Alternating-Sign Matrices and Domino Tilings (Part I)*

NOAM ELKIES Harvard University, Cambridge, MA 02138

GREG KUPERBERG University of California at Berkeley, Berkeley, CA 94720

MICHAEL LARSEN University of Pennsylvania, Philadelphia, PA 19104

JAMES PROPP Massachusetts Institute of Technology, Cambridge, MA 02139

Received October 10, 1991; Revised March 24, 1992

Abstract. We introduce a family of planar regions, called Aztec diamonds, and study tilings of these regions by dominoes. Our main result is that the Aztec diamond of order n has exactly $2^{n(n+1)/2}$ domino tilings. In this, the first half of a two-part paper, we give two proofs of this formula. The first proof exploits a connection between domino tilings and the alternating-sign matrices of Mills, Robbins, and Rumsey. In particular, a domino tiling of an Aztec diamond corresponds to a compatible pair of alternating-sign matrices. The second proof of our formula uses monotone triangles, which constitute another form taken by alternating-sign matrices; by assigning each monotone triangle a suitable weight, we can count domino tilings of an Aztec diamond.

Keywords: tiling, domino, alternating-sign matrix, monotone triangle, representation, square ice

1. Introduction

The Aztec diamond of order n is the union of those lattice squares $[a, a + 1] \times [b, b + 1] \subset \mathbb{R}^2$ $(a, b \in \mathbb{Z})$ that lie completely inside the tilted square $\{(x, y) : |x| + |y| \le n + 1\}$. (Figure 1 shows the Aztec diamond of order 3.) A domino is a closed 1×2 or 2×1 rectangle in \mathbb{R}^2 with corners in \mathbb{Z}^2 , and a tiling of a region R by dominoes is a set of dominoes whose interiors are disjoint and whose union is R. In this paper we will show that the number of domino tilings of the Aztec diamond of order n is $2^{n(n+1)/2}$. We will furthermore obtain more refined enumerative information regarding two natural statistics of a tiling: the number of vertical tiles and the "rank" of the tiling (to be defined shortly).

Fix a tiling T of the Aztec diamond of order n. Every horizontal line y = k divides the Aztec diamond into two regions of even area; it follows that the

*Part II will appear in the next issue.



Fig. 1. Aztec diamond of order 3.



Fig. 2. The rotation move.

number of dominoes that straddle the line must be even. Letting k vary, we see that the total number of vertical dominoes must be even; accordingly, we define v(T) as half the number of vertical tiles in T.

The most intuitively accessible definition of the rank statistic r(T) comes by way of the notion of an "elementary move," which is an operation that converts one domino tiling of a region into another by removing two dominoes that form a 2×2 block and putting them back rotated by 90° (see Figure 2). It will be shown that any domino tiling of an Aztec diamond can be reached from any other by a sequence of such moves; we may therefore define the *rank* of the tiling T as the minimum number of moves required to reach T from the "all-horizontals" tiling (shown in Figure 5(a)). Thus, all-horizontals tiling itself has rank 0, and the tiling shown on the right side of Figure 2 (viewed as a tiling of the order-1 Aztec diamond) has rank 1.

Let

$$AD(n; x, q) = \sum_{T} x^{\nu(T)} q^{r(T)},$$

where T ranges over all domino tilings of the order-n Aztec diamond; this is a polynomial in x and q. The main result of this paper is

THEOREM:

AD(n; x, q) =
$$\prod_{k=0}^{n-1} (1 + xq^{2k+1})^{n-k}$$
.

As important special cases we have

$$AD(n; q) = \prod_{k=0}^{n-1} (1 + q^{2k+1})^{n-k},$$

$$AD(n; x) = (1 + x)^{n(n+1)/2},$$

$$AD(n) = 2^{n(n+1)/2},$$

where we adopt the convention that an omitted variable is set equal to 1.

In this two-part article, we will give four ways of understanding the formula for AD(n). The first exploits the relationship between tilings of the Aztec diamond and the still fairly mysterious alternating-sign matrices introduced by Mills, Robbins, and Rumsey in [12]. Our second proof yields the formula for AD(n) as a special case of a theorem on monotone triangles (combinatorial objects closely related to alternating-sign matrices and introduced in [13]). The third proof comes from the representation theory of the general linear group. The last proof yields the more general formula for AD(n; x, q) and also leads to a bijection between tilings of the order-*n* diamond and bit strings of length n(n + 1)/2. We conclude by pointing out some connections between our results and the square ice model studied in statistical mechanics.

2. Height functions

It is not at all clear from the definition of rank given in Section 1 just how one would calculate the rank of a specific tiling; for instance, it happens that the all-verticals tiling of the order-*n* Aztec diamond has rank n(n+1)(2n+1)/6 and that every other tiling has strictly smaller rank, but it is far from obvious how one would check this. Therefore, we will now give a more technical definition of the rank and prove that it coincides with the definition given above. We use the vertex-marking scheme described in [21]; it is a special case of the boundary-invariants approach to tiling problems introduced in [3].

