JOURNAL OF COMBINATORIAL THEORY, Series A 40, 55–61 (1985)

Ballot Sequences and a Determinant of Good's

I. P. GOULDEN AND D. M. JACKSON

Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada

Communicated by the Managing Editors

Received September 15, 1983

1. INTRODUCTION

Consider the $n \times n$ determinant

$$D_n(x) = |x\mathbf{I} - n\mathbf{AC}|,$$

where $\mathbf{A} = [a_{ij}]_{n \times n}$ with $a_{ij} = \delta_{i,j-1} + \delta_{i-1,j}$, $\mathbf{C} = \mathbf{I} - n^{-1}\mathbf{J}$, $\mathbf{I}$ is the $n \times n$ identity matrix and $\mathbf{J}$ is the $n \times n$ matrix of all 1's. This determinant arises in the calculation of cumulants of a statistic analogous to Pearson's chi-squared for a multinomial sample (Sibson [5]), and its value has been conjectured by Good [2].

In this paper we use enumerative methods to prove Good's conjecture. The particular combinatorial structures we use in the enumeration are generalized ballot sequences, which correspond to random walks with two reflecting barriers. The generating function for this set is determined in Section 2. In Section 3, this generating function is combined with the result that $|\mathbf{I} + \mathbf{B}| = 1 + \mathrm{trace}(\mathbf{B})$ if $\mathrm{rank}(\mathbf{B}) = 1$, to give Good's determinant. A linear recurrence for $D_n(x)$ is also obtained.

It is worthwhile noting at the outset that the above determinant result is closely related to the following form of the Sherman–Morrison [4] "rank one update" formula

$$(\mathbf{I} + \mathbf{B})^{-1} = \mathbf{I} - (1 + \mathrm{trace}\,\mathbf{B})^{-1}\,\mathbf{B}$$

if $\mathrm{rank}(\mathbf{B}) = 1$. As may be seen in Goulden and Jackson ([3, Chap. 4]), matrices with rank equal to one arise naturally as incidence matrices in sequence enumeration problems. The two above results are particularly useful both in this context and in the evaluation of permanents by means of the MacMahon master theorem.

0097-3165/85 $3.00

2. Generalized Ballot Sequences

Let n be a fixed positive integer. A generalized ballot sequence from i to j on m steps is a sequence $(a_1, a_2, \ldots, a_{m+1})$ in which $1 \leqslant a_k \leqslant n$ for $k = 1, \ldots, m+1$, $|a_{k+1} - a_k| = 1$ for $k = 1, \ldots, m$, and $m \geqslant 0$, $a_1 = i$, $a_{m+1} = j$.

Let $F_{ij}(z)$ be the ordinary generating function for generalized ballot sequences from i to j with respect to the number of steps. The following well-known results involve the denominator polynomials Q_k, $k \geqslant 0$ for the continued fraction $1/1 - z^2/1 - z^2/1 - \cdots$, and proofs of (1) and (2) can be found in Flajolet [1] and Goulden and Jackson [3].

Proposition 2.1. *If* $\{Q_k(z) \mid k \geqslant 0\}$ *satisfies* $Q_k = Q_{k-1} - z^2 Q_{k-2}$, $k \geqslant 2$, *with* $Q_0 = Q_1 = 1$, *then*

(1) $F_{11}(z) = F_{nn}(z) = Q_{n-1}/Q_n$,

(2) $F_{1n}(z) = F_{n1}(z) = z^{n-1}/Q_n$,

(3) $Q_k = (\alpha^{k+1} - \beta^{k+1})/(\alpha - \beta)$, *where* $\alpha, \beta = (1 \pm \sqrt{1 - 4z^2})/2$, $k \geqslant 0$.

Proof. (3) $\{Q_k \mid k \geqslant 0\}$ satisfy a second-order linear difference equation with constant coefficients, so $Q_k = a\alpha^k + b\beta^k$, where α, β are the roots of the quadratic equation $1 - y + z^2 y^2 = 0$. The initial conditions $Q_0 = Q_1 = 1$ give $a + b = 1$, $a\alpha + b\beta = 1$ and the result follows. ∎

We have the following expression for the generating function

$$F_n(z) = \sum_{i=1}^{n} \sum_{j=1}^{n} F_{ij}(z)$$

for all generalized ballot sequences.

