SIAM J. DISC. MATH.
Vol. 3, No. 4, pp. 574–581, November 1990

013

ON THE NUMBER OF DISTINCT FORESTS*

LAJOS TAKÁCS†

Abstract. This paper contains a simple explicit formula, a recurrence formula and an asymptotic expansion for the number of distinct forests with n labeled vertices.

Key words. enumeration of forests, exact and asymptotic formulas

AMS(MOS) subject classifications. 05C05, 05C30

1. Introduction. A forest is a simple graph that has no cycles. In other words, a forest is a simple graph, all of whose components are trees. Denote by $F(n)$ the number of distinct forests having vertex set $\{1, 2, \cdots, n\}$. In 1959, Dénes [3] and Rényi [7] gave two explicit expressions for $F(n)$. Rényi also proved that

(1) $$\lim_{n\to\infty} F(n)/n^{n-2} = \sqrt{e} = 1.6487212707\cdots.$$

See also Riordan [8]. In 1980 Stanley [10] observed that $F(n)$ can also be interpreted as the number of different score vectors in a tournament in which each pair of n players $1, 2, \cdots, n$ plays a game. For each player a game may result in either a win, a tie, or a loss. Denote by v_i the total number of wins of player i. Then $(v_1, v_2, \cdots, v_n)$ is the score vector of the tournament. In 1981, Kleitman and Winston [6] proved that there is a one-to-one correspondence between the elements of the set of forests with n labeled vertices and the elements of the set of score vectors in a tournament of n players. The sequence $F(1), F(2), \cdots, F(10)$ is listed as Sequence 714 in the book of Sloane [9].

2. New results. In what follows we derive a simple explicit formula for $F(n)$. Namely we prove that

(2) $$F(n) = \frac{n!}{n+1}\sum_{j=0}^{[n/2]}(-1)^j\frac{(2j+1)(n+1)^{n-2j}}{2^j j!(n-2j)!}$$

if $n \geqq 1$, or equivalently,

(3) $$F(n) = H_n(n+1) - nH_{n-1}(n+1)$$

for $n \geqq 1$ where $H_n(x)$ is the nth Hermite polynomial defined by

(4) $$H_n(x) = n!\sum_{j=0}^{[n/2]}\frac{(-1)^j x^{n-2j}}{2^j j!(n-2j)!}$$

for $n \geqq 0$. We have $H_0(x) = 1$, $H_1(x) = x$ and

(5) $$H_n(x) = xH_{n-1}(x) - (n-1)H_{n-2}(x)$$

for $n \geqq 2$. It is convenient to use the recurrence formula (5) for calculating numerically $H_{n-1}(n+1)$ and $H_n(n+1)$ in (3). Table 1 contains $F(n)$ for $n \leqq 24$.

We can also express $F(n)$ in the following way:

(6) $$F(n) = \frac{n^n e^{1/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2-ix}\left(1+\frac{ix}{n}\right)^n ix\,dx$$

* Received by the editors August 3, 1989; accepted for publication October 4, 1989.
† Department of Mathematics and Statistics, Case Western Reserve University, Cleveland, Ohio 44106.

TABLE 1

n	$F(n)$	n	$F(n)$	n	$F(n)$	n	$F(n)$
1	1	7	36,961	13	$3.52563011\cdots 10^{12}$	19	$1.02392064\cdots 10^{22}$
2	2	8	561,948	14	$1.10284283\cdots 10^{14}$	20	$4.86909744\cdots 10^{23}$
3	7	9	10,026,505	15	$3.74835769\cdots 10^{15}$	21	$2.44766976\cdots 10^{25}$
4	38	10	205,608,536	16	$1.37557910\cdots 10^{17}$	22	$1.29692217\cdots 10^{27}$
5	291	11	4,767,440,679	17	$5.42117905\cdots 10^{18}$	23	$7.22423439\cdots 10^{28}$
6	2932	12	123,373,203,208	18	$2.28359487\cdots 10^{20}$	24	$4.22040860\cdots 10^{30}$

for $n \geqq 1$. This integral representation is convenient to obtain the asymptotic expansion of $F(n)/n^n$ as $n \to \infty$. We shall prove that

$$F(n) \sim n^n e^{1/2} \sum_{\nu=0}^{\infty} \frac{f_\nu}{n^\nu} \tag{7}$$

as $n \to \infty$ where the coefficients f_ν ($\nu = 0, 1, 2, \cdots$) are given in Table 2.