It will be conceptually helpful to extend a tiling T of the Aztec diamond to a tiling T^+ of the entire plane by tiling the complement of the Aztec diamond by horizontal dominoes in the manner shown in Figure 3 for n = 3. Let G be the graph with vertices $\{(a, b) \in \mathbb{Z}^2 : |a| + |b| \le n + 1\}$ and with an edge between (a, b) and (a', b') precisely when |a - a'| + |b - b'| = 1. Color the lattice squares of \mathbb{Z}^2 in a black-white checkerboard fashion so that the line $\{(x, y) : x + y = n + 1\}$ that bounds the upper-right border of the Aztec diamond passes through only white squares. Call this the standard (or even) coloring. Orient each edge of G so that a black square lies to its left and a white square lies to its right; this gives the standard orientation of the graph G, with arrows circulating clockwise around white squares and counterclockwise around black squares. (Figure 4 shows the case n = 3.) Write $u \to v$ if uv



Fig. 3. Extending a tiling to the whole plane.

is an edge of G whose standard orientation is from u to v. Call v = (a, b) a boundary vertex of G if |a| + |b| = n or n + 1, and let the boundary cycle be the closed zigzag path (-n-1, 0), (-n, 0), (-n, 1), (-n+1, 1), (-n+1, 2), ..., (-1, n), (0, n), (0, n+1), (0, n), (1, n), ..., (n+1, 0), ..., (0, -n-1), ..., (-n-1, 0). Call the vertex <math>v = (a, b) even if it is the upper-left corner of a white square (i.e., if a + b + n + 1 is even) and odd otherwise, so that in particular the four corner vertices (-n - 1, 0), (n + 1, 0), (0, -n - 1), (0, n + 1) are even.

If one traverses the six edges that form the boundary of any domino, one will follow three edges in the positive sense and three edges in the negative sense. Also, every vertex v of G lies on the boundary of at least one domino in T^+ . Hence, if for definiteness one assigns height 0 to the leftmost vertex (-n-1, 0) of G, there is for each tiling T a unique way of assigning integervalued heights $H_T(v)$ to all the vertices v of G, subject to the defining constraint that if the edge uv belongs to the boundary of some tile in T^+ with $u \to v$, then $H_T(v) = H_T(u) + 1$. The resulting function $H_T(\cdot)$ is characterized by two properties:

- (i) H(v) takes on the successive values $0, 1, 2, \ldots, 2n + 1, 2n + 2, 2n + 1, \ldots, 0, \ldots, 2n + 2, \ldots, 0$ as v travels along the boundary cycle of G;
- (ii) if $u \to v$, then H(v) is either H(u) + 1 or H(u) 3.

The former is clear, since every edge of the boundary cycle is part of the boundary of a tile of T^+ . To see that (ii) holds, note that if the edge uv belongs to T^+ (i.e., is part of the boundary of a tile of T^+), then H(v) = H(u) + 1, whereas if uv does not belong to T^+ , then it bisects a domino of T^+ , in which case we see (by considering the other edges of that domino) that H(v) = H(u) - 3.

In the other direction, notice that every height function $H(\cdot)$ satisfying (i) and (ii) arises from a tiling T and that the operation $T \mapsto H_T$ is reversible: given a



Fig. 4. Oriented edges of the square grid.

function H satisfying (i) and (ii), we can place a domino covering every edge uv of G with |H(u) - H(v)| = 3, obtaining thereby a tiling of the Aztec diamond, which will coincide with the original tiling T in the event $H = H_T$. Thus there is a bijection between tilings of the Aztec diamond and height functions $H(\cdot)$ on the graph G that satisfy (i) and (ii). For a geometric interpretation of $H(\cdot)$, see [21].

Figure 5 shows the height functions corresponding to two special tilings of the Aztec diamond, namely, (a) the all-horizontals tiling T_{\min} and (b) the all-verticals tiling T_{\max} . Since $H_T(v)$ is independent of T modulo 4, we are led to define the reduced height

$$h_T(v) = (H_T(v) - H_{T_{\min}}(v))/4.$$

The reduced-height function of T_{\min} is thus constantly zero; the reduced-height function of T_{\max} is as shown in part (c) of Figure 5. Last, we define the rank statistic

$$r(T) = \sum_{v \in G} h_T(v)$$

It is easy to verify that if one performs an elementary rotation on a 2-by-2 block centered at a vertex v (a "v-move" for short), the effect is to leave $h_T(v')$ alone for all $v' \neq v$ and to either increase or decrease $h_T(v)$ by 1; we call the moves raising and lowering, respectively.

We may now verify that r(T) (as defined by the preceding equation) is equal to the number of elementary moves required to get from T to T_{\min} . Since $r(T_{\min}) = 0$ and since an elementary move merely changes the reduced height of a single vertex by ± 1 , at least r(T) moves are required to get from T to T_{\min} . It remains to verify that for every tiling T there is a sequence of moves leading from T to T_{\min} in which only r(T) moves are made. To find such a sequence, let



Fig. 5. Height functions of the all-horizontals and all-verticals tilings.

 $T_0 = T$ and iterate the following operation for i = 0, 1, 2, ...: Select a vertex v_i at which $h_{T_i}(\cdot)$ achieves its maximum value. If $h_{T_i}(v_i) = 0$, then $T_i = T_{\min}$ and we are done. Otherwise, we have $h_{T_i}(v_i) > 0$, so that $H_{T_i}(v_i) > H_{T_{\min}}(v_i)$ with v_i not on the boundary of G (since h_{T_i} vanishes on the boundary). Reducing $H_{T_i}(v_i)$ by 4 preserves the legality of the height labeling and corresponds to performing a v_i -move on T_i , yielding a new tiling T_{i+1} with $r(T_{i+1}) = r(T_i) - 1$. By repeating this process, we continue to reduce the rank statistic by 1 until the procedure terminates at $T_{r(T)} = T_{\min}$.

Thus, we have shown that every tiling of the Aztec diamond may be reached from every other by means of moves of the sort described. This incidentally furnishes another proof that the number of dominoes of each orientation (horizontal or vertical) must be even, since this is clearly true of T_{\min} and since every move annihilates two horizontal dominoes and creates two vertical ones, or vice versa.

