# ENUMERATION OF LINEAR GRAPHS FOR MAPPINGS OF FINITE SETS

By John Riordan

*Bell Telephone Laboratories, Inc.*

**1. Introduction.** A finite set of $n$ elements may be mapped onto itself in $n^n$ ways, since each element may be mapped independently on any element. Each mapping is a permutation with unlimited repetition. A linear graph is found for a mapping by drawing a sling if $i$ is mapped on $i$, a line from $i$ to $j$ if $i$ is mapped on $j$. Hence each linear graph consists of one or more components (connected parts) and each component has a single cycle (closed path) of length $k = 1, 2, \cdots$, if a sling is regarded as of length 1, and a pair of points connected by two lines a cycle of length 2. Also, all points of the graph are labeled and directedness of the lines has significance for the enumeration only in that, for $k > 2$, the lines in cycle may be directed in two ways. Note that slings and multiple lines between pairs of points usually are banned in graph enumerations.

The study of such mappings may be given a probability setting by assigning a probability for each mapping. In the simplest case, the probability is the same for all, and the mappings are said to be random. Random mappings have been considered in 1953 by N. Metropolis and S. Ulam [7] who raised the question of the expected number of components, answered in 1954 by Martin D. Kruskal [6]. H. Rubin and R. Sitgreaves in 1954 considered other random variables associated with the mappings, including the number of lines in cycles. Both Jay E. Folkert [2] and Bernard Harris [4] have considered the enumeration by number of components, both with and without slings. Finally Leo Katz [5] has enumerated the connected graphs with slings, while Alfréd Rényi [8] has given the corresponding result for the classical case where slings and multiple lines are banned.

In the present paper new and simpler results are obtained for the enumerations both by number of components and by number of lines in cycle of the unrestricted and various restricted graphs.

**2. The number of components in unrestricted graphs.** The graphs in question are those corresponding to the complete set $(n^n)$ of mappings described above. Let $T_{nk}$ be the number of such graphs with $n$ distinct (labeled) points and $k$ components. Let $C_k$ be the number of such connected graphs with $k$ labeled points. Then if the enumerator by number of parts is

$$T_n(x) = \sum_{k=1}^{n} T_{nk}x^k, \qquad T_0(x) = 1,$$

and if

$$C(y) = \sum_{n=1}^{\infty} C_n \frac{y^n}{n!},$$

Received May 19, 1961.

the basic equation for the enumeration is

$$T(x, y) = \sum_{n=0}^{\infty} T_n(x) \frac{y^n}{n!} = \exp [xC(y)]. \tag{1}$$

Equation (1) is a familiar form and is derived easily by an argument for a similar situation given by E. N. Gilbert [3], which in present terms is as follows. In the graphs with $n + 1$ labeled points and $k$ components, the point labeled $n + 1$ belongs in a component with $j$ points, $j = 0, 1, \cdots$ while the remaining $n - j$ labeled points belong to a graph with $k - 1$ components. Hence

$$T_{n+1,k} = \sum_{j=0}^{n} \binom{n}{j} C_{j+1} T_{n-j,k-1} . \tag{2}$$

Forming the enumerator $T_{n+1}(x)$ leads to

$$T_{n+1}(x) = x \sum_{j=0}^{n} \binom{n}{j} C_{j+1} T_{n-j}(x). \tag{3}$$

Then

$$\begin{aligned} \frac{\partial T(x, y)}{\partial y} &= \sum_{n=0}^{\infty} T_{n+1}(x) \frac{y^n}{n!} \\ &= x \frac{\partial C(y)}{\partial y} T(x, y) \end{aligned}$$

may be integrated with respect to $y$ to give (1).

Since $T_n(1) = n^n$, and $T(1, y) = \exp C(y)$, equation (1) effectively determines $C(y)$ and hence $T(x, y)$. Indeed, writing $T_n = T_n(1)$, the equation

$$T(1, y) = \exp yT = \exp C(y), \qquad T^n \equiv T_n$$

is the equation relating ordinary moments and cumulants. Hence by equation (51) of Chapter 2 of [9],

$$C_n = Y_n(fT_1, \cdots, fT_n), \qquad f^k \equiv f_k = (-1)^{k-1}(k - 1)! \tag{4}$$

with $Y_n$ a multivariable Bell polynomial. Also, since (1) may be rewritten

$$T(x, y) = (\exp yT)^x, \qquad T^n \equiv T_n ,$$

it follows from [9; problem 24 of Chapter 2] that

$$T_n(x) = Y_n(gT_1, \cdots, gT_n), \quad g^k \equiv g_k = x(x - 1) \cdots (x - k + 1) \tag{5}$$

an equation given (effectively) both by Folkert and by Harris.

