C. Lunn and J. K. Senior, *Isomerism and configuration*, J. Phys. Chem. vol. 33 (1929) pp. 1027–1079.
13. R. Z. Norman, *On the number of linear graphs with given blocks*, Doctoral dissertation, University of Michigan, 1954.
14. R. Otter, *The number of trees*, Ann. of Math. vol. 49 (1948) pp. 583–599.
15. G. Pólya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen, und chemische Verbindungen*, Acta Math. vol. 68 (1937) pp. 145–254.
16. R. J. Riddell, *Contributions to the theory of condensation*, Doctoral dissertation, Department of Physics, University of Michigan, 1951.
17. ———, *On the number of Feynman-Dyson graphs*, Physical Review vol. 91 (1953) pp. 1243–1248.
18. R. J. Riddell and G. E. Uhlenbeck, *On the theory of the virial development of the equation of state of monoatomic gases*, J. Chem. Phys. vol. 21 (1953) pp. 2056–2064.
19. J. Riordan and C. E. Shannon, *The number of two terminal series-parallel networks*, J. of Math. and Phys. vol. 21 (1942) pp. 83–93.
20. H. Seifert and W. Threlfall, *Lehrbuch der Topologie*, Leipzig, 1934; reprinted New York, 1947.
21. G. E. Uhlenbeck, *Some basic problems of statistical mechanics*, Gibbs Lecture, American Mathematical Society, December 1950.

UNIVERSITY OF MICHIGAN,
ANN ARBOR, MICH.