The partial ordering on the set of tilings of an Aztec diamond given by height

functions has a pleasant interpretation in terms of a two-person game. Let T, T' be tilings of the Aztec diamond of order n. We give player A the tiling T and player B the tiling T'. On each round, A makes a rotation move and B has the choice of either making the identical move (assuming it is available to her) or passing. Here, to make an *identical move* means to find an identically situated 2-by-2 block in the identical orientation and give it a 90° twist. If, after a certain number of complete rounds (i.e., moves by A and countermoves by B), A has solved her puzzle (that is, reduced the tiling to the all-horizontals tiling) and B has not, then A is deemed the winner; otherwise, B wins. Put $T' \leq T$ if and only if B has a winning strategy in this game. It is easy to verify (without even considering any facts about tilings) that the relation \preceq is reflexive, asymmetric, and transitive. In fact, $T' \leq T$ if and only if $h_T(v) \leq h_{T'}(v)$ for all $v \in G$. Moreover, the ideal strategy for either player is to make only lowering moves—although in the case $T' \leq T$, it turns out that B can win by copying A whenever possible, regardless of whether such moves are lowering or raising.

3. Alternating-sign matrices

An alternating-sign matrix is a square matrix (*n*-by-n, say) all of whose entries are 1, -1, and 0, such that every row sum and column sum is 1 and such that the nonzero entries in each row and column alternate in sign; for instance,

is a 4-by-4 alternating-sign matrix. (For an overview of what is currently known about such matrices, see [16].) Let \mathcal{A}_n denote the set of *n*-by-*n* alternating-sign matrices.

If A is an n-by-n alternating-sign matrix with entries a_{ij} $(1 \le i, j \le n)$, we may define

$$a_{ij}^* = i + j - 2\left(\sum_{i'=1}^{i}\sum_{j'=1}^{j}a_{i'j'}\right)$$

for $0 \le i, j \le n$. We call the (n + 1)-by-(n + 1) matrix A^* the skewed summation of A. (It is a variant of the corner-sum matrix of [17].) The matrices A^* that arise in this way are precisely those such that $a_{i0}^* = a_{0i}^* = i$ and $a_{in}^* = a_{ni}^* = n - i$ for $0 \le i \le n$ and such that adjacent entries of A' in any row or column differ by 1. Note that $a_{ij} = \frac{1}{2}(a_{i-1,j}^* + a_{i,j-1}^* - a_{i-1,j-1}^* - a_{i,j}^*)$, so that an alternating-sign matrix can be recovered from its skewed summation. Thus, the alternating-sign matrix A defined above has

as its skewed summation.

Our goal is to show that the domino tilings of the Aztec diamond of order n are in 1-to-1 correspondence with pairs (A, B), where $A \in \mathcal{A}_n$, $B \in \mathcal{A}_{n+1}$, and A, B jointly satisfy the compatibility relation introduced in [17]. We will do this by means of the height functions defined in Section 2.

Given a tiling T of the order-n Aztec diamond, we construct matrices A' and B' that record $H_T(v)$ for v odd and even, respectively (where $v = (x, y) \in G$ is even or odd according to the parity of x + y + n + 1). We let

$$a'_{ij} = H_T(-n+i+j, -i+j)$$

for $0 \le i, j \le n$ and

$$b'_{ij} = H_T(-n-1+i+j, -i+j)$$

for $0 \le i, j \le n + 1$; thus, the tiling of Figure 6 gives the matrices

	/ 1	3	5	7	٥ /		0	2	4	6	8	10
A' =		5	7	5	7	and $B' =$	2	4	6	4	6	8
	5	7	5	7	5		4	6	8	6	4	6
	7	5	3	, 5	3		6	4	6	4	6	4
		7	5	3	1		8	6	4	2	4	2
		•	-	-			10	8	6	4	2	0

Note that the matrix elements on the boundary of A' and B' are independent of the particular tiling T. Also note that in both matrices consecutive elements in any row or column differ by exactly 2. Therefore, under suitable normalization A' and B' can be seen as skewed summations of alternating-sign matrices A and B. Specifically, by setting $a_{ij}^* = (a_{ij}' - 1)/2$ and $b_{ij}^* = b_{ij}'/2$, we arrive at matrices A^* , B^* , which, under the inverse of the skewed summation operation, yield the



Fig. 6. An Aztec tiling converts into two alternating-sign matrices.

matrices A, B that we desire:

<i>A</i> * =	$ \left(\begin{array}{c} 0\\ 1\\ 2\\ 3\\ 4 \end{array}\right) $	1 2 3 2 3	2 3 2 1 2	3 2 3 2 1	4 3 2 1 0	and $B^* =$	$ \left(\begin{array}{c} 0\\ 1\\ 2\\ 3\\ 4\\ 5 \end{array}\right) $	1 2 3 2 3 4	2 3 4 3 2 3	3 2 3 2 1 2	4 3 2 3 2 1	5 4 3 2 1 0	;
<i>A</i> =) () : 1 () () 1)	1 1 0 1	0 1 0 0	and $B =$	(0 0 1 0 0	0 0 0 1	1 0 0 0	0 1 -1 0 1	0 0 1 0 0).	

Conversely, A and B determine A' and B', which determine H_T , which determines T.

There is an easy way of reading off A and B from the domino tiling T, without using height functions. First, note that the even vertices in the interior of the Aztec diamond of order n are arranged in the form of a tilted n-by-n square.

Also note that each such vertex is incident with 2, 3, or 4 dominoes belonging to the tiling T; if we mark each such site with a 1, 0, or -1 (respectively), we get the entries of A, where the upper-left corner of each matrix corresponds to positions near the left corner of the diamond. Similarly, the odd vertices of the Aztec diamond (including those on the boundary) form a tilted (n+1)-by-(n+1)square. If we mark each such site with a -1, 0, or 1 according to whether it is incident with 2, 3, or 4 dominoes of the extended tiling T^+ , we get the entries of B. (We omit the proof that this construction agrees with the one we gave earlier, since it is only the first one that we actually need.)