Simpler results appear when $C(y)$ is formulated independently. Note first that the number of connected graphs with a single sling and no other cycles is equal to the number of rooted trees with labeled points, since the sling effectively identifies the root. Thus if $C(y,k)$ is the enumerator for graphs with a single cycle of $k$, $C(y, 1) = R(y) = \sum_{n=1} n^{n-1}y^n/n!$, the enumerator for rooted trees. Next $C(y, 2) = R^2(y)/2$, since there is a rooted tree at each of the two points on the cycle, and these two points may be interchanged. Finally $C(y, k) = R^k(y)/k$ by a similar argument. The corresponding result when the directedness of the lines in cycle is ignored is $R^k(y)/2k$, $k = 3, 4, \cdots$, (cf., equation (16) of T. L. Austin et al. [1]); the difference flows from the different orders, $k$ and $2k$, respectively, of the associated cyclic and dihedral groups. Thus finally

$$\begin{aligned} C(y) &= C(y, 1) + C(y, 2) + \cdots + C(y, k) + \cdots \\ &= R(y) + R^2(y)/2 + \cdots + R^k(y)/k + \cdots \\ &= \log (1 - R(y))^{-1}. \end{aligned} \tag{6}$$

Now notice that if $u = R(y)$, $ue^{-u} = y$ [9; equation (45) of Chapter 6], and the Lagrange formula gives

$$f(u) = f(0) + \sum_{n=1}^{\infty} D^{n-1}[f'(u)e^{nu}]_{u=0} \frac{y^n}{n!}, \qquad D = \frac{d}{du} \tag{7}$$

(the prime denotes a derivative). Also, using the Leibniz formula for differentiation of a product,

$$D^{n-1}[f'(u)e^{nu}]_{u=0} = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{n-1-k} f_{k+1} \tag{8}$$

where

$$f(u) = \sum_{k=0}^{\infty} f_k \frac{u^k}{k!} = \exp uf, \qquad f^n \equiv f_n. \tag{9}$$

Then by (6), (7) and (8)

$$C_n = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{n-1-k} k! = (n-1)! \sum_{k=0}^{n-1} \frac{n^k}{k!} \tag{10}$$

the result found by Leo Katz [5].

Now return to equation (1). Using (6), it becomes

$$T(x, y) = (1 - R(y))^{-x}. \tag{11}$$

Since

$$(1 - u)^{-x} = \exp u\, c(x),$$

$$c^n(x) \equiv c_n(x) = x(x+1) \cdots (x+n-1),$$

where $c_n(x)$ is the enumerator of permutations (without repetition) by number of cycles, it follows from (7), (8) and (9) that

$$T_n(x) = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{n-1-k} c_{k+1}(x). \tag{12}$$

Then, if $c_n(x) = \sum c_{nj} x^j$ (the numbers $c_{nj}$ are the signless Stirling numbers of the first kind)

$$T_{n,j} = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{n-1-k} c_{k+1,j} .$$

Note that $T_{n1} = C_n$ follows from $c_{k+1,1} = k!$ and (10). Note also that $T_n(1) = n^n$ and (12) produce the identity

$$n^n = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{n-1-k} (k+1)!,$$

which is one of the forms associated with Abel's generalization of the binomial formula.

The generating function for the probability distribution of the number of components of random mappings is

$$P_n(x) = n^{-n} T_n(x) = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{-1-k} c_{k+1}(x). \tag{13}$$

The corresponding binomial moment generating function is

$$B_n(x) = P_n(1+x) = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{-1-k} c_{k+1}(1+x). \tag{14}$$

But $c_n(1+x) = x^{-1} c_{n+1}(x)$. Hence, the $j$th binomial moment, $B_{nj}$, is given by

$$B_{nj} = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{-1-k} c_{k+2,j+1} . \tag{15}$$

In particular the mean is

$$B_{n1} = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{-1-k} (k+1)! \left[1 + \frac{1}{2} + \cdots + \frac{1}{k+1}\right]$$
$$= \sum_{j=1}^{n} j^{-1} \sum_{k=j-1}^{n-1} \binom{n-1}{k} n^{-1-k} (k+1)! = \sum_{j=1}^{n} \frac{A_{nj}}{j} .$$

Then since $A_{nn} = n!n^{-n}$ and

$$A_{nj} = A_{n,j+1} + \binom{n-1}{j-1} n^{-j} j!,$$

it follows by mathematical induction that $A_{nj} = n!/(n-j)!\,n^j$ and

$$B_{n1} = \sum_{j=1}^{n} \frac{n!}{(n-j)!\,n^j j}$$

in agreement with Martin D. Kruskal [6].

