Journal of Algebraic Combinatorics 1 (1992), 219–234

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

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 continue the study of the family of planar regions dubbed Aztec diamonds in our earlier article and study the ways in which these regions can be tiled by dominoes. Two more proofs of the main formula are given. The first uses the representation theory of $GL(n)$. The second is more combinatorial and produces a generating function that gives not only the number of domino tilings of the Aztec diamond of order n but also information about the orientation of the dominoes (vertical versus horizontal) and the accessibility of one tiling from another by means of local modifications. Lastly, we explore a connection between the combinatorial objects studied in this paper and the square-ice model studied by Lieb.

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

5. Grassmann Algebras

The resemblance between the formula

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

given in Section 4 and the Weyl dimension formula for representations of $GL(n)$ is not coincidental. In fact, the identity

$$\sum_A 2^{N(A)} = 2^{n(n-1)/2}$$

*Part I of this paper appears in Journal of Algebraic Combinatorics, Vol. 1, No. 2, September 1992, 111–132.

can be proved by pure representation theory. The idea is to relate the rules for consecutive rows in Gelfand triangles to the decomposition of $GL(n)$-representations as $GL(n-1) \times GL(1)$ representations.

Let V be a finite-dimensional vector space, let $\Lambda^i(V)$ be the ith exterior power of V, and let

$$\Lambda(V) = \bigoplus_{i=0}^{n} \Lambda^i(V)$$

be the exterior algebra generated by V (regarded here as a module, not an algebra). It is elementary that if V and W are finite-dimensional vector spaces,

$$\Lambda^2(V \oplus W) = \Lambda^2(V) \oplus \Lambda^2(W) \oplus (V \otimes W),$$

$$\Lambda(V \oplus W) = \Lambda(V) \otimes \Lambda(W).$$

Writing $G(V)$ for $\Lambda(\Lambda^2(V))$, we get $\dim(G(V)) = 2^{\binom{\dim(V)}{2}}$ and

$$G(V \oplus W) = G(V) \otimes G(W) \otimes \Lambda(V \otimes W).$$

We now recall the Cartan–Weyl theory of weights of irreducible representations of Lie groups in the case of $GL(n)$ (due to Schur); for more details, see [5]. If $V = \mathbf{C}^n$, then $GL(V) = GL(n, \mathbf{C})$ contains the group T of diagonal matrices $\mathrm{diag}(x_1, \dots, x_n)$. The analytic homomorphisms $T \to \mathbf{C}^*$ are precisely the Laurent monomials $x_1^{a_1} \cdots x_n^{a_n}$, $a_i \in \mathbf{Z}$. If ρ is a finite-dimensional (analytic) representation of $GL(n, \mathbf{C})$, its restriction to T is a direct sum of l-dimensional analytic representations (called *weights*) and the restriction of the trace of ρ to T is a Laurent polynomial in the x_i; we represent a weight of ρ by the sequence of exponents occurring in the corresponding Laurent monomial in $\mathrm{tr}(\rho|_T)$. For instance, the trace function of the identity representation is the sum of the diagonal elements, $x_1 + \cdots + x_n$, so the weights are the basis vectors $(0, \dots, 0, 1, 0, \dots, 0)$. The operations of linear algebra can be translated into operations on trace polynomials. Thus the trace of a direct sum of representations is the sum of the traces, the trace of a tensor product is the product of traces, the trace of the kth exterior power is the kth elementary symmetric function of the constituent monomials, and so on. The irreducible representations ρ of $GL(n, \mathbf{C})$ are indexed by *dominant weights* $\lambda = (\lambda_1, \dots, \lambda_n)$, where $\lambda_i \in \mathbf{Z}$ and $\lambda_1 \leq \cdots \leq \lambda_n$; among all weights that occur in $\rho|_T$ and satisfy this inequality, λ has the greatest norm. For instance, the dominant weight of the identity representation is $(0, 0, \dots, 0, 1)$.