The legality constraint (ii) from the Section 2 tells us that for $1 \le i, j \le n$, the internal entries b'_{ij} of the matrix B' must be equal to

either
$$a'_{i-1,j-1} - 1$$
 or $a'_{i-1,j-1} + 3$,
either $a'_{i-1,j} - 3$ or $a'_{i-1,j} + 1$,
either $a'_{i,j-1} - 3$ or $a'_{i,j-1} + 1$, and
either $a'_{i,j} - 1$ or $a'_{i,j} + 3$.

Thus, in all but one of the six possible cases for the submatrix

$$\left(egin{array}{ccc} a'_{i-1,\,j-1} & a'_{i-1,\,j} \ a'_{i,\,j-1} & a'_{i,\,j} \end{array}
ight)$$

shown in Table 1, the value of b'_{ij} is uniquely determined; only in the case

$$\left(\begin{array}{rrr} 2k-1 & 2k+1\\ 2k+1 & 2k-1 \end{array}\right)$$

arising from $a_{ij} = 1$ does b'_{ij} have two possible values, namely, 2k - 2 and 2k + 2.

It now follows that if we hold A fixed, the number of (n + 1)-by-(n + 1) alternating-sign matrices B such that the pair (A, B) yields a legal height function is equal to $2^{N_+(A)}$, where $N_+(A)$ is the number of +1's in the n-by-n alternating-sign matrix A. That is,

$$AD(n) = \sum_{A \in \mathcal{A}_n} 2^{N_{\star}(A)}.$$
 (1)

Switching the roles of A and B, we may by a similar argument prove

$$AD(n) = \sum_{B \in \mathcal{A}_{n+1}} 2^{N_{-}(B)},$$
 (2)

$\left(egin{array}{ccc} a'_{i-1,j-1} & a'_{i-1,j} \ a'_{i,j-1} & a'_{i,j} \end{array} ight)$	a_{ij}	b_{ij}^\prime
$\left(\begin{array}{ccc} 2k-1 & 2k+1\\ 2k+1 & 2k+3 \end{array}\right)$	0	2 <i>k</i> + 2
$\left(\begin{array}{ccc} 2k+3 & 2k+1\\ 2k+1 & 2k-1 \end{array}\right)$	0	2 <i>k</i> + 2
$\left(\begin{array}{ccc} 2k+1 & 2k-1\\ 2k+3 & 2k+1 \end{array}\right)$	0	2 <i>k</i>
$\left(\begin{array}{ccc} 2k+1 & 2k+3\\ 2k-1 & 2k+1 \end{array}\right)$	0	2k
$\left(\begin{array}{ccc} 2k-1 & 2k+1\\ 2k+1 & 2k-1 \end{array}\right)$	1	2k - 2 or $2k + 2$
$\left(\begin{array}{ccc} 2k+1 & 2k-1\\ 2k-1 & 2k+1 \end{array}\right)$	-1	2k

Table 1. From domino tilings to alternating-sign matrices.

where $N_{-}(\cdot)$ gives the number of -1's in an alternating-sign matrix. Replacing n by n-1 and B by A in (2), we get

$$AD(n-1) = \sum_{A \in \mathcal{A}_n} 2^{N_-(A)}.$$
 (3)

On the other hand, $N_+(A) = N_-(A) + n$ for all $A \in \mathcal{A}_n$, so (1) tells us that

$$AD(n) = 2^n \sum_{A \in \mathcal{A}_n} 2^{N_-(A)}.$$
 (4)

Combining (3) and (4), we derive the recurrence relation

$$\mathrm{AD}(n)=2^{n}\mathrm{AD}(n-1),$$

which suffices to prove our formula for AD(n). (Mills, Robbins, and Rumsey [12] prove

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

as a corollary to their Theorem 2. The result also appears in [17].)

In the remainder of this section we discuss tilings and alternating-sign matrices from the point of view of lattice theory. Specifically, we show that the tilings of an order-*n* Aztec diamond correspond to the lower ideals (or down-sets) of a partially ordered set P_n , whereas the *n*-by-*n* alternating-sign matrices correspond to the lower ideals of a partially ordered set Q_n , such that P_n consists of a copy of Q_n interleaved with a copy of Q_{n+1} . (For terminology associated with partially ordered sets, see [20].)

We start by observing that the set of legal height functions H on the order-nAztec diamond is a poset in the obvious component-wise way, with $H_1 \ge H_2$ if $H_1(v) \ge H_2(v)$ for all $v \in G$. Moreover, the consistency conditions (i) and (ii) are such that if H_1 and H_2 are legal height functions, then so are $H_1 \lor H_2$ and $H_1 \land H_2$, defined by $(H_1 \lor H_2)(v) = \max(H_1(v), H_2(v))$ and $(H_1 \land H_2)(v) = \min(H_1(v), H_2(v))$; thus, our partially ordered set is actually a distributive lattice.

 A_n , the set of *n*-by-*n* alternating-sign matrices, also has a lattice structure. Given $A_1, A_2 \in A_n$, we form their skewed summations A_1^*, A_2^* and declare $A_1 \ge A_2$ if every entry of A_1^* is greater than or equal to the corresponding entry of A_2^* . This partial ordering on alternating-sign matrices is intimately connected with the partial ordering on tilings: If T_1, T_2 are tilings, then $T_1 \ge T_2$ if and only if $A_1 \ge A_2$ and $B_1 \ge B_2$, where (A_i, B_i) is the pair of alternating-sign matrices corresponding to the tiling T_i (i = 1, 2).

For each vertex v = (x, y) of the graph G associated with the order-n Aztec diamond (with $x, y \in \mathbb{Z}$, $|x| + |y| \le n + 1$), let $m = H_{T_{\min}}(v)$ and $M = H_{T_{\max}}(v)$, so that $m, m + 4, \ldots, M$ are the possible values of $H_T(v)$, and introduce points $(x, y, m), (x, y, m + 4), \ldots, (x, y, M - 4) \in \mathbb{Z}^3$ lying above the vertex v = (x, y). (Note that if (x, y) is on the boundary of G, m = M, so the set of points above (x, y) is empty.) Let P denote the set of all such points as v ranges over the vertex-set of G. We make P a directed graph by putting an edge from $(x, y, z) \in P$ to $(x', y', z') \in P$ provided z = z' + 1 and |x - x'| + |y - y'| = 1; we then make P a partially ordered set by putting $(x, y, z) \ge (x', y', z')$ if there is a sequence of arrows leading from (x, y, z) to (x', y', z').

