Determinants and Alternating Sign Matrices

DAVID P. ROBBINS AND HOWARD RUMSEY, JR.

Institute for Defense Analyses, Thanet Road, Princeton, New Jersey 08540


1. Introduction

By a connected minor of a matrix we mean a minor formed from consecutive rows and consecutive columns. We will give formulas which express the determinant of a matrix in terms of certain connected minors. These formulas are obtained by using a method of Dodgson [2] (sometimes known as Lewis Carroll) for computing determinants. Both formulas involve alternating sign matrices, which we shall describe below and which are discussed at greater length in [6, 7]. Alternating sign matrices appear to have extremely close connections with the descending plane partitions introduced by Andrews in [1]. Their study led in [6] to a proof of Macdonald's conjecture [1; 5, p. 53] on cyclically symmetric plane partitions.

Let $M$ be an $n$ by $n$ matrix. We denote its entries by $M_{ij}$ where $i$ and $j$ vary from 1 to $n$. We will assume until the last section that all the connected minors of $M$ are non-zero. Let $M_{\text{NW}}$, $M_{\text{NE}}$, $M_{\text{SW}}$, $M_{\text{SE}}$ be the $n-1$ by $n-1$ connected minors in the northwest, northeast, southwest, and southeast corners of $M$ and let $M_{C}$ be the central $n-2$ by $n-2$ minor. Then it is known that

$$M_{C} \det M = M_{\text{NW}}M_{\text{SE}} - M_{\text{NE}}M_{\text{SW}}.$$

(1)
If we assign the value 1 to minors of size 0, then the identity is valid for \( n \geq 2 \). This identity was derived by Dodgson in [2] as a special case of an identity due to Jacobi [4, p. 9; 8, p. 77] relating the minors of a matrix to the minors of its adjugate (or cofactor) matrix. We will call it Dodgson's identity. Dodgson pointed out in [2] that (1) could be used iteratively to compute the determinant of a matrix. Given all the values of all connected minors of sizes \( k - 1 \) and \( k \), one uses (1) to compute all the connected minors of size \( k + 1 \). Since we know the minors of sizes 0 and 1, we can use (1) repeatedly to obtain \( \det M \). The method is mentioned in some texts on numerical analysis. For example, see [3].

If we think of the connected minors of size \( k \) and \( k + 1 \) as known, then Dodgson's method will yield a formula expressing the determinant of \( M \) as a rational function of these minors. In Theorem 1 we give an explicit formula for this rational function. It has several interesting properties. In particular, it can be written as a sum of terms each of which is plus or minus a monomial in which each variable appears with exponent 1, -1, or 0. In Theorem 2 we give a modification of this formula expressing \( \det M \) in terms of minors of three consecutive sizes, \( k - 1 \), \( k \), and \( k + 1 \). This formula has a similar form but it has the additional property that the denominator in each monomial involves only the minors of size \( k \). In both formulas the terms are indexed by sets of alternating sign matrices.

2. THE RATIONAL FUNCTIONS FOR MINORS OF TWO CONSECUTIVE SIZES

For each pair of positive integers \((i, j)\) we introduce indeterminates \( x_{ij} \) and \( y_{ij} \). In addition we let \( \lambda \) be another indeterminate. We will work in the field of rational functions in the \( x \)'s, \( y \)'s, and \( \lambda \) with coefficients in an arbitrary field. For any rational function \( R \) in our indeterminates we let \( T(R) \) be the result of simultaneously replacing

\[
y_{ij} \quad \text{by} \quad x_{ij}
\]

and

\[
x_{ij} \quad \text{by} \quad \frac{x_{ij}x_{i+1,j+1} + \lambda x_{i+1,j}x_{ij} + 1}{y_{i+1,j+1}}.
\]

Let \( R_k \) be the rational function which results when we apply the substitution \( T \) to the rational function \( y_{11} \) \( k \) times in succession. For example, we have

\[
R_0 = y_{11}, \quad R_1 = x_{11}, \quad R_2 = \frac{x_{11}x_{22} + \lambda x_{21}x_{12}}{y_{22}}.
\]
Let $\lambda = -1$. In $R_k$ substitute the connected minors of size $n-k+1$ and size $n-k$ whose upper left entries are $M_{ij}$ for $x_{ij}$ and $y_{ij}$, respectively. Then $R_k$ becomes the determinant of $M$. One may verify this directly for $k=0, 1, 2$. It follows for larger $k$ by induction once we observe that the substitution $T$ corresponds to expressing the size $n-k+1$ connected minors in terms of the size $n-k$ and $n-k-1$ connected minors with the help of Dodgson's identity. Thus a formula for the $R$'s will yield a formula for the determinant. We derive such a formula with \( \lambda \) indeterminate since the derivation is no more difficult and has some interesting applications.