**3. The number of lines in cycles in unrestricted graphs.** The enumeration by number of lines in cycle requires only slight modification. First the enumerator of connected graphs by number of lines in cycle (variable $x$) and by number of labeled points (variable $y$) is

$$C(x, y) = \sum C_n(x)y^n/n! \tag{16}$$
$$= xR(y) + \cdots + x^kR^k(y)/k + \cdots$$
$$= \log (1 - xR(y))^{-1}.$$

Hence, following the derivation of (10),

$$C_n(x) = \sum_{k=0}^{n-1} \binom{n-1}{k} n^{n-1-k}k!\, x^{k+1}. \tag{17}$$

Let $q_{nk}$ be the number of graphs with $n$ labeled points and $k$ lines in cycle, and write $q_n(x)$ for its generating function. Then, by the argument used for (1), with $q_0(x) = 1$,

$$q(x, y) = \sum q_n(x)y^n/n! = \exp C(x, y) \tag{18}$$

and, using (16),

$$q(x, y) = (1 - xR(y))^{-1}. \tag{18a}$$

It follows at once that $q_0 = 1$ and

$$q_n(x) = \sum_0^{n-1} \binom{n-1}{k} n^{n-1-k}(k + 1)!\, x^{k+1}, \qquad n = 1, 2, \cdots, \tag{19}$$

in agreement with Bernard Harris [4]. The binomial moment generating function of the corresponding probability distribution is

$$B_n(x) = \sum_1^n \binom{n-1}{k-1} n^{-k}k!(1 + x)^k \tag{20}$$

and

$$B_{nj} = \sum_1^n \binom{n-1}{k-1} \binom{k}{j} n^{-k}k!. \tag{20a}$$

It may also be noted that comparison of (17) and (19) or the relation

$$q(x, y) = 1 + x(d/dx)C(x, y)$$

leads to (the prime denotes a derivative)

$$q_n(x) = xC'_n(x), \qquad n = 1, 2, \cdots.$$

Further

$$(\partial/\partial x)q(x, y) = \exp yq'(x) = R(y)(1 - xR(y))^{-2}, \qquad (q'(x))^n = q'_n(x),$$

and since

$$(1 - R(y))yR'(y) = R(y), \exp yq'(1) = yR'(y)(1 - R(y))^{-1} = yC'(y);$$

or

$$q'_n(1) = n^n B_{n1} = nC_n ,$$

a result appearing in Bernard Harris [4].

**4. Restricted graphs.** The enumerations above are readily adapted to restrictions expressed in terms of cycle lengths. The simplest of these is that slings be barred, but it is just as easy to bar cycles of length $r$.

Considering the enumeration by number of components, write $T_n(x, r)$ for the enumerator with no cycles of $r$, $C(y, r)$ for the corresponding enumerator of connected graphs. Then by (6)

$$C(y, r) = \log (1 - R(y))^{-1} - R^r(y)/r,$$

which corresponds to

$$C_n(r) = C_n - \binom{n-1}{r-1} n^{n-r} r!. \tag{21}$$

Next

$$\begin{aligned} T(x, y, r) &= \sum T_n(x, r) y^n/n! = \exp [xC(y, r)], \\ &= (1 - R(y))^{-x} \exp [-xR^r(y)/r], \end{aligned}$$

which corresponds to $T_0(r) = 1$ and

$$T_n(x, r) = \sum_0^{n-1} \binom{n-1}{k} n^{n-1-k} d_{k+1}(x, r), \tag{22}$$

where

$$e^{-xu^r/r}(1 - u)^{-x} = \exp u\, d(x, r), \qquad d^n(x, r) \equiv d_n(x, r).$$