We set $a_i = \lambda_i + i$, so the finite-dimensional irreducible representations of $GL(n, \mathbf{C})$ are indexed by $(a_1, \dots, a_n)$, where $a_i \in \mathbf{Z}$ and $a_1 < \cdots < a_n$. The Weyl character formula for $GL(n)$ says that the trace function for $\rho(a_1, \dots, a_n)$ is

$$\frac{\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) x_1^{a_{\sigma(1)}} \cdots x_n^{a_{\sigma(n)}}}{\prod_{1\le i\le n} x_i \prod_{1\le j<i\le n} (x_i - x_j)}.$$

The numerator of this expression can be written

$$\det \begin{pmatrix} x_1^{a_1} & x_2^{a_1} & \cdots & x_n^{a_1} \\ x_1^{a_2} & x_2^{a_2} & \cdots & x_n^{a_2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{a_n} & x_2^{a_n} & \cdots & x_n^{a_n} \end{pmatrix}.$$

Subtracting $x_1^{a_{i+1}-a_i}$ times row i from row $i+1$, for $i = n-1, n-2, \ldots, 1$, we obtain

$$x_1^{a_1} \sum_{\tau\in S_{n-1}} \operatorname{sgn}(\tau) \prod_{i=1}^{n-1} \left(x_{i+1}^{a_{\tau(i)+1}} - x_{i+1}^{a_{\tau(i)}} x_1^{a_{\tau(i)+1}-a_{\tau(i)}} \right).$$

The trace functions for the representations V, $\Lambda^2(V)$, and $G(V)$ (that is, for the action of $GL(n)$ on these spaces induced by the action of $GL(n)$ on V) are given by $\Sigma_{1\le i\le n} x_i$, $\Sigma_{1\le j<i\le n} x_i x_j$, and $\Sigma_{1\le j<i\le n}(1 + x_i x_j)$, respectively. The trace function for $\rho(a_1, \ldots, a_n) \otimes G(V)$ is therefore

$$x_1^{a_1} \sum_{\tau\in S_{n-1}} \operatorname{sgn}(\tau) \prod_{i=1}^{n-1} \left(x_{i+1}^{a_{\tau(i)+1}} - x_{i+1}^{a_{\tau(i)}} x_1^{a_{\tau(i)+1}-a_{\tau(i)}} \right) \prod_{i=1}^{n} \frac{1}{x_i} \prod_{1\le j<i\le n} \frac{1 + x_i x_j}{x_i - x_j};$$

this is equal to

$$x_1^{a_1-1} \sum_{\tau\in S_{n-1}} \operatorname{sgn}(\tau) \prod_{i=1}^{n-1} (\Sigma_1(i, \tau) + \Sigma_2(i, \tau)) \prod_{i=2}^{n} \frac{1}{x_i} \prod_{2\le j<i\le n} \frac{1 + x_i x_j}{x_i - x_j},$$

where

$$\Sigma_1(i, \tau) = \sum_{k=a_{\tau(i)}}^{a_{\tau(i)+1}-1} x_{i+1}^{k} x_1^{a_{\tau(i)+1}-1-k},$$

$$\Sigma_2(i, \tau) = \sum_{l=a_{\tau(i)}+1}^{a_{\tau(i)+1}} x_{i+1}^{l} x_1^{a_{\tau(i)+1}-1-l}.$$

Viewing $GL(n-1, \mathbf{C})$ as the subgroup of $GL(n, \mathbf{C})$ consisting of all matrices of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix},$$

we can restrict $\rho(a_1, \ldots, a_n) \otimes G(\mathbf{C}^n)$ to $GL(n-1)$. At the level of traces on the diagonal, this amounts to setting $x_1 = 1$ to obtain

$$\left(2 \sum_{b_1=a_1}^{a_2}{}^{*}\right) \cdots \left(2 \sum_{b_{n-1}=a_{n-1}}^{a_n}{}^{*}\right) \sum_{\tau \in S_{n-1}} \operatorname{sgn}(\tau) x_2^{b_{\tau(1)}} \cdots x_n^{b_{\tau(n-1)}} \prod_{i=2}^{n} \frac{1}{x_i} \prod_{2 \le j < i \le n} \frac{1 + x_i x_j}{x_i - x_n}.$$