Theorem 2.2.

$$F_n(z) = \frac{n}{1-2z} - \frac{2z}{(1-2z)^2} + \frac{2z^2}{(1-2z)^2} (Q_{n-1} + z^{n-1})/Q_n.$$

Proof. If $(a_1, \ldots, a_{m+1})$ is a generalized ballot sequence from i to j, with $m \geqslant 1$, then $a_2 = i-1$ or $a_2 = i+1$, and $a_m = j-1$ or $a_m = j+1$. Thus

$$F_n(z) - n = z\left(F_n(z) - \sum_{j=1}^{n} F_{nj}(z)\right) + z\left(F_n(z) - \sum_{j=1}^{n} F_{1j}(z)\right)$$

$$\sum_{j=1}^{n} F_{1j}(z) - 1 = z\left(\sum_{j=1}^{n} F_{1j}(z) - F_{1n}(z)\right) + z\left(\sum_{j=1}^{n} F_{1j}(z) - F_{11}(z)\right)$$

and

$$\sum_{j=1}^{n} F_{nj}(z)-1=z\left(\sum_{j=1}^{n} F_{nj}(z)-F_{nn}(z)\right)+z\left(\sum_{j=1}^{n} F_{nj}(z)-F_{n1}(z)\right).$$

Eliminate $\sum_{j=1}^{n} F_{nj}(z)$ and $\sum_{j=1}^{n} F_{1j}(z)$ between these three equations to give

$$\begin{aligned}F_n(z)=(1-2z)^{-1}\{n-z(1-2z)^{-1}\\ \times(2-z\{F_{11}(z)+F_{1n}(z)+F_{n1}(z)+F_{nn}(z)\})\}.\end{aligned}$$

The result follows by applying Proposition 2.1. ∎

The coefficient of z^m in $F_n(z)$, denoted by $[z^m]\,F_n(z)$, gives the number of generalized ballot sequences on m steps. This number can be determined as an explicit but cumbersome triple summation involving a binomial coefficient, but we do not give the details here.

In the next section we evaluate Good's determinant $D_n(x)$ as a polynomial in x and derive explicit expressions for its coefficients. F_n is used in these calculations since, as we shall see, it is a factor of Good's determinant.

3. Good's Determinant

As an initial simplification, note that $D_n(x)=n^{-1}x^nE_n(nx^{-1})$, where $E_n(z)=n\,|\mathbf{I}-z\mathbf{AC}|$. Now $E_n(z)$ can be evaluated in terms of the generating function $F_n(z)$, with the aid of the following result.

Proposition 3.1. $E_n(z)=|\mathbf{I}-z\mathbf{A}|\ \mathrm{trace}(\mathbf{I}-z\mathbf{A})^{-1}\,\mathbf{J}$.

Proof.

$$\begin{aligned}E_n(z)&=n\,|\mathbf{I}-z\mathbf{A}+n^{-1}z\mathbf{AJ}|\\&=n\,|\mathbf{I}-z\mathbf{A}|\ |\mathbf{I}+n^{-1}z(\mathbf{I}-z\mathbf{A})^{-1}\,\mathbf{AJ}|\\&=n\,|\mathbf{I}-z\mathbf{A}|\,\{1+n^{-1}z\ \mathrm{trace}(\mathbf{I}-z\mathbf{A})^{-1}\,\mathbf{AJ}\},\end{aligned}$$

since rank $(\mathbf{J})=1$. But $z\mathbf{A}=\mathbf{I}-(\mathbf{I}-z\mathbf{A})$ and $\mathrm{trace}(\mathbf{J})=n$, and the result follows. ∎

By interpreting $\mathrm{trace}(\mathbf{I}-z\mathbf{A})^{-1}\,\mathbf{J}$ in the preceding expression for $E_n(z)$ combinatorially, we obtain from Section 2 an expression for $E_n(z)$ in terms of the denominator polynomials.

THEOREM 3.2.

$$E_n(z)=\left\{\frac{n}{1-2z}-\frac{2z}{(1-2z)^2}\right\}Q_n+\frac{2z^2}{(1-2z)^2}Q_{n-1}+\frac{2z^{n+1}}{(1-2z)^2}.$$

Proof. By expanding $G_n=|\mathbf{I}-z\mathbf{A}|$ by its first row, we see that G_n and Q_n satisfy the same recurrence equation, with the same initial conditions, so $|\mathbf{I}-z\mathbf{A}|=Q_n$.

From Goulden and Jackson [3, p. 245],

$$\operatorname{trace}(\mathbf{I}-z\mathbf{A})^{-1}\,\mathbf{J}=F_n(z),$$

since the (i, j)-element of $\mathbf{A}^m$ is the number of generalized ballot sequences from i to j on m steps. Thus, from Proposition 3.1, $E_n(z)=Q_n(z)\,F_n(z)$ and the result follows from Theorem 2.2. ▮

The determinants D_n and E_n are polynomials of degree n. To calculate $[x^k]\,D_n(x)$ we first observe that this is equal to $[x^{-1}]\,x^{-(k+1)}\,D_n(x)$, and then use the following result for the formal residue of a Laurent series. Its proof is given in Goulden and Jackson [3, p. 15].