3. The First Formula for the $R$'s

We need to introduce some extra notation in order to state our formula conveniently.

**Definition.** A square matrix is an alternating sign matrix if it satisfies:

(i) all entries are 1, -1, or 0,
(ii) every row and column has sum 1,
(iii) in every row and column the non-zero entries alternate in sign.

Note that in every row and column of an alternating sign matrix the first and last non-zero entries must be a 1 so that all the partial sums of every row and column must be 0 or 1.

We denote by $\mathcal{A}_n$ the set of $n$ by $n$ alternating sign matrices. All permutation matrices are alternating sign matrices. These are the only alternating sign $1$ by $1$ and $2$ by $2$ matrices. There are exactly seven $3$ by $3$ alternating sign matrices, the six $3$ by $3$ permutation matrices and the matrix $\begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. 

Numerous conjectures concerning alternating sign matrices are described in [7]. For example, there is overwhelming evidence that the number $|\mathcal{A}_n|$ of $n$ by $n$ alternating sign matrices is given by

$$|\mathcal{A}_n| = \prod_{i=0}^{n-1} \frac{(3i+1)!}{(n+1)!},$$

but this has not yet been proved.
Let $A$ be an $n$ by $n$ matrix. We define the corner sum matrix $\bar{A}$ of $A$ by

$$\bar{A}_{ij} = \sum A_{kl}$$

where the sum ranges over all pairs of integers $(k, l)$ with $k \leq i$ and $l \leq j$ and we regard $A_{kl}$ as 0 if $k$ or $l$ is out of the range $1, \ldots, n$. $A$ can be recovered from $\bar{A}$ by taking mixed second differences:

$$A_{ij} = \bar{A}_{ij} - \bar{A}_{i-1,j-1} - \bar{A}_{i-1,j} + \bar{A}_{i-1,j-1}.$$

Observe that the differences $\bar{A}_{ij} - \bar{A}_{i-1,j}$ and $\bar{A}_{ij} - \bar{A}_{i,j-1}$ are partial sums of the rows and columns of $A$. Using this observation one may prove the following characterization of alternating sign matrices in terms of corner sums. We omit the details.

**Lemma 1.** An $n$ by $n$ matrix $A$ is an alternating sign matrix if and only if $\bar{A}$ satisfies:

(i) $\bar{A}_{ii} = \bar{A}_{nn} = i$ for $i = 1, \ldots, n$,

(ii) $\bar{A}_{ij} - \bar{A}_{i-1,j}$ and $\bar{A}_{ij} - \bar{A}_{i,j-1}$ are always 0 or 1 for $1 \leq i, j \leq n$.

**Definition.** Let $A$ be a $k - 1$ by $k - 1$ alternating sign matrix and $B$ a $k$ by $k$ alternating sign matrix. We say that $A$ and $B$ are compatible and write $A \preceq B$ if we have all the inequalities

(i) $\bar{A}_{ij} \geq \bar{B}_{ij}$,

(ii) $\bar{A}_{ij} \geq \bar{B}_{i+1,j+1} - 1$,

(iii) $\bar{A}_{ij} \leq \bar{B}_{i+1,j}$,

(iv) $\bar{A}_{ij} \leq \bar{B}_{i,j+1}$,

for $0 \leq i, j \leq k - 1$.

**Remark.** The preceding form of the definition of compatibility is the only one that we shall need. There is, however, a more symmetrical version in which we require only (i) and the analogous conditions on the other three corners of $A$ and $B$.

One may verify that the single 1 by 1 alternating sign matrix is compatible with both of the 2 by 2 alternating sign matrices and that identity matrices of two consecutive sizes are always compatible.

We shall need two notions concerning alternating sign matrices which are similar to inversions for ordinary permutation matrices. We define the first now. We call these flips.
For an \( n \) by \( n \) alternating sign matrix \( A \) we define its number of flips, \( F(A) \), by
\[
F(A) = |I_n| - |A|.
\]
Here the \( |A| \) is the sum of the entries \( A_{ij} \) for \( 1 \leq i, j \leq n \). \( I_n \) is the \( n \) by \( n \) identity matrix. The number of flips is always a non-negative integer.