(As in Section 4, the notation Σ^* indicates a sum where the endpoints are counted with multiplicity $\frac{1}{2}$.) This is visibly the sum of the traces of the $GL(n-1)$-representations

$$\rho(b_1, \ldots, b_{n-1}) \otimes G(\mathbf{C}^{n-1})$$

counted with appropriate multiplicities. In fact, since two representations of $GL(n-1)$ are the same if and only if their trace polynomials coincide, this gives a formula for the restriction of the representation $\rho(a_1, \ldots, a_n) \otimes G(\mathbf{C}^n)$ to $GL(n-1)$. Iterating this process, we see that $W(1, 2, \ldots, n)$ (as defined in Section 4) is the value obtained by substituting $x_1 = x_2 = \cdots = x_n = 1$ in the trace function of $G(\mathbf{C}^n)$ viewed as a $GL(n)$-representation or, in other words, the trace function of $G(\mathbf{C}^n)$ on $GL(0) = 1$, which is simply the dimension of $G(\mathbf{C}^n)$.

6. Domino Shuffling

The even (or standard) coloring of the Aztec diamond, as defined earlier, is the black–white checkerboard coloring in which the interior squares along the northeast border are black. In this section it will be convenient to also consider the other checkerboard coloring, which we call *odd*. We will continue to call a vertex of a checkerboard-colored region *even* if it is the upper-left corner of a white square, and we will call it *odd* otherwise, only now this notion depends on the checkerboard coloring chosen as well as on the coordinates of the vertex.

In general, a union of squares in a bicolored checkerboard will be called *even* if the leftmost square in its top row is white, and it will be called *odd* if that square is black. Thus the left half of Figure 12 shows an even Aztec diamond, an even 2-by-2 block and two even dominoes (along with an even vertex), and the right half of Figure 12 shows odd objects of the same kind. Hereafter, a 2-by-2 block will be called simply a *block*.

Given a tiling of a colored region by dominoes, we may remove all the odd blocks to obtain an *odd-deficient tiling*. In general, an odd-deficient domino tiling of a region in the plane is a partial tiling that has no odd blocks and that can be extended to a complete tiling of that region by adding only odd blocks. An odd-deficient tiling of the Aztec diamond of order n with its even coloring is uniquely determined by the heights of its even vertices, as recorded in the matrix B of Section 3; thus these odd-deficient tilings are in one–one correspondence with alternating-sign matrices of order $n + 1$.

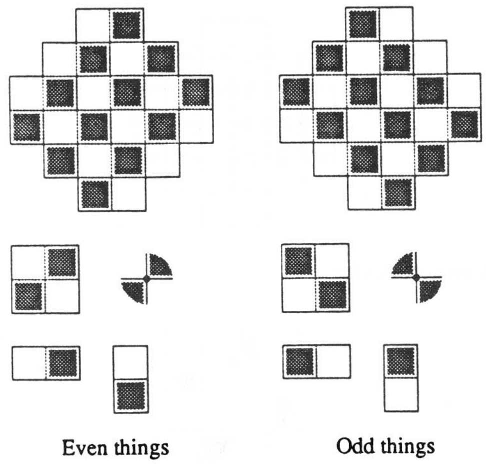


Figure 12. Even things and odd things.

Given a partial tiling $\mathcal{T}$ of the plane, let $U_{\mathcal{T}}$ be the union of the dominoes belonging to $\mathcal{T}$. Observe that if $\mathcal{T}$ is odd-deficient, then the boundary of $U_{\mathcal{T}}$ has corners only at odd vertices.

The functions $v(T)$ and $r(T)$ defined earlier can be expressed in the form

$$v(T) = \sum_{d \in T} v(d)$$

and

$$r(T) = \sum_{d \in T} r(d)$$

for suitable functions $v(\cdot)$ and $r(\cdot)$ on the set of dominoes, which we now define. If the domino d is horizontal, let $v(d) = r(d) = 0$; if d is vertical, let $v(d) = \frac{1}{2}$ and let $r(d)$ be assigned according to the location of the center of d following the pattern set down in Figure 13 for the case $n = 3$. (More formally, we may declare that if d is the vertical domino with upper-left corner at (i, j), then $r(d) = (-1)^{i+j+n}(i + n + 1)$.) Clearly, $v(T)$ is the sum of $v(d)$ over all dominoes $d \in T$. As for $r(T)$, note that

$$r(T_{\min}) = 0 = \sum_{d \in T_{\min}} r(d);$$

also note that a move that increases $h(T)$ by 1 either creates two vertical dominoes d_1, d_2 that satisfy $r(d_1) + r(d_2) = 1$ or annihilates two vertical dominoes d_1, d_2 that satisfy $r(d_1) + r(d_2) = -1$. Thus by induction $r(T) = \Sigma_{d \in T} r(d)$ for all tilings T.