PROPOSITION 3.3. *Let $h(z)$ be a Laurent series with a finite number of terms with negative exponents, and $r(z)$ be a power series with leading term of order $\rho>0$. Then*

$$\rho[z^{-1}]\,h(z)=[u^{-1}]\,h(r(u))\,r'(u).$$

The main theorem follows by combining the previous result with Theorem 3.2.

THEOREM 3.4.

$$D_n(x)=\sum_{k=0}^{n} d_k x^k$$

where

$$d_k=n^{n-k-1}k\left\{(\psi_s+1)\binom{-k-1}{s}+2\sum_{j=0}^{s-1}\binom{-k-1}{j}\right\}$$

and $s=\lfloor(n-k)/2\rfloor$, $\psi_s=0$ *if* s *even*, $\psi_s=1$ *if* s *odd.*

Proof. We have

$$d_k=[x^k]\,n^{-1}x^nE_n(nx^{-1})=n^{n-k-1}[z^{n-k}]\,E_n(z).$$

But $[z^{n-k}]\,E_n(z) = [z^{-1}]\,z^{-(n-k+1)}E_n(z)$. Now let $z = u(1+u^2)^{-1}$. Then $\alpha = (1+u^2)^{-1}$, $\beta = u^2(1+u^2)^{-1}$, so $Q_i = (1-u^{2i+2})(1-u^2)^{-1}(1+u^2)^{-i}$ from Proposition 2.1(3). Also $1-2z = (1-u)^2(1+u^2)^{-1}$, $dz/du = (1-u^2)(1+u^2)^{-2}$. Thus from Proposition 3.3 with $\rho = 1$, $[z^{n-k}]\,E_n(z) = [u^{-1}]\,u^{-(n-k+1)}(1+u^2)^{n-k+1}\,E_n(u(1+u^2)^{-1})(1-u^2)(1+u^2)^{-2}$ and from Theorem 3.2,

$$
\begin{aligned}
[z^{n-k}]\,E_n(z) &= [u^{n-k}](1+u^2)^{n-k-1}\left(\left\{\frac{n(1+u^2)}{(1-u)^2}\right.\right.\\
&\quad \left.\left.-\frac{2u(1+u^2)}{(1-u)^4}\right\}\frac{(1-u^{2n+2})}{(1+u^2)^n}+\frac{2u^2(1-u^{2n})}{(1-u)^4\,(1+u^2)^{n-1}}\right)+0\\
&= [u^{n-k}](1+u^2)^{-k}\,\{n(1-u)^{-2}-2u(1-u)^{-3}\}+0\\
&= n\sum_{i=0}^{\lfloor (n-k)/2\rfloor}\binom{-k}{i}(n-k-2i+1)\\
&\quad -2\sum_{i=0}^{\lfloor (n-k-1)/2\rfloor}\binom{-k}{i}\binom{n-k-2i+1}{2}\\
&= \sum_{i=0}^{s}\binom{-k}{i}(n-k-2i+1)\{(k+i)+i\}\\
&= k\sum_{i=0}^{s}\binom{-k-1}{i}(n-k-2i+1)\\
&\quad -k\sum_{i=1}^{s}\binom{-k-1}{i-1}(n-k-2i+1)
\end{aligned}
$$

and the result follows by replacing the second summation index i by $j+1$. ∎

The following corollary is Good's [2] conjecture.

COROLLARY 3.5. (1) *If* $n-k=2m+1$ *then* $d_k = n^{2m}2(n-2m-1)\sum_{j=0}^{m}\binom{2m-n}{j}$.

(2) *If* $n-k=2m$ *then* $d_k = n^{2m-1}(n-2m)\sum_{j=0}^{m}\binom{2m-n}{j}$.

Proof. (1) This is immediate from Theorem 3.4.

(2) In this case, Theorem 3.4 gives

$$
\begin{aligned}
d_k &= n^{2m-1}(n-2m)\left\{\binom{2m-n-1}{m}+2\sum_{j=0}^{m-1}\binom{2m-n-1}{j}\right\}\\
&= n^{2m-1}(n-2m)\left\{\sum_{j=0}^{m-1}\binom{2m-n-1}{j}+\sum_{j=0}^{m}\binom{2m-n-1}{j}\right\}.
\end{aligned}
$$

Now in the second summation set

$$\binom{2m-n-1}{j}=\binom{2m-n}{j}-\binom{2m-n-1}{j-1},$$

and the result follows. ∎

Finally, we give a four-term linear recurrence equation for $E_n(z)$.

PROPOSITION 3.6. *$D_n(x)=(x/(x-2n))\,H_n(x,n)$, where $\{H_i(x,\lambda)\mid i\geqslant 0\}$ satisfy the recurrence*

$$H_i=(x+\lambda)\,H_{i-1}-\lambda(x+\lambda)\,H_{i-2}+\lambda^3H_{i-3}, \qquad i\geqslant 2,$$

where $H_{-1}=0$, $H_0=1$, $H_1=x-2$.

Proof. From Theorem 3.1, $E_n(z)=R_n(z,n)$, where, for $n\geqslant 0$,

$$R_n(z,\lambda)=\left\{\frac{\lambda}{1-2z}-\frac{2z}{(1-2z)^2}\right\}Q_n(z)+\frac{2z^2}{(1-2z)^2}Q_{n-1}(z)+\frac{2z^{n+1}}{(1-2z)^2},$$

and $Q_{-1}(z)=0$. Let $R(z,\lambda,y)=\sum_{n\geqslant 0}R_n(z,\lambda)\,y^n$. Then, from the above equation,

$$R(z,\lambda,y)=\left\{\frac{\lambda}{1-2z}-\frac{2z}{(1-2z)^2}+\frac{2z^2y}{(1-2z)^2}\right\}$$
$$\times\sum_{n\geqslant 0}Q_n(z)\,y^n+\frac{2z}{(1-2z)^2\,(1-zy)}.$$

But we immediately deduce $\sum_{n\geqslant 0}Q_n(z)\,y^n=(1-y+z^2y^2)^{-1}$ from Proposition 2.1. If $T(x,\lambda,y)=\sum_{n\geqslant 0}T_n(x,\lambda)\,y^n=\lambda^{-1}R(\lambda x^{-1},\lambda,y)$, then from the above expression for R,

$$T(x,\lambda,y)=\frac{x}{x-2\lambda}\left(\left\{1-\frac{2}{x-2\lambda}+\frac{2\lambda y}{x-2\lambda}\right\}\right.$$
$$\left.\times(1-xy+\lambda^2y^2)^{-1}+\frac{2}{(x-2\lambda)(1-\lambda y)}\right).$$

Now $\lambda^{-1}R(\lambda x^{-1},\lambda,xy)=\sum_{n\geqslant 0}\lambda^{-1}R_n(\lambda x^{-1},\lambda)\,x^ny^n$, so $T_n(x,n)=D_n(x)$, from the result at the beginning of Section 2. Thus $D_n(x)=(x/(x-2n))$ $H_n(x,n)$, where $T_n(x,\lambda)=(x/(x-2\lambda))\,H_n(x,\lambda)$, so from the expression for T,

$$\sum_{n\geqslant 0}H_n(x,\lambda)\,y^n=\left\{1-\frac{2}{x-2\lambda}+\frac{2\lambda y}{x-2\lambda}\right\}$$
$$\times(1-xy+\lambda^2y^2)^{-1}+\frac{2}{(x-2\lambda)(1-\lambda y)}.$$

Multiplying this equation by $(1-xy+\lambda^2y^2)(1-\lambda y)$ yields

$$(1-(x+\lambda)\,y+\lambda(x+\lambda)\,y^2-\lambda^3y^3)\sum_{n\geqslant 0} H_n(x,\lambda)\,y^n = 1-(\lambda+2)\,y,$$

and the result follows by equating coefficients of y^i. ∎

ACKNOWLEDGMENTS

The authors would like to thank Ira Gessel for bringing this problem to their attention. This work has been supported by Grants U0073 and A8235 from the Natural Sciences and Engineering Council of Canada.

REFERENCES

1. P. FLAJOLET, Combinatorial aspects of continued fractions, *Discrete Math.* **32** (1980), 125–162.
2. I. J. GOOD, The cumulants of an analogue of Pearson's chi-squared, *J. Statist. Comput. Simulation* **15** (1982), 171–181.
3. I. P. GOULDEN AND D. M. JACKSON, "Combinatorial Enumeration," Wiley, New York, 1983.
4. J. SHERMAN AND W. J. MORRISON, Adjustments of an inverse matrix corresponding to changes in the elements of a given row or a given column of the original matrix, *Ann. Math. Statist.* **20** (1949), 621.
5. R. SIBSON, Discussion of Barndorff–Neilsen and D. R. Cox, *J. Roy. Statist. Soc. Ser. B* **41** (1979), 304–305.