Let \( u_{ij} \) be any array of rational functions defined for pairs of positive integers \( (i, j) \) and let \( A \) be a \( k \) by \( k \) matrix with integer entries. It is convenient to define
\[
u^A = \prod_{i,j=1}^{k} u_{ij}^{A_{ij}}.
\]

Given any array of rational functions \( u_{ij} \), we define two new arrays \( s(u) \) and \( d(u) \) by
\[
s(u)_{ij} = u_{i+1,j+1}
\]
and
\[
d(u)_{ij} = u_{ij}u_{i+1,j+1} + \lambda u_{i+1,j}u_{i,j+1}.
\]

When we do arithmetic operations on arrays of rational functions, these operations are done pointwise. For example, if \( u \) and \( v \) are two arrays, then their product \( uv \) is defined by \( (uv)_{ij} = u_{ij}v_{ij} \).

Now we can state the formula for \( R_k \).

**Theorem 1.** For \( k \geq 2 \),
\[
R_k = \sum \lambda^{F(B) - F(A)} x^B s(y)^A,
\]
where the sum is over all compatible pairs \( (A, B) \) of alternating sign matrices \( B \) in \( \mathcal{A}_k \) and \( A \) in \( \mathcal{A}_{k-1} \).

Here \( R_k \) is the rational function defined at the beginning of Section 2. \( \mathcal{A}_k \) is the set of \( k \) by \( k \) alternating sign matrices, \( F \) is the function giving the number of flips of an alternating sign matrix, and \( s \) is a shift operator. If \( \lambda = -1 \) and, for some \( n \) by \( n \) matrix \( M \), \( x \) and \( y \) are the arrays of connected minors of \( M \) of sizes \( n-k+1 \) and \( n-k \), then \( R_k \) will be its determinant.

**Proof.** We may check directly that the theorem is true when \( k = 2 \). We proceed by induction. Suppose we already know the theorem for \( k \). Then, making the substitution \( T \), we have
\[ R_{k+1} = \sum_{A \approx B} \lambda^{F(B)} d(x)^B s(y)^{-B} s(x)^{-A} \]
\[ = \sum_{B} \left( \lambda^{F(B)} d(x)^B s(y)^{-B} \sum_{A, A \approx B} \lambda^{-F(A)} s(x)^{-A} \right). \]  \hspace{1cm} (3)

On the other hand we must show that

\[ R_{k+1} = \sum_{B \approx C} \lambda^{F(C)} - F(B) x^C s(y)^{-B} \]
\[ = \sum_{B} \left( \lambda^{-F(B)} s(y)^{-B} \sum_{C, C \approx B} \lambda^{F(C)} x^C \right). \] \hspace{1cm} (4)

In these two equations \( A, B, \) and \( C \) are, respectively, alternating sign matrices of sizes \( k - 1, k, \) and \( k + 1. \)

Thus we need to show that the expressions on the right sides of (3) and (4) are equal.

Now fix a size \( k \) alternating sign matrix \( B. \) It will suffice to show that

\[ \sum_{A, A \approx B} \lambda^{-F(A)} s(x)^{-A} = \sum_{C, C \approx B} \lambda^{F(C)} x^C. \] \hspace{1cm} (5)

The rest of our proof depends on a description of the sets of size \( k - 1 \) and size \( k + 1 \) alternating sign matrices that are compatible with our fixed size \( k \) alternating sign matrix \( B. \)

Let us begin with a description of the \( A's. \) Our definition of compatibility (2) can be restated as requiring that the entries of \( A \) satisfy the inequality

\[ \max_{i, j} (A_{i,j}, A_{i,j+1} - 1) \leq A_{i,j} \leq \min(\bar{B}_{i,j+1}, \bar{B}_{i+1,j}) \] \hspace{1cm} (6)

for all \((i, j)\) with \(0 \leq i, j \leq k - 1.\)

**Lemma 2.** For every \((i, j)\) with \(0 \leq i, j \leq k - 1,\) the maximum on the left of (6) is either equal to the minimum on the right side of (6) or is 1 less than it. Moreover the left side is less than the right side exactly when \(B_{i+1,j+1} = -1.\)

**Proof.** Part (ii) of Lemma 1 implies that the max does not exceed the min in (6). Also their difference cannot exceed the difference between \(\bar{B}_{i,j}\) and \(\bar{B}_{i,j+1}\) which is \(\leq 1.\) The difference is 1 if and only if

\[ \bar{B}_{i,j} + 1 = \bar{B}_{i+1,j+1} = \bar{B}_{i,j+1} = \bar{B}_{i+1,j}. \] \hspace{1cm} (7)

Taking mixed second differences, we have \(B_{i+1,j+1} = -1.\) Conversely, if \(B_{i+1,j+1} = -1,\) then the partial row sum \(\sum_{s \leq j+1} B(i+1,s) = 0.\) This
implies that \( B_{i+1,j+1} = B_{i,j+1} \). Similarly, \( B_{i+1,j+1} = B_{i,j+1} \). The first equality in (7) now follows by taking differences.