The polynomial $d_n(x, r)$ of course is the enumerator by number of cycles of permutations without cycles of length $r$. Its recurrence relation has been given in problem 16 of Chapter 4 of [9], and in present notation, reads

$$d_{n+1}(x, r) = (n + x)d_n(x, r) - (n)_{r-1} x d_{n-r+1}(x, r) + (n)_r x d_{n-r}(x, r),$$

where $(n)_r = n(n - 1) \cdots (n - r + 1)$. The binomial moments for the corresponding probability distribution have the generating function

$$b_n(x, r) = d_n(1 + x, r)/d_n(1, r).$$

Hence the corresponding generating function for graphs is given by

$$\begin{aligned} T_n(1, r) B_n(x, r) &= \sum_0^{n-1} \binom{n-1}{k} n^{n-1-k} d_{k+1}(1 + x, r), \\ &= \sum_1^{n} \binom{n-1}{k-1} n^{n-k} d_k(1, r) b_k(x, r), \end{aligned} \tag{23}$$

or

$$T_n(1, r)B_{nj}(r) = \sum_1^n \binom{n-1}{k-1} n^{n-k} d_k(1, r)b_{kj}(r). \tag{23a}$$

It is worth noting that

$$d_n(1 + x, r) = (d(1, r) + d(x, r))^n, \qquad d^n(x, r) \equiv d_n(x, r),$$
$$= \sum_0^n \binom{n}{k} d_{n-k}(1, r)\, d_k(x, r).$$

For $r = 1$, $T_n(1, 1) = (n - 1)^n$ and $d_n(1, 1) = D_n$, the subfactorial of $n(D_n = \Delta^n 0!)$, which with (22) leads to the interesting identity

$$(n - 1)^n = \sum_0^{n-1} \binom{n-1}{k} n^{n-1-k} D_{k+1}, \qquad n = 1, 2, \cdots.$$

Another identity may be obtained as follows. First

$$T(x, y, 1) = e^{-xR(y)}T(x, y);$$

then

$$e^{-xR(y)} = 1 - x\sum_{n=1} (n - x)^{n-1}y^n/n!.$$

Hence

$$T_n(x, 1) = T_n(x) - x \sum_1^n \binom{n}{k} T_{n-k}(x)\ (k - x)^{k-1}.$$

For $x = 1$ this corresponds to the identity

$$n^n = (n - 1)^n + \sum_1^n \binom{n}{k} (n - k)^{n-k}(k - 1)^{k-1} \qquad (0^0 = 1).$$

For the enumeration by number of lines in cycles, first

$$C(x, y, r) = \log (1 - xR(y))^{-1} - x^r R^r(y)/r,$$

and

$$C_n(x, r) = C_n(x) - \binom{n-1}{r-1} n^{n-r} r!\, x^{r+1}. \tag{24}$$

Then

$$q(x, y, r) = \exp C(x, y, r),$$
$$= (1 - xR(y))^{-1} \exp [-x^r R^r(y)/r],$$

so that $q_0(x, r) = 1$ and

$$q_n(x, r) = \sum_0^{n-1} \binom{n-1}{k} n^{n-1-k} d_{k+1}(1, r)x^{k+1}. \tag{25}$$

It is clear that the procedure followed reduces enumerations for any restricted graphs with restrictions specified by cycle lengths to the enumeration of the corresponding restricted permutations.

## REFERENCES

[1] T. L. Austin, R. E. Fagen, W. F. Penney and John Riordan, "The number of components in random linear graphs," *Ann. Math. Stat.*, Vol. 30 (1959), pp. 747–754.

[2] Jay E. Folkert, "The distribution of the number of components of a random mapping function," Unpublished Ph.D. dissertation, Michigan State University, 1955.

[3] E. N. Gilbert, "Enumeration of labelled graphs," *Canad. J. Math.*, Vol. 8 (1956), pp. 405–411.

[4] Bernard Harris, "Probability distributions related to random mappings," *Ann. Math. Stat.*, Vol. 31 (1960), pp. 1045–1062.

[5] Leo Katz, "Probability of indecomposability of a random mapping function," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 512–517.

[6] Martin D. Kruskal, "The expected number of components under a random mapping function," *Amer. Math. Monthly*, Vol. 61 (1954), pp. 392–397.

[7] N. Metropolis and S. Ulam, "A property of randomness of an arithmetical function," *Amer. Math. Monthly*, Vol. 60 (1953) pp. 252–253.

[8] Alfréd Rényi, "On connected graphs, I," *Publications of the Math. Inst. of the Hung. Acad. of Sciences*, Vol. 4 (1959), pp. 385–388.

[9] John Riordan, *An Introduction to Combinatorial Analysis*, John Wiley and Sons, New York, 1958.

[10] H. Rubin and R. Sitgreaves, "Probability distributions related to random transformations on a finite set," Tech. Report No. 19A, Applied Mathematics and Statistics Laboratory, Stanford University, 1954.