Explicitly,

$$f_\nu = \sum_{i=1}^{[\nu/2]} (-1)^{i-1} \frac{C(\nu, 2i-1)}{2^{\nu-i}(\nu-i)!} \tag{8}$$

for $\nu \geqq 2$ where $C(\nu, j)$, $(0 \leqq j \leqq \nu - 1)$, is the number of permutations of $(1, 2, \cdots, 2\nu - j)$ with $\nu - j$ cycles each of length >1. $C(\nu, j) = 0$ if $j \geqq \nu$. We have also

$$f_{\nu+1} = f_\nu + \sum_{i=0}^{[(\nu-1)/2]} (-1)^i \frac{C(\nu, 2i)}{2^{\nu-i-1}(\nu-i-1)!} \tag{9}$$

for $\nu \geqq 1$. Table 3 contains $C(\nu, j)$ for $0 \leqq j \leqq \nu - 1$ and $\nu \leqq 10$.

We can calculate $C(\nu, j)$ for $0 \leqq j \leqq \nu - 1$ by the recurrence formula

$$C(\nu+1, j) = [C(\nu, j) + C(\nu, j-1)](2\nu + 1 - j) \tag{10}$$

where $0 \leqq j \leqq \nu$, $\nu \geqq 1$, $C(1, 0) = 1$, $C(1, j) = 0$ for $j > 0$ and $C(\nu, -1) = 0$ for $\nu \geqq 1$. In particular, $C(\nu, 0) = 1.3\cdots(2\nu - 1) = (2\nu)!/2^\nu \nu!$ for $\nu \geqq 1$ and $C(\nu, \nu - 1) = \nu!$ for $\nu \geqq 1$.

The fact that the coefficients $C(\nu, j)$, $0 \leqq j \leqq \nu - 1$, satisfy (10) can be proved in the following way: By definition $C(\nu + 1, j)$ is the number of permutations of $(1, 2, \cdots, 2\nu - j + 2)$ with $\nu - j + 1$ cycles each of length >1. There are two possibilities: (a) Element $2\nu - j + 2$ belongs to a cycle which contains more than two elements. If we remove element $2\nu - j + 2$, the remaining permutation of $(1, 2, \cdots, 2\nu - j + 1)$ contains $\nu - j + 1$ cycles of length >1. There are $C(\nu, j - 1)$ such permutations and ele-

TABLE 2

ν	f_ν	ν	f_ν
0	0	6	−17,207/384
1	0	7	−3,607/768
2	1	8	1,408,301/3072
3	5/2	9	8,181,503/6144
4	11/8	10	−137,483,257/61440
5	−203/16	11	−24,971,924,401/983040

TABLE 3
$C(\nu, j)$

ν \ j	0	1	2	3	4	5	6	7	8	9
1	1									
2	3	2								
3	15	20	6							
4	105	210	130	24						
5	945	2520	2380	924	120					
6	10395	34650	44100	26432	7308	720				
7	135135	540540	866250	705320	303660	64224	5040			
8	2027025	9459450	18288270	18858840	11098780	3678840	623376	40320		
9	34459425	183783600	416215800	520059540	389449060	177331440	47324376	6636960	362880	
10	654729075	3928374450	10199989800	14980405440	13642629000	7934927000	2920525608	647536032	76998240	3628800

ment $2\nu - j + 2$ can be included in $2\nu - j + 1$ ways. (b) Element $2\nu - j + 2$ belongs to a cycle which contains two elements. The companion element of $2\nu - j + 2$ can be chosen in $2\nu - j + 1$ ways. The remaining $2\nu - j$ elements form a permutation with $\nu - j$ cycles each of length >1. The number of such permutations is $C(\nu, j)$. This proves (10).

From (10) it follows by mathematical induction that

$$\sum_{i=0}^{\nu-1} (-1)^i C(\nu, i) = 1 \tag{11}$$

for all $\nu \geqq 1$. By (10) it follows also that (8) and (9) are equivalent.

In proving the above results we shall make use of some properties of the Hermite polynomials and of the Stirling numbers of the first kind.