**Lemma 3.** Any matrix \( A \) with integer entries for which \( A \) satisfies (6) is an alternating sign matrix.

**Proof.** We show that \( A \) satisfies the conditions of Lemma 1. First we have

\[
A_{i,j} \leq \min(\overline{B}_{i,j+1}, \overline{B}_{i+1,j}) \leq \max(\overline{B}_{i,j+1}, \overline{B}_{i+1,j+2}) \leq \overline{A}_{i,j+1},
\]

which shows that \( A \) is weakly increasing in the second coordinate. Also

\[
A_{i,j+1} - A_{ij} \leq \min(\overline{B}_{i,j+2}, \overline{B}_{i+1,j+1}) - \max(\overline{B}_{ij}, \overline{B}_{i+1,j+1} - 1)
\]

\[
\leq \overline{B}_{i+1,j+1} - (\overline{B}_{i+1,j+1} - 1) = 1.
\]

which shows that the steps are always 0 or 1. If \( 1 \leq i \leq k-1 \), then

\[
i \leq \max(\overline{B}_{ik-1}, \overline{B}_{i+1,k-1}) \leq \overline{A}_{i,k-1} \leq \min(\overline{B}_{ik}, \overline{B}_{i+1,k-1}) \leq i,
\]

which shows that \( \overline{A} \) satisfies the condition (i) of Lemma 1. 

Let \( A \) be a matrix with integer entries. We define its "positive part" \( A^+ \) to be the result of replacing all the negative entries of \( A \) by 0 (and leaving the remaining entries unchanged). By the "negative part" \( A^- \) of \( A \) we mean the result of replacing all the positive entries of \( A \) by 0 and then negating the matrix. With this notation we have \( A = A^+ - A^- \).

Now let \( A_0 \) be the \( k-1 \) by \( k-1 \) alternating sign matrix compatible with \( B \) with minimum corner sum matrix; that is, from (6), the matrix \( A_0 \) with

\[
\overline{A}_{ij} = \max(\overline{B}_{ij}, \overline{B}_{i+1,j+1} - 1)
\]

for all \((i, j)\) with \( 0 \leq i, j \leq k-1 \). Then all other compatible \( A \)'s can be obtained from this one by adding 1's to a subset of those \( \overline{A}_{ij} \)'s for which \( B_{i+1,j+1} = -1 \). For these pairs \((i, j)\) the effect of one of these additions on \( A_0 \) is to add 1 to the entries of \( A_0 \) with indices \((i, j)\) and \((i+1, j+1)\) and to subtract 1 from those entries with indices \((i+1, j)\) and \((i, j+1)\). The effect on \( s(x)^{-A_0} \) is to multiply it by \( f_{i+1,j+1} \) where

\[
f_{ij} = \frac{x_{ij+1}x_{i+1,j}}{x_{ij}x_{i+1,j+1}}.
\]

Since 1 has been added to a single entry of the corner sum matrix, the effect
on $\lambda^{-F(A^0)}$ is to multiply by $\lambda$. Thus the sum on the left side of (5) can be written as

$$\sum_{A, A \approx B} \lambda^{-F(A^0)} s(x)^{-A} = \lambda^{-F(A^0)} s(x)^{-A^0} \prod (1 + \lambda f_{i+1,j+1}) \quad (8)$$

The product on the right side is over all pairs $(i, j)$ with $B_{i+1,j+1} = -1$. After replacing $i$ and $j$ by $i - 1$ and $j - 1$ in this product, we can rewrite (8) as

$$\sum_{A, A \approx B} \lambda^{-F(A^0)} s(x)^{-A^0} \left( \frac{d(x)}{xs(x)} \right)^{B_{i,j}} \quad (9)$$

Now we make a similar calculation for the sum on the right of (5). If we rewrite the conditions for compatibility in terms of the $c$'s, we see that they imply that

$$\max(g_{i,j}, 1, B_{i,j}, \alpha) \leq C_{i,j} \leq \min(B_{i,j}, B_{i+1,j+1} + 1) \quad (10)$$

for all $(i, j)$ with $1 \leq i, j \leq k$. Then we have the following two lemmas analogous to Lemmas 2 and 3. We omit their proofs.

**Lemma 4.** For every $(i, j)$ with $1 \leq i, j \leq k$, the maximum on the left of (10) is either equal to the minimum on the right side of (10) or 1 less than it. Moreover the left side is less than the right side exactly when $B_{i,j} = 1$.

**Lemma 5.** Any $k + 1$ by $k + 1$ matrix $C$ with integer entries, such that $C$ satisfies (10) and the condition (i) of Lemma 1, is an alternating sign matrix.