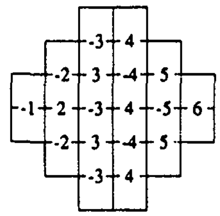


Figure 13. r-weights of vertical dominoes.

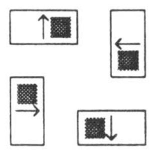


Figure 14. Directions for shuffling dominoes.

We therefore have

$$\mathrm{AD}(n; x, q) = \sum_T \prod_{d \in T} x^{v(d)} q^{r(d)}.$$

We now prove

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

by using a process called *domino shuffling*, which is a certain involution on the set of odd-deficient tilings of an infinite checkerboard. If d is a domino on a colored region, we define $S(d)$, the *shuffle* of d, as the domino obtained by moving d one unit to the left or up if it is even and one unit to the right or down if it is odd (see Figure 14). Graphically, one can place an arrow joining the two noncorner vertices on the boundary of d, pointing from the even vertex to the odd vertex; this indicates the direction in which d will shuffle.

Clearly, S is an involution on the set of dominoes on an infinite checkerboard. Two dominoes form an odd block if and only if each is the shuffle of the other; if d and $S(d)$ are horizontal, then $r(d) + r(S(d)) = 0$ whereas if d and $S(d)$ are vertical, then $r(d) + r(S(d)) = -1$.

Given a partial tiling $\mathcal{T}$ we define $S(\mathcal{T})$, the *shuffle* of $\mathcal{T}$, to be the collection of all $S(d)$ with $d \in \mathcal{T}$.

LEMMA. *Domino shuffling is an involution on the odd-deficient tilings of an infinite checkerboard.*

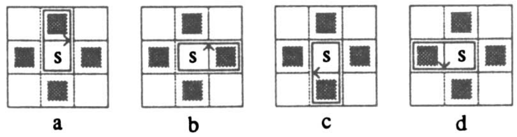


Figure 15. Four domino positions.

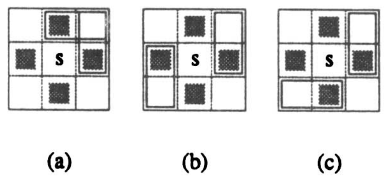


Figure 16. Three pairs of impossible dominoes.

Proof. Let $\tilde{T}$ be an odd-deficient tiling of the plane, with T an extension to a true tiling of the plane. We first show that $S(\tilde{T})$ is a partial tiling, that is, that no two dominoes of $S(\tilde{T})$ overlap. Assume otherwise, and suppose that a white square s is covered by two dominoes in $S(\tilde{T})$. That is, $S(\tilde{T})$ contains two of the four dominoes a, b, c, d shown in Figure 15 (with arrows indicating the directions in which they shuffle). There are six cases to be considered and ruled out:

(i) $a, b \in S(\tilde{T})$: $\tilde{T}$ must contain the dominoes $S^{-1}(a) = S(a)$ and $S^{-1}(b) = S(b)$, but $S(a)$ and $S(b)$ overlap (see Figure 16(a)).
(ii) $c, d \in S(\tilde{T})$: Same reasoning as in case (i).
(iii) $a, c \in S(\tilde{T})$: $S(a), S(c) \in \tilde{T}$ (see Figure 16(b)). The full tiling T must cover s but cannot include b or d (since T already includes $S(a)$ and $S(c)$, which conflict with those two dominoes); hence T must include a or c. However, in the former case $a \in T$ forms an odd block with $S(a) \in T$, so that $S(a) \notin \tilde{T}$ after all; the case $c \in T$ leads to a similar contradiction.
(iv) $b, d \in S(\tilde{T})$: Same reasoning as in case (iii).
(v) $a, d \in S(\tilde{T})$: Same reasoning as in cases (iii) and (iv), though the geometry is somewhat different (see Figure 16(c)).
(vi) $b, c \in S(\tilde{T})$: Same reasoning as in case (v).