To each height function H we may assign a subset $I_H \subseteq P$, with $I_H = \{(x, y, z) \in P : z < H(x, y)\}$. This operation is easily seen to be a bijection between the legal height functions H and the lower ideals of the partially ordered set P. Indeed, the natural lattice structure on the set of height functions H (with $H_1 \leq H_2$ precisely if $H_1(v) \leq H_2(v)$ for all $v \in G$) makes it isomorphic to the lattice J(P) of lower ideals of P, and the rank r(T) of a tiling T (as defined above) equals the rank of H_T in the lattice, which in turn equals the cardinality of I_{H_T} .

Note that for all $(x, y, z) \in P$, $x + y + z \equiv n + 1 \mod 2$. The poset P decomposes naturally into two complementary subsets P^{even} and P^{odd} , where a point $(x, y, z) \in P$ belongs to P^{even} if z is even and P^{odd} if z is odd. The vertices of P^{even} form a regular tetrahedral array of side n, resting on a side (as opposed to a face); that is, it consists of a 1-by-n array of nodes, above which lies a 2-by-(n-1) array of nodes, above which lies a 3-by-(n-2) array of nodes, and so on, up to the n-by-1 array of nodes at the top. The partial ordering of P restricted to P^{even} makes P^{even} a poset in its own right, with (x, y, z) covering (x', y', z') when z = z' + 2 and |x - x'| = |y - y'| = 1. Similarly, the vertices of P^{odd} form a tetrahedral array of side n-1; each vertex of P^{odd} lies at the center of a small tetrahedron with vertices in P^{even} . P^{odd} , like P^{even} , is a poset in itself, with (x, y, z) covering (x', y', z) covering (x', y', z') when z = z' + 2 and |x - x'| = |y - y'| = 1.

Our correspondence between height functions H_T and pairs (A, B) of alternating-sign matrices tells us that \mathcal{A}_n , as a lattice, is isomorphic to $J(P_n^{\text{odd}})$, and \mathcal{A}_{n+1} is isomorphic to $J(P_n^{\text{even}})$. Indeed, under this isomorphism, $A \in \mathcal{A}_n$ and $B \in \mathcal{A}_{n+1}$ are compatible if and only if the union of the down-sets of P^{even} and P^{odd} corresponding to A and B is a down-set of $P = P^{\text{even}} \cup P^{\text{odd}}$. (This coincides with the notion of compatibility given in [17].) If we let Q_n denote the tetrahedral poset P_n^{odd} (so that P_n^{even} is isomorphic to Q_{n+1}), then we see that P_n indeed consists of a copy of Q_n interleaved with a copy of Q_{n+1} .

As an aid to visualizing the poset P and its lower ideals, we may use stacks of marked 2-by-2-by-4/3 bricks resting on a special multilevel tray. The bottom face of each brick is marked by a line joining midpoints of two opposite edges, and the top face is marked by another such line, skew to the mark on the bottom face (see Figure 7). These marks constrain the ways in which we allow ourselves to stack the bricks. To enforce these constraints, whittle away the edges of the brick on the top and bottom faces that are parallel to the marks on those faces and replace each mark by a ridge, as in Figure 8; the rule is that a ridge on the bottom face of a brick must fit into the space between two whittled-down edges (or between a whittled-down edge and empty space). The only exception to this rule is at the bottom of the stack, where the ridges must fit into special furrows in the tray. Figure 9 shows the tray in the case n = 4; it consists of four levels, three of which float in midair. On the bottom level the outermost two of the three gently sloping parallel lines running from left to right should be taken as ridges and the one in between should be taken as a furrow. Similarly, in the higher levels of the tray the outermost lines are ridges and the innermost two are furrows. We require that the bricks resting on the table must occupy only the n obvious discrete positions; no intermediate positions are permitted. Also, a brick cannot be placed unless its base is fully supported by the tray, a tray and a brick, or two bricks.

In stacking the bricks, one quickly sees that in a certain sense one has little freedom in how to proceed; any stack one can build will be a subset of the stack shown in Figure 10 in the case n = 4. Indeed, if one partially orders the



Fig. 7. A cube with marked top and bottom.



Fig. 8. A beveled cube.

bricks in Figure 10 by the transitive closure of the relation "is resting on," then the poset that results is the poset P defined earlier and the admissible stacks correspond to lower ideals of P in the obvious way. Moreover, the markings visible to an observer looking down on the stack yield a picture of the domino tiling that corresponds to that stack.

4. Monotone triangles

Let A^* be the skewed summation of an *n*-by-*n* alternating-sign matrix A. Notice that the *i*th row $(0 \le i \le n)$ begins with an *i* and ends with an n - i, so that reading from left to right we must see *i* descents and n - i ascents; that is, there are exactly *i* values of *j* in $\{1, 2, ..., n\}$ satisfying $a'_{i,j} = a'_{i,j-1} - 1$, and the remaining n - i values of *j* satisfy $a'_{i,j} = a'_{i,j-1} + 1$. Form a triangular array



Fig. 9. The tray in which the cubes sit.



Fig. 10. Cubes stacked in the domino-tiling fashion.

whose ith row $(1 \le i \le n)$ consists of those values of j for which $a'_{i,j} = a'_{i,j-1} - 1$; e.g., for

,

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \text{ and } A' = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 1 & 2 & 3 \\ 2 & 1 & 2 & 1 & 2 \\ 3 & 2 & 3 & 2 & 1 \\ 4 & 3 & 2 & 1 & 0 \end{pmatrix}$$

we get the triangle

Note that j occurs in the *i*th row of the monotone triangle exactly if the sum of the first i entries in column j of the alternating-sign matrix is 1.