3. Hermite polynomials. The mth Hermite polynomial is defined by

$$H_m(x) = (-1)^m e^{x^2/2} D^m e^{-x^2/2} \tag{12}$$

for $m = 0, 1, 2, \cdots$ where $D = d/dx$ is the differential operator. Equivalently, we can write that

$$H_m(x) = e^{-D^2/2} x^m \tag{13}$$

for $m = 0, 1, 2, \cdots$. Definition (13) has many advantages over (12). For example, by (13), it follows immediately that

$$\sum_{m=0}^{\infty} \frac{H_m(x)}{m!} z^m = e^{-D^2/2} e^{xz} = e^{-z^2/2} e^{xz} \tag{14}$$

and $DH_m(x) = mH_{m-1}(x)$ for $m \geqq 1$. We mention that

$$H_m(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2} (x - iu)^m \, du \tag{15}$$

for $m \geqq 0$. This follows from (13) or from (5). Furthermore, we have

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} x^n \, dx = i^n H_n(0) = \begin{cases} 1.3 \cdots (2m-1) & \text{if } n = 2m, \\ 0 & \text{if } n = 2m-1, \end{cases} \tag{16}$$

and $m = 1, 2, \cdots$.

4. Stirling numbers of the first kind. We define $S(n, k)$ for $0 \leqq k \leqq n$ as the number of permutations of $(1, 2, \cdots, n)$ with k cycles, and $S(0, 0) = 1$. We define also $S(n, k) = 0$ for $k > n \geqq 0$. The numbers $S(n, k)$, $0 \leqq k \leqq n$, are called Stirling numbers of the first kind. They can easily be determined by the recurrence formula

$$S(n+1, k) = S(n, k-1) + nS(n, k) \tag{17}$$

for $n \geqq 0$ and $k \geqq 1$ where $S(0, 0) = 1$, $S(n, 0) = 0$ for $n \geqq 1$, and $S(0, k) = 0$ for $k \geqq 1$. Formula (17) reflects the fact that in a permutation of $(1, 2, \cdots, n+1)$, the element $n+1$ may form a cycle by itself or it may belong to a cycle of length > 1. See Table 4 for $S(n, k)$, $0 \leqq k \leqq n \leqq 6$.

By (17) it follows that

$$\sum_{k=0}^{n} S(n, k) x^k = x(x+1) \cdots (x+n-1) \tag{18}$$

for $n \geqq 1$.

TABLE 4
$S(n, k)$

n \ k	0	1	2	3	4	5	6
0	1						
1	0	1					
2	0	1	1				
3	0	2	3	1			
4	0	6	11	6	1		
5	0	24	50	35	10	1	
6	0	120	274	225	85	15	1

Obviously,

$$S(n,n-\nu)=\sum_{i=0}^{\nu-1} C(\nu,i)\binom{n}{2\nu-i} \tag{19}$$

for $1 \leqq \nu \leqq n$ where $C(\nu, i)$, $0 \leqq i \leqq \nu - 1$, is the number of permutations of $(1, 2, \cdots, 2\nu - i)$ with $\nu - i$ cycles each of length >1. For, a permutation of $(1, 2, \cdots, n)$ with $n - \nu$ cycles may have $n + i - 2\nu$ cycles of length 1 and $\nu - i$ cycles of length >1 where $i = 0, 1, \cdots, \nu - 1$. The representation (19) was found by Jordan [4], [5 p. 150] in 1933. See also Ward [11] and Carlitz [1].

From (19) we obtain by inversion that

$$C(\nu,i)=\sum_{j=0}^{\nu-i}(-1)^{\nu-i-j}\binom{2\nu-i}{\nu+j}S(\nu+j,j) \tag{20}$$

for $0 \leqq i \leqq \nu - 1$.

5. The determination of $F(n)$. By definition $F(n)$ is the number of distinct forests having vertex set $\{1, 2, \cdots, n\}$. A forest may consist of $r = 1, 2, \cdots, n$ distinct trees of orders $t_1, t_2, \cdots, t_r$ where $t_1 + t_2 + \cdots + t_r = n$. By a formula of Cayley [2] the number of distinct trees with t labeled vertices is t^{t-2}. Thus we obtain that

$$F(n)=\sum_{r=1}^{n}\frac{1}{r!}\sum_{t_1+\cdots+t_r=n}\frac{n!}{t_1!t_2!\cdots t_r!}t_1^{t_1-2}t_2^{t_2-2}\cdots t_r^{t_r-2} \tag{21}$$

for $n \geqq 1$. Write $F(0) = 1$. Then

$$\sum_{n=0}^{\infty}\frac{F(n)}{n!}x^n=\exp\left\{\sum_{n=1}^{\infty}\frac{n^{n-2}}{n!}x^n\right\} \tag{22}$$