Hence a white square cannot be covered by two dominoes of $S(\tilde{T})$. The proof for black squares is similar. Therefore, $S(\tilde{T})$ is a partial tiling of the checkerboard.

We must also show that $S(\tilde{T})$ is odd deficient. $S(\tilde{T})$ cannot contain any odd blocks because the inverse shuffle (which is the same as the shuffle) of an odd block is an odd block. It remains to show that the boundary of $U_{S(\mathcal{T})}$ has corners only at odd vertices. Let v be an even vertex. It is easily checked that v is a corner of $U_{\mathcal{T}}$ if and only if $U_{\mathcal{T}}$ contains unequal numbers of black squares and white squares adjacent to v (and similarly for $U_{S(\mathcal{T})}$). A domino $d \in \mathcal{T}$ may cover, of the four squares adjacent to v, one black square, one white square, or one square of each color. In these three cases $S(d)$ covers one white square, one black square, or no squares at all, respectively. Thus the even vertex v could be a corner of $U_{S(\mathcal{T})}$ only if it was already a corner of $U_{\mathcal{T}}$. However, we assumed $\mathcal{T}$ was odd deficient, so that its only corners were at odd vertices. □

Assume now that $\mathcal{T}$ is an odd-deficient tiling, not of the entire plane, but of the order-$(n-1)$ Aztec diamond. We can use the above to show that $S(\mathcal{T})$ is an odd-deficient tiling of the order-n diamond. It is clear that for every domino $d \in \mathcal{T}$, $S(d)$ lies in the order-n diamond; what is less pictorially obvious is that the complement of $S(\mathcal{T})$ relative to the order-n diamond must be a union of odd blocks. One way to see this is to tile the complement of the order-$(n-1)$ Aztec diamond with horizontal dominoes and thereby obtain an odd-deficient tiling $\mathcal{T}^+$ of the entire plane. Then, by the lemma, $S(\mathcal{T}^+)$ is an odd-deficient tiling of the plane; some of its missing odd blocks lie in two semi-infinite strips of height 2 to the left and right of the order-n diamond, and all the others must lie strictly inside the order-n diamond. None of these blocks crosses the boundary of the order-n diamond, so if we add these blocks to $S(\mathcal{T})$, we get a complete tiling of the order-n diamond.

Consider now an odd-deficient tiling $\mathcal{T}$ of the order-$(n-1)$ Aztec diamond, with $\mathcal{T}_{\text{vert}}$ equal to the set of vertical tiles of $\mathcal{T}$; let

$$v(\mathcal{T}) = \sum_{d \in \mathcal{T}} v(d) = \sum_{d \in \mathcal{T}_{\text{vert}}} v(d)$$

and

$$r(\mathcal{T}) = \sum_{d \in \mathcal{T}} r(d) = \sum_{d \in \mathcal{T}_{\text{vert}}} r(d)$$

(recall that $v(d) = r(d) = 0$ for all horizontal dominoes d). Let

$$\mathrm{AD}(n-1, \mathcal{T}; x, q) = \sum x^{v(T)} q^{r(T)},$$

where the sum is over all tilings T that extend $\mathcal{T}$; we have

$$\mathrm{AD}(n-1; x, q) = \sum_{\mathcal{T}} \mathrm{AD}(n-1, \mathcal{T}; x, q),$$

where the sum is over all partial tilings $\mathcal{T}$ of the order-$(n-1)$ Aztec diamond. Say that $\mathcal{T}$ is missing m odd blocks, so that it gives rise to 2^m distinct complete tilings T; then it is easily seen that

$$\mathrm{AD}(n-1, \bar{T}; x, q) = (1 + xq^{-1})^m \prod_{d \in \bar{T}} x^{v(d)} q^{r(d)}. \tag{8}$$

$S(\bar{T})$ is an odd-deficient tiling of the order-n Aztec diamond with its odd coloring, missing $m + n