A monotone triangle of size n is a triangular array of natural numbers with strict increase from left to right along its n rows and with nonstrict increase from left to right along its diagonals, as in the array above. If the bottom row of a monotone triangle is $1 \ 2 \ \cdots \ n$, we call the array a *complete monotone triangle*. It is not difficult to show that the preceding construction gives a bijection between the n-by-n alternating-sign matrices and the complete monotone triangles of size n. Moreover, the +1's in the alternating-sign matrix correspond to entries in some row of the triangle that do not occur in the preceding row. (This correspondence, as well as the notion of a monotone triangle, was introduced in [13].) It follows from the foregoing that AD(n) is the sum, over all complete monotone triangles of size n, of 2 to the power of the number of entries in the monotone triangle that do not occur in the preceding row. Since a monotone triangle of size n has exactly n(n + 1)/2 entries, we may divide both sides of the equation $AD(n) = 2^{n(n+1)/2}$ by $2^{n(n+1)/2}$ and paraphrase it as the claim that the sum, over all complete monotone triangles of size n, of $\frac{1}{2}$ to the power of the number of entries in the monotone triangle that do occur in the preceding row is precisely 1.

Define the weight of a monotone triangle (of any size) as $\frac{1}{2}$ to the power of the number of entries that appear in the preceding row, and let $W(a_1, a_2, \ldots, a_k)$ be the sum of the weights of the monotone triangles of size k with bottom row $a_1 a_2 \cdots a_k$. (For now we may assume $a_1 < a_2 < \cdots < a_k$, although we will relax this restriction shortly.) Our goal is to prove that $W(1, 2, \ldots, n) = 1$ for all n.

To this end, observe that we have the recurrence relation

$$W(a_1, a_2, \ldots, a_n) = \sum_{b_1=a_1}^{a_2} \sum_{b_2=a_2}^{a_3} \cdots \sum_{b_{n-1}=a_{n-1}}^{a_n} W(b_1, b_2, \ldots, b_{n-1})$$
(5)

for all n, where \sum^* is the modified summation operator

$$\sum_{i=r}^{s} f(i) = \frac{1}{2}f(r) + f(r+1) + f(r+2) + \dots + f(s-1) + \frac{1}{2}f(s);$$

the number of factors of $\frac{1}{2}$ that contribute to the coefficient of $W(b_1, b_2, \ldots, b_{n-1})$ is exactly the number of b_i 's that also occur among the a_i 's. Observe that the operator \sum^* resembles definite integration in that

$$\sum_{i=r}^{s} f + \sum_{i=s}^{t} f = \sum_{i=r}^{t} f$$
(6)

for r < s < t. Indeed, if we extend \sum^* by defining

$$\sum_{i=r}^{r} {}^{*}f = 0$$

for all r and

$$\sum_{i=r}^{s} f = -\sum_{i=s}^{r} f$$

for r > s then (6) holds for all integers r, s, t. Hence, by starting from the base relation $W(a_1) = 1$, (5) can be applied iteratively to define $W(\cdot)$ as a function of a_1, a_2, \ldots, a_n regardless of whether $a_1 < a_2 < \cdots < a_n$ or not.

Notice that if $f(x) = x^m$, then

$$\sum_{r=s}^{t} f(r)$$

is a polynomial in s and t of degree m + 1 of the form

$$\frac{t^{m+1}-s^{m+1}}{m+1}$$
 + terms of lower order.

More generally, if f(x, y, z, ...) is a polynomial in x, y, z, ... with a highest-order term $cx^m y^{m'} z^{m''} ...$, then

$$\sum_{r=s}^{t} f(r, y, z, \ldots)$$

is a polynomial in s, t, y, z, ... of degree (deg f)+1 with highest-order terms

$$\frac{c}{m+1}t^{m+1}y^{m'}z^{m''}\cdots$$
 and $-\frac{c}{m+1}s^{m+1}y^{m'}z^{m''}\cdots$

We will now use (5) and (6) to prove the general formula

$$W(a_1, a_2, \ldots, a_n) = \prod_{1 \le i < j \le n} \frac{a_j - a_i}{j - i}.$$
 (7)

(This immediately yields W(1, 2, ..., n) = 1, which, as we have seen, implies $AD(n) = 2^{n(n+1)/2}$.)

The proof is by induction. When n = 1, we have $W(a_1) = 1$, so that (7) is satisfied. Suppose now that we have

$$W(b_1, b_2, \ldots, b_{n-1}) = \prod_{1 \le i < j \le n-1} \frac{b_j - b_i}{j-i}$$

for all $b_1, b_2, \ldots, b_{n-1}$. Since $W(b_1, b_2, \ldots, b_{n-1})$ is a polynomial of degree (n-1)(n-2)/2 with a highest-order term

$$\frac{b_2^1 b_3^2 \cdots b_{n-1}^{n-2}}{1! 2! \cdots (n-2)!},$$

the recurrence relation (5) and the observations made in the proceeding paragraph imply that $W(a_1, a_2, \ldots, a_n)$ is a polynomial of degree

$$(n-1)(n-2)/2 + (n-1) = n(n-1)/2$$

with a highest-order term

$$\frac{a_2^1 a_3^2 \cdots a_n^{n-1}}{1! 2! \cdots n!}$$

To complete the proof, we need show only that $W(a_1, a_2, \ldots, a_n)$ is skew symmetric in its arguments, for this implies that it is divisible by $(a_2 - a_1)(a_3 - a_1) \cdots (a_n - a_{n-1})$, a polynomial of the same degree (namely, n(n-1)/2) as itself, and a comparison of the coefficients of leading terms yields (7). It suffices to show that interchanging any two consecutive arguments of W changes the sign of the result. For convenience, we illustrate with n = 4:

$$W(a_2, a_1, a_3, a_4) = \sum_{b_1=a_2}^{a_1} \sum_{b_2=a_1}^{a_3} \sum_{b_3=a_3}^{a_4} W(b_1, b_2, b_3)$$