for $|x| \leqq e^{-1}$. If $|x| \leqq e^{-1}$, then the equation

$$ye^{-y}=x \tag{23}$$

has exactly one root $y = y(x)$ in the unit disk $|y| \leqq 1$, and by Lagrange's expansion we obtain that

$$y-\frac{y^2}{2}=\sum_{n=1}^{\infty}\frac{n^{n-2}}{n!}x^n \tag{24}$$

for $|x| \leqq e^{-1}$. Thus by (22) and (14)

$$\sum_{n=0}^{\infty} \frac{F(n)}{n!} x^n = e^{y-y^2/2} = \sum_{r=0}^{\infty} \frac{H_r(1)}{r!} y^r. \tag{25}$$

Since by Lagrange's expansion

$$y^r = [y(x)]^r = r \sum_{n=r}^{\infty} \frac{n^{n-r}}{n(n-r)!} x^n \tag{26}$$

for $|x| \leqq e^{-1}$ and $r \geqq 1$, by (25) and (26) we get

$$F(n) = \sum_{r=1}^{n} \binom{n-1}{r-1} n^{n-r} H_r(1) \tag{27}$$

for $n \geqq 1$. Now by (13)

$$H_r(1) = [e^{-D^2/2} x^r]_{x=1}, \tag{28}$$

and it follows from (27) that

(29)

$$F(n) = \left[e^{-D^2/2} \sum_{r=1}^{n} \binom{n-1}{r-1} n^{n-r} x^r \right]_{x=1} = [e^{-D^2/2} x(n+x)^{n-1}]_{x=1}$$

$$= \left[e^{-D^2/2}(n+x)^n \right]_{x=1} - n[e^{-D^2/2}(n+x)^{n-1}]_{x=1} = H_n(n+1) - nH_{n-1}(n+1)$$

for $n \geqq 1$. This proves (3), and by (4) we obtain (2).

By using (5) we can also express (3) in the following form:

$$F(n) = H_{n+1}(n+1) - nH_n(n+1) \tag{30}$$

for $n \geqq 0$. By (15) and (30) we obtain (6) and the following formula

$$F(n) = \sum_{r=0}^{n} \binom{n}{r} n^{n-r} H_{r+1}(1) \tag{31}$$

for $n \geqq 0$.

6. The asymptotic expansion of $F(n)$. Since by (18)

$$\binom{n}{r} = \frac{1}{r!} \sum_{j=0}^{r} (-1)^{r-j} S(r,j) n^j \tag{32}$$

for $0 \leqq r \leqq n$, by (31) we have

$$F(n)n^{-n} = \sum_{r=0}^{n} \frac{1}{r!} H_{r+1}(1) \sum_{j=0}^{r} (-1)^{r-j} S(r,j) \frac{1}{n^{r-j}} = e^{1/2} \sum_{\nu=0}^{n} \frac{f_\nu(n)}{n^\nu} \tag{33}$$

where

$$(-1)^\nu e^{1/2} f_\nu(n) = \sum_{j=0}^{n-\nu} \frac{S(j+\nu, j)}{(j+\nu)!} H_{j+\nu+1}(1) \tag{34}$$

for $0 \leqq \nu \leqq n$. The limit $\lim_{n\to\infty} f_\nu(n) = f_\nu$ exists for $\nu = 0, 1, 2, \cdots$ and we have

$$(-1)^\nu e^{1/2} f_\nu = \sum_{j=0}^{\infty} \frac{S(j+\nu, j)}{(j+\nu)!} H_{j+\nu+1}(1). \tag{35}$$

If $\nu = 0$, then $S(j+\nu, j) = 1$ and

$$e^{1/2} f_0 = \sum_{j=0}^{\infty} \frac{H_{j+1}(1)}{j!} = [e^{-D^2/2} e^x x]_{x=0} = e^{1/2} H_1(0) = 0, \tag{36}$$

that is, $f_0 = 0$. If $\nu \geqq 1$, then by (19)

(37)

$$\begin{aligned}(-1)^\nu e^{1/2} f_\nu &= \sum_{j=0}^{\infty} \frac{H_{j+\nu+1}(1)}{(j+\nu)!} \sum_{k=0}^{\nu-1} C(\nu,k) \binom{j+\nu}{2\nu-k} \\ &= \sum_{k=0}^{\nu-1} C(\nu,k) \frac{1}{(2\nu-k)!} \sum_{j=\nu-k}^{\infty} \frac{H_{j+\nu+1}(1)}{(j+k-\nu)!} = e^{1/2} \sum_{k=0}^{\nu-1} \frac{C(\nu,k)}{(2\nu-k)!} H_{2\nu+1-k}(0).\end{aligned}$$