Now we let $C^0$ be the $k + 1$ by $k + 1$ alternating sign matrix compatible with $B$ that has maximum corner sum matrix. That is, we let

$$C^0_{i,j} = \min(B_{i,j}, B_{i-1,j+1} + 1)$$

for all $(i, j)$ with $1 \leq i, j \leq k$ and require that it satisfy the condition (i) of Lemma 1. Then all other compatible $C$'s can be obtained from this one by subtracting 1's from a subset of those $C^0$'s for which $B_{i,j} = 1$. The effect of one of these subtractions on $C^0$, for these pairs $(i, j)$, is to subtract 1 from the entries of $C^0$ with indices $(i, j)$ and $(i+1, j+1)$ and to add 1 to those entries with indices $(i+1, j)$ and $(i, j+1)$. The effect on $x^{C^0}$ is to multiply it by $f_{i,j}$. Since 1 has been subtracted from a single entry of the corner sum matrix, the effect on $\lambda^{F(C^0)}$ is to multiply by $\lambda$. Thus the sum on the left side of (5) can be written as

$$\sum_{C, B \approx C} \lambda^{F(C)} x^C = \lambda^{F(C^0)} x^{C^0} \prod (1 + \lambda f_{i,j}) \quad (11)$$
where the product on the right side of (11) is over all pairs \((i, j)\) with \(B_{ij} = 1\). We can rewrite (11) as

\[
\sum_{C \cong B} z_{F(B)}^{-1} f^C = z_{F(B)}^{-1} f^C \left( \frac{d(x)}{xs(x)} \right)^{B^*}. \tag{12}
\]

Now, combining (5), (9), and (12), we see that we have reduced the proof of Theorem 1 to showing that

\[
z_{2F(B)}^{2F(B)} d(x)^B z^{-F(A^0)} s(x)^{-A^0} \left( \frac{d(x)}{xs(x)} \right)^B = z_{F(C^0)}^{-1} f^C \left( \frac{d(x)}{xs(x)} \right)^{B^*}.
\]

For this it suffices to show that

\[
s(x)^{A^0} x^{C^0} = (xs(x))^B \tag{13}
\]

and that

\[
F(A^0) + F(C^0) = 2F(B). \tag{14}
\]

From the definitions of \(\overline{A}^0\) and \(\overline{C}^0\) we have

\[
\overline{A}^0_{i,j-1} + \overline{C}^0_{i,j} = \max(\overline{B}_{i,j-1}, \overline{B}_{i,j} - 1) + \min(\overline{B}_{j,i}, \overline{B}_{i-1,j} + 1)
\]

whenever \(1 \leq i, j \leq k\). One may verify directly that (15) also holds if \(i\) or \(j\) is 0, and, from the condition (i) of Lemma 1, that (15) also holds if \(i\) or \(j\) is \(k + 1\). Thus we may take mixed second differences and conclude, recalling that terms with out of range indices are to be regarded as zero, that

\[
A^0_{i-1,j} + C^0_{i,j} = B_{ij} + B_{i-1,j-1} \tag{16}
\]

for all \(i\) and \(j\). But the two sides of (16) are the exponents of \(x_{ij}\) on the two sides of (13). Thus we have proved (13).

It is easily verified that

\[
|F_k| = \frac{k(k + 1)(2k + 1)}{6}.
\]

Now using the definition of \(F\), we can reduce the proof of Eq. (14) to showing that

\[
|A^0| + |C^0| = 2|B| + 2k + 1. \tag{17}
\]
If we sum both sides of (15) for \(1 \leq i, j \leq k\), then we obtain
\[
|\mathcal{A}_0^i| + |\mathcal{C}_0^i| - [1 + 2 + \cdots + k + (k + 1) + k + \cdots + 1] = |\mathcal{B}| - [1 + 2 + \cdots + (k - 1) + k + (k - 1) + \cdots + 1],
\]
which is equivalent to (17). This completes the proof of Theorem 1.

4. Formula in Terms of Minors of Three Consecutive Sizes

Now we describe our alternate formula. We introduce a new array \(z\) of rational functions which satisfy
\[
xs(z) = d(y).
\]
If \(\lambda = -1\) and the \(x_{ij}\) and \(y_{ij}\) represent the size \(n - k + 1\) and size \(n - k\) minors with upper left entry having indices \((i, j)\), then, still assuming that none of the connected minors are 0, \(z_{ij}\) will be the size \(n - k - 1\) minor whose upper left entry has indices \((i, j)\). We will express \(R_k\) as a sum of monomials in the \(x\)'s, \(y\)'s, and \(z\)'s with only \(y\)'s in the denominator.

Using Theorem 1 and substituting \(y\) for \(x\) in (9), we have
\[
R_k = \sum_{A \prec B} \lambda^{F(B) - F(A)} x^B y^A
\]
\[
= \sum_{B} \lambda^{F(B)} x^B \sum_{A, A \prec B} \lambda^{-F(A)} y^A
\]
\[
= \sum_{B} \lambda^{F(B)} - F(A_0^B) x^B S(y) - A_0^B \left( \frac{d(y)}{S(y)} \right)^B.
\]

Here we have used the notation \(A^0(B)\) to indicate the dependence of \(A^0\) on \(B\). Using the definition of the \(z\)'s and simplifying, we have
\[
R_k = \sum_{B \in \mathcal{S}_k} \lambda^{F(B) - F(A_0^B)} x^B + s(z)^B \frac{s(y)^A}{s(y)^{A_0^B} \left( S(y) \right)^B}.
\]

Introduce the abbreviations
\[
P(B) = F(B) - F(A_0^B) \quad \text{(18)}
\]
and
\[
B^*_y = A_0^y + B^*_{i+1,j+1} + B^*_{ij}. \quad \text{(19)}
\]
In terms of these we have proved
ALTERNATING SIGN MATRICES

Theorem 2a. For \( k \geq 2 \) we have

\[
R_k = \sum_{B \in \mathcal{A}_k} \lambda_{P(B)} \frac{x^B s(z)^B}{s(y)^B^*}.
\]

Here \( R_k \) is the rational function defined at the beginning of Section 2. \( \mathcal{A}_k \) is the set of \( k \) by \( k \) alternating sign matrices. We shall see in Theorem 2c that the function \( P \) generalizes the function giving the number of inversions of a permutation matrix. \( B^+ \) and \( B^- \) are the positive and negative parts of \( B \) defined in Section 3. \( B^* \) is a \( k-1 \) by \( k-1 \) matrix depending on \( B \). Its construction is defined earlier in this section. An interesting alternate description is given in Theorem 2b below. The function \( s \) is a shift operator defined in Section 3. If \( \lambda = -1 \) and, for some \( n \) by \( n \) matrix \( M \), \( x, y, \) and \( z \) are the arrays of connected minors of \( M \) of sizes \( n-k+1, n-k, \) and \( n-k-1 \), then \( R_k \) will be its determinant. Since only the \( y \)'s appear in the denominators, this is an expression for the determinant that has the form promised earlier.

While the preceding theorem was easy to prove, we would prefer a simpler description of the two functions \( B^* \) and \( P(B) \). We derive these descriptions in the remainder of this section.

We begin with a geometric description of \( B^* \). This will yield the form in which our results were originally discovered. Let \( B \) be any \( k \) by \( k \) alternating sign matrix, \( k \geq 2 \). Let \( B_{i,j} \) and \( B_{i,j+1} \) be a pair of adjacent entries in a row of \( B \). If \( x_i = 1 \), we place an arrow between these two entries which points to the left. If the sum is 0, we place an arrow pointing to the right. We place vertical arrows similarly between every pair of adjacent entries in the same column. For example, if

\[
R = \begin{pmatrix}
0 & 1 & 0 & 0 \\
0 & -1 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix},
\]

then, after we insert the arrows, we have

\[
\begin{pmatrix}
0 & \rightarrow & 1 & \leftarrow & 0 & \leftarrow & 0 \\
\downarrow & \uparrow & \downarrow & \downarrow & \downarrow \\
1 & \leftarrow & -1 & \rightarrow & 0 & \rightarrow & 1 \\
\uparrow & \downarrow & \downarrow & \uparrow & \uparrow \\
0 & \rightarrow & 0 & \rightarrow & 1 & \leftarrow & 0 \\
\uparrow & \downarrow & \uparrow & \uparrow & \uparrow \\
0 & \rightarrow & 1 & \leftarrow & 0 & < & 0
\end{pmatrix}.
\]
Fill in the boxes formed by the arrows according to the following rules:

(i) If two opposite sides of a box have parallel arrows, then fill the box with a 0.

(ii) If the arrows circulate around the box in either direction, then fill the box with a -1.

(iii) Otherwise fill the box with a 1.

For example, if \( B \) is as above, then the resulting picture is

\[
\begin{array}{ccccccc}
0 & \rightarrow & 1 & \leftrightarrow & 0 & \leftrightarrow & 0 \\
\downarrow & 1 & \uparrow & 1 & \downarrow & 0 & \downarrow \\
1 & \leftrightarrow & -1 & \rightarrow & 0 & \rightarrow & 1 \\
\uparrow & 1 & \downarrow & 0 & \downarrow & 1 & \uparrow \\
0 & \rightarrow & 0 & \rightarrow & 1 & \leftrightarrow & 0 \\
\uparrow & 0 & \downarrow & 1 & \uparrow & 0 & \uparrow \\
0 & \rightarrow & 1 & \leftrightarrow & 0 & \leftrightarrow & 0
\end{array}
\]

**Theorem 2b.** If \( B \) is any \( n \) by \( n \) alternating sign matrix, then \( B^* \) is the \( n - 1 \) by \( n - 1 \) matrix formed inside the boxes of arrows according to the preceding rules.

For example, when \( B \) is given as above, then

\[
B^* = \begin{pmatrix}
1 & 1 & 0 \\
1 & 0 & 1 \\
0 & 1 & 0
\end{pmatrix}
\]

To begin our proof of Theorem 2b we need to interpret the rules for constructing \( B^* \) algebraically. Define matrices of partial sums on the first and second subscripts of \( B \) by

\[
B_{ij}^{(1)} = \sum_{r \leq i} B_{rj}
\]

and

\[
B_{ij}^{(2)} = \sum_{s \leq j} B_{is}.
\]

Then \( B_{ij}^{(1)} = 1 \) exactly when the arrow between \( B_{ij} \) and \( B_{i+1,j} \) points upward and \( B_{ij}^{(2)} = 1 \) exactly when the arrow between \( B_{ij} \) and \( B_{i,j+1} \) points to the left. Then the rules (i), (ii), and (iii) above are equivalent to

\[
B_{ij}^* = (B_{i,j+1}^{(1)} - B_{ij}^{(1)})(B_{i+1,j}^{(2)} - B_{ij}^{(2)}),
\]

(20)
when \(1 \leq i, j \leq k - 1\) and \(B_{ij}^* = 0\) for other values of \(i\) and \(j\). Our proof proceeds by restating the formula (19), which defines \(B^*\), in terms of the \(B^{(1)}\)'s and the \(B^{(2)}\)'s and then showing that it is equivalent to (20).

Since negative entries never appear on the boundary of an alternating sign matrix, we may conclude from (19) that \(B_{ij}^* = 0\) unless \(1 \leq i, j \leq k - 1\). Rewrite (19) as

\[
B_{ij}^* = B_{ij}^+ + B_{i+1,j+1}^- + (A_{ij}^0 - B_{ij}).
\]  

(21)

We will find expressions for each of the three terms on the right of (21) in terms of the \(B^{(1)}\)'s and the \(B^{(2)}\)'s.

Suppose that \(1 \leq i, j \leq k - 1\). We have

\[
B_{ij}^+ = B_{ij}^+ B_{ij}^{(1)} = (B_{ij}^{(2)} - B_{i,j+1}^{(2)}) B_{ij}^{(1)}
\]  

(22)

since \(B_{ij}^{(1)} = 0\) when \(B_{ij} = -1\). Similarly,

\[
B_{i+1,j+1}^- = -(B_{i+1,j+1}^{(1)} - B_{i,j+1}^{(1)}) B_{i+1,j}^{(2)}.
\]  

(23)

From the definition of \(A^0\) we have

\[
A_{ij}^0 - B_{ij} = \max(B_{ij}, B_{i+1,j+1} - 1) - B_{ij}
\]

if \(0 \leq i, j \leq k - 1\). The right side has the value 1 when \(B_{i+1,j+1}\) exceeds \(B_{ij}\) by 2 and is 0 otherwise. It follows that

\[
A_{ij}^0 - B_{ij} = (B_{i+1,j+1} - B_{i+1,j})(B_{i+1,j} - B_{ij}) = B_{i+1,j+1}^{(1)} B_{i+1,j}^{(2)}
\]

(24)

\[
=B_{i+1,j+1} - B_{i,j+1})(B_{ij} - B_{i,j+1}) = B_{i+1,j+1}^{(2)} B_{i,j+1}^{(1)}.
\]

When we take mixed second differences of both sides of (24), we obtain

\[
A_{ij}^0 - B_{ij} = B_{i+1,j+1}^{(1)} B_{i+1,j}^{(2)} - B_{i,j+1}^{(1)} B_{i+1,j}^{(2)} - B_{ij}^{(1)} B_{i,j+1}^{(2)} + B_{ij}^{(1)} B_{i,j}^{(2)},
\]  

(25)

for \(1 \leq i, j \leq k - 1\) and where, in the third term on the right, we have used the second form in (24). Now (20) and Theorem 2b follow by adding Eqs. (22), (23), and (25).