= $\left(-\sum_{b_1=a_1}^{a_2}\right) \left(\sum_{b_2=a_1}^{a_2} \sum_{b_2=a_2}^{a_3}\right) \left(\sum_{b_3=a_3}^{a_4}\right) W(b_1, b_2, b_3).$

The skew symmetry of $W(b_1, b_2, b_3)$ kills off one of the two terms:

$$\sum_{b_1=a_1}^{a_2} \sum_{b_2=a_1}^{a_2} \sum_{b_3=a_3}^{a_4} W(b_1, b_2, b_3) = \sum_{b_2=a_1}^{a_2} \sum_{b_1=a_1}^{a_2} \sum_{b_3=a_3}^{a_4} W(b_2, b_1, b_3)$$

by re-labeling
$$= \sum_{b_1=a_1}^{a_2} \sum_{b_2=a_1}^{a_2} \sum_{b_3=a_3}^{a_4} W(b_2, b_1, b_3)$$

by commutativity
$$= -\sum_{b_1=a_1}^{a_2} \sum_{b_2=a_1}^{a_2} \sum_{b_3=a_3}^{a_4} W(b_1, b_2, b_3)$$

by skew symmetry,

implying that the term vanishes. Hence,

$$W(a_2, a_1, a_3, a_4) = -\sum_{b_1=a_1}^{a_2} \sum_{b_2=a_2}^{a_3} \sum_{b_3=a_3}^{a_4} W(b_1, b_2, b_3)$$

= -W(a_1, a_2, a_3, a_4),

as claimed. Similarly, $W(a_1, a_2, a_4, a_3) = -W(a_1, a_2, a_3, a_4)$. A slightly more complicated calculation, involving a sum of four terms of which three vanish, gives $W(a_1, a_3, a_2, a_4) = -W(a_1, a_2, a_3, a_4)$. The argument for the skew symmetry of $W(a_1, a_2, \ldots, a_n)$ is much the same for n in general, although the notation is more complex; we omit the details.

Having shown that W is skew symmetric in its arguments, we have completed the proof of (7), which yields the formula for AD(n) as a consequence.

We have given our own proof of (7) in order to keep this article self-contained, but we should mention that the formula is equivalent to Theorem 2 in [13] and is a special case of the main result of [22]. It can also be derived from identity (5.11) of [11, p. 120] by setting t = -1 and making the observation immediate from [11, p. 104] that $P_{\lambda}(x_1, \ldots, x_n; -1) = s_{\mu}(x_1, \ldots, x_n) \prod_{i < j} (x_i + x_j)$ whenever λ , μ are partitions satisfying $\lambda = \mu + (n - 1, n - 2, \ldots, 0)$. We thank one of the referees for bringing some of these connections to our attention.

128

Some further remarks are in order. First, it is noteworthy that

$$\prod_{1 \le i < j \le n} \frac{a_j - a_i}{j - i}$$

is an integer provided a_1, \ldots, a_n are integers; this can be proved in a messy but straightforward manner by showing that every prime p must divide the numerator at least as many times as it divides the denominator. Alternatively, one can show that this product is equal to the determinant of the *n*-by-*n* matrix whose *i*, *j*th entry is the integer

$$\left(\begin{array}{c}a_i\\j-1\end{array}\right)$$

(see [15] and [18]).

Second, formula (7) has a continuous analogue: If we take V(x) = 1 for all real x and inductively define

$$V(x_1, x_2, \ldots, x_n) = \int_{x_1}^{x_2} \int_{x_2}^{x_3} \cdots \int_{x_{n-1}}^{x_n} V(y_1, y_2, \ldots, y_{n-1}) \, dy_{n-1} \cdots \, dy_2 \, dy_1,$$

then essentially the same argument shows that

$$V(x_1, x_2, \ldots, x_n) = \prod_{1 \leq i < j \leq n} \frac{x_j - x_i}{j - i}$$

This has the following probabilistic interpretation: Given n real numbers $x_1 < x_2 < \cdots < x_n$, let $X_{i,i} = x_i$ for $1 \le i \le n$, and for all $1 \le i < j \le n$ let $X_{i,j}$ be a number chosen uniformly at random in the interval $[x_i, x_j]$. Then the probability that $X_{i,j} \le X_{i+1,j}$ and $X_{i,j} \le X_{i,j+1}$ for all suitable i, j is

$$\prod_{1 \le i < j \le n} \frac{1}{j-i} = \frac{1}{1! 2! \cdots (n-1)!}.$$

We do not know a more direct proof of this fact than the one outlined here.

Third, the usual (unstarred) summation operator does not satisfy a relation such as (6), so the method used here will not suffice to count unweighted monotone triangles. (Mills, Robbins, and Rumsey [13] offer abundant evidence that the number of complete monotone triangles of size n is

$$\prod_{k=0}^{n-1} \frac{(3k+1)!}{(n+k)!}$$

but no proof has yet been found.) However, the operators

$$\sum_{i=r}^{s} L = \sum_{i=r}^{s-1}$$



Fig. 11. Paths in Aztec diamonds.

and

$$\sum_{i=r}^{s} R = \sum_{i=r+1}^{s}$$

do satisfy an analogue of (6), and one can exploit this to give streamlined proofs of some formulas in the theory of plane partitions; details will appear elsewhere.