Here we used that

$$\begin{aligned}\sum_{j=\nu-k}^{\infty} \frac{H_{j+\nu+1}(1)}{(j+k-\nu)!} &= \left[e^{-D^2/2} \sum_{j=\nu-k}^{\infty} \frac{x^{j+\nu+1}}{(j+k-\nu)!} \right]_{x=1} = [e^{-D^2/2} e^x x^{2\nu+1-k}]_{x=1} \\ &= e^{1/2} H_{2\nu+1-k}(0).\end{aligned} \tag{38}$$

If $k = 2i$ $(0 \leqq i \leqq \nu)$, then $H_{2\nu+1-k}(0) = 0$ and if $k = 2i-1$ $(1 \leqq i \leqq \nu)$, then by (4)

$$H_{2\nu-2i+2}(0) = \frac{(-1)^{\nu-i+1}(2\nu-2i+2)!}{2^{\nu-i+1}(\nu-i+1)!}. \tag{39}$$

Since $H_3(0) = 0$, by (37) $f_1 = 0$, and if $\nu \geqq 2$, then by (37) and (39)

$$f_\nu = \sum_{i=1}^{[\nu/2]} (-1)^{i-1} \frac{C(\nu, 2i-1)}{2^{\nu-i}(\nu-i)!}. \tag{40}$$

For $0 \leqq \nu \leqq n$ we have

$$|f_\nu - f_\nu(n)| \leqq e^{-1/2} \sum_{k=n+1}^{\infty} \frac{S(k, k-\nu)}{k!} |H_{k+1}(1)| \tag{41}$$

where

$$\frac{|H_{k+1}(1)|}{\sqrt{(k+1)!}} \leqq e^{1/4} 1.086435 \cdots < 1.396 \tag{42}$$

and

$$\lim_{k\to\infty} \frac{S(k, k-\nu)}{k^{2\nu}} = \frac{C(\nu, 0)}{(2\nu)!} = \frac{1}{2^\nu \nu!}. \tag{43}$$

Accordingly, for every $\nu \geqq 0$, $\lim_{n \to \infty} f_\nu(n) = f_\nu$, and the rate of convergence is exponential. Consequently

$$F(n)n^{-n} \sim e^{1/2} \sum_{\nu=0}^{\infty} \frac{f_\nu}{n^\nu} \tag{44}$$

as $n \to \infty$ where $f_0 = f_1 = 0$ and f_ν for $\nu \geqq 2$ is given by (40). This proves (7) and (8).

REFERENCES

[1] L. CARLITZ, *Note on the numbers of Jordan and Ward*, Duke Math. J., 38 (1971), pp. 783–790.

[2] A. CAYLEY, *A theorem on trees*, Quarterly Journal of Pure and Applied Mathematics, 23 (1889), pp. 376–378. (The Collected Mathematical Papers of Arthur Cayley. Vol. XIII. Cambridge University Press, 1897, pp. 26–28.)

[3] J. DÉNES, *The representation of a permutation as the product of a minimal number of transpositions, and its connection with the theory of graphs*, Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 4 (1959), pp. 63–71.

[4] C. JORDAN, *On Stirling's numbers*, Tôhoku Mathematical Journal, 37 (1933), pp. 254–278.

[5] ———, *Calculus of Finite Differences*, Budapest, 1939. Reprinted by Chelsea, New York, 1947.

[6] D. J. KLEITMAN AND K. J. WINSTON, *Forests and score vectors*, Combinatorica, 1 (1981), pp. 49–54.

[7] A. RÉNYI, *Some remarks on the theory of trees*, Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 4 (1959), pp. 73–85.

[8] J. RIORDAN, *Forests of labeled trees*, J. Combin. Theory, 5 (1968), pp. 90–103.

[9] N. J. A. SLOANE, *A Handbook of Integer Sequences*, Academic Press, New York, 1973.

[10] R. STANLEY, *Decompositions of rational convex polytopes*, Ann. Discrete Math., 6 (1980), pp. 333–342.

[11] M. WARD, *The representation of Stirling's numbers and Stirling's polynomials as sums of factorials*, Amer. J. Math., 56 (1934), pp. 87–95.