Next we give a simplified description of the function \(P(B)\). For any \(k\) by \(k\) alternating sign matrix \(B\) we define the number \(I(B)\) of inversions of \(B\) by

\[
I(B) = \sum_{i,j=1}^{k} B_{i-1,j}^{(1)} B_{i,j-1}^{(2)}.
\]  

(26)
This definition generalizes the usual notion of inversions for permutation matrices. It is equivalent to the definition given in [7] which is

\[ I(B) = \sum B_{ij} B_{rs}, \]

where the sum is over all \( i, j, r, s \) with \( i < r \) and \( j > s \).

**Theorem 2c.** \( P(B) = I(B) - N(B) \), where \( N(B) \) is the number of negative entries in \( B \).

**Remark.** If \( B_{ij} = -1 \), then the corresponding term of (26) is a 1. We call these inversions "negative inversions" and the remaining positive terms of (26) "positive inversions." Thus Theorem 2c states that \( P(B) \) is the number of positive inversions of \( B \). In [7] evidence is given which suggests that the positive inversions of an alternating sign matrix correspond to the nonspecial parts of a descending plane partition.

**Proof of Theorem 2c.** Using the definition of \( F \) and (18), we have

\[ P(B) = |A^0| - |\overline{B}| + k^2. \]

Now sum equation (24) for \( 0 \leq i, j \leq k - 1 \). We obtain

\[ P(B) = |A^0| - (|\overline{B}| - (1 + \cdots + k + \cdots + 1)) \]

\[ = \sum_{i,j=1}^{k} B_{ij}^{(1)} B_{i,j-1}^{(2)} \]

\[ = \sum_{i,j=1}^{k} (B_{i-1,j}^{(1)} B_{i,j-1}^{(2)} + B_{ij} B_{i,j-1}^{(2)}) \]

\[ = I(B) - N(B). \]

5. **Examples. The \( \lambda \)-Determinant**

One may define determinants inductively by assigning the value 1 to size 0 determinants, the usual value to 1 by 1 determinants, and then obtaining larger determinants using Dodgson's identity. We now consider a similar function, which we call the \( \lambda \)-determinant. We denote the \( \lambda \)-determinant of a matrix \( M \) by \( \det_{\lambda} M \). We begin the induction the same way but we require instead that it satisfy the following \( \lambda \) form of Dodgson's identity:

\[ M_C \det_{\lambda} M = M_{NW} M_{SE} + \lambda M_{NE} M_{SW}, \]
for square matrices $M$ of size $\geq 2$. Here the subscripted $M$'s represent $\lambda$-minors rather than ordinary minors as in Section 1.

Let $M$ be an $n$ by $n$ matrix with all its $\lambda$-minors non-zero. From the recursion defining $R_k$ it is straightforward to check that when we substitute for the $x$'s and $y$'s the connected $\lambda$-minors of size $n - k + 1$, $n - k$ of the matrix $M$, then $R_k$ becomes the $\lambda$-determinant of $M$.

We may use the case $k = n$ of Theorem 2 to obtain a formula for the $\lambda$-determinant of $M$. Then the $y$'s are all 1's, so that the $z$'s satisfy $s(z) x = 1 + \lambda$. After substituting the entries of $M$ for the $x$'s, we obtain, in the obvious notation,

$$\det_\lambda M = \sum_{B \in \mathcal{A}_n} \lambda^{P(B)} (1 + \lambda)^{N(B)} M^B.$$  

This formula makes sense whenever the entries of $M$ are non-zero so it provides an extension of the definition of the $\lambda$-determinant to this wider class of matrices. The $\lambda$-determinant, even when extended, will continue to satisfy the $\lambda$ form of Dodgson's identity.

If we take $\lambda = -1$, then the $\lambda$-determinant becomes the ordinary determinant and all the terms on the right side of (27) except those corresponding to permutation matrices vanish. Thus (27) becomes the usual expression for a determinant as a sum over all permutations.

If we let $M$ be the all 1's matrix, then, using the $\lambda$ form of Dodgson's identity, we can evaluate the $\lambda$-determinant as $(1 + \lambda)^{n(n-1)/2}$. Combining this observation with the case $\lambda = 1$ of (27), we obtain

$$\sum_{B \in \mathcal{A}_n} 2^{N(B)} = 2^{n(n-1)/2},$$

a result which also appears in [7]. Note that if we could somehow replace the 2 by a 1 inside the summation sign in the preceding equation, we would be able to enumerate the $n$ by $n$ alternating sign matrices.

We remark that the $\lambda$-determinant shares some properties with ordinary determinants. For example, if $M$ is the $n$ by $n$ Vandermonde matrix with entries $M_{ij} = a_i^{j-1}$, $i, j = 1, \ldots, n$, then

$$\det_\lambda M = \prod (a_r + \lambda a_s),$$

where the product is over all pairs $(r, s)$ with $1 \leq s < r \leq n$.

REFERENCES