Fourth, we should note that the function $W(a_1, a_2, \ldots, a_m)$ has significance for tilings of the Aztec diamond of order n, even outside the case with m = nand $a_i = i$ for $1 \le i \le m$. Suppose $m \le n$ and $a_m \le n$, and let Π be the path in the graph G that starts at (-m, n-m) whose 2j - 1st and 2jth steps head south and east, respectively, if $j \in \{a_1, \ldots, a_n\}$ and otherwise head east and south, respectively, for $1 \le j \le n$, ending at the vertex (n - m, -m); Figure 11(a) shows Π when n = 4, m = 2, $(a_1, a_2) = (2, 3)$. It is not difficult to show that the number of domino tilings of the portion of the Aztec diamond that lies above Π is $2^{m(m+1)/2}W(a_1, a_2, \ldots, a_m)$.

Fifth (and last), we should note that the role played by the matrix A at the beginning of the section (in expressing AD(n) in terms of a weighted sum over complete monotone triangles of size n) could have been played just as well by the matrix B, giving rise to an alternative formula expressing AD(n) as the sum, over all complete monotone triangles of size n + 1, of 2 to the power of the number of entries above the bottom row that do not occur in the succeeding row. However, by dividing by $2^{n(n+1)/2}$, we reduce the claim AD(n) = $2^{n(n+1)/2}$ to the same claim as before (the sum of the weights of all the fractionally weighted complete monotone triangles of any given size is equal to 1). This gives a second significance of $W(\cdots)$ for tilings of the Aztec diamond. Specifically, suppose $m \le n + 1$ and $a_m \le n$, and let Π be the path in the graph G that starts at (-m, n + 1 - m) whose 2j - 1st and 2jth steps head east and south, respectively, if $j \in \{a_1, \ldots, a_n\}$ and otherwise head south and east, respectively, for $1 \le j \le n + 1$, ending at the vertex (n + 1 - m, -m); Figure 11(b) shows

 Π when $n = 4, m = 2, (a_1, a_2) = (2, 3)$. It can be shown that the number of domino tilings of the portion of the Aztec diamond that lies above Π is $2^{m(m-1)/2}W(a_1, a_2, \ldots, a_m)$.

Acknowledgements

Noam Elkies is supported by the Harvard Society of Fellows and by National Science Foundation grant DMS-87-18965. Greg Kuperberg is supported by National Science Foundation Graduate Research Fellowship. Michael Larsen is supported by National Science Foundation grants DMS-8610730 and DMS-8807203. James Propp is supported by National Science Foundation Postdoctoral Research Fellowship.

References

- 1. R.J. Baxter, Exactly Solved Models in Statistical Mechanics, Academic Press, New York, 1982.
- 2. E.F. Beckenbach, ed., Applied Combinatorial Mathematics, John Wiley, New York, 1964.
- 3. J.H. Conway and J.C. Lagarias, "Tilings with polyominoes and combinatorial group theory," J. Combin. Theory Ser. A 53 (1990), 183-208.
- 4. C. Fan and F.Y. Wu, "General lattice model of phase transitions," Phys. Rev. B2 (1970), 723-733.
- 5. J.A. Green, Polynomial Representations of GL_n , Lecture Notes in Mathematics 830, Springer, Berlin, 1980.
- 6. W. Jockusch, "Perfect matchings and perfect squares," Ph.D. thesis, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 1992.
- 7. P.W. Kasteleyn, "The statistics of dimers on a lattice, I: the number of dimer arrangements on a quadratic lattice," *Physica* 27 (1961), 1209-1225.
- 8. P.W. Kasteleyn, "Graph theory and crystal physics," in *Graph Theory and Theoretical Physics*, F. Harary, ed., Academic Press, New York, 1967, pp. 43-110.
- 9. E. Lieb, "Residual entropy of square ice," Phys. Rev. 162 (1967), 162-172.
- 10. L. Lovász, Combinatorial Problems and Exercises, North-Holland, Amsterdam, 1979, prob. 4.29.
- 11. I.G. Macdonald, Symmetric Functions and Hall Polynomials, Oxford University Press, Oxford, 1979.
- 12. W.H. Mills, D.P. Robbins, and H. Rumsey, Jr., "Proof of the Macdonald conjecture," Invent. Math. 66 (1982), 73-87.
- 13. W.H. Mills, D.P. Robbins, and H. Rumsey, Jr., "Alternating sign matrices and descending plane partitions," J. Combin. Theory Ser. A 34 (1983), 340-359.
- 14. J.K. Percus, Combinatorial Methods, Courant Institute, New York, 1969.
- 15. G. Pólya and S. Szegö, Problems and Theorems in Analysis, Vol. II, Springer, New York, 1976, prob. 132, p. 134.
- 16. D.P. Robbins, "The story of 1, 2, 7, 42, 429, 7436, ...," Math. Intelligencer 13 (1991), 12-19.
- 17. D.P. Robbins and H. Rumsey, Jr., "Determinants and alternating sign matrices," Adv. Math. 62 (1986), 169-184.
- A.E. Spencer, "Problem E 2637," Amer. Math. Monthly 84 (1977), 134–135; solution published in 85 (1978), 386–387.
- R. Stanley, "A baker's dozen of conjectures concerning plane partitions," in Combinatoire Énumérative, G. Labelle and P. Leroux, eds., Lecture Notes in Mathematics 1234, Springer-Verlag, Berlin, 1986, pp. 285-293.

- 20. R. Stanley, Enumerative Combinatorics, Vol. I. Wadsworth and Brooks/Cole, Belmont, MA, 1986.
- W. Thurston, "Conway's tiling groups," Amer. Math. Monthly 97 (1990), 757-773.
 T. Tokuyama, "A generating function of strict Gelfand patterns and some formulas on characters of general linear groups," J. Math. Soc. Japan 40 (1988), 671-685.
- 23. B.-Y. Yang, "Three enumeration problems concerning Aztec diamonds," Ph.D. thesis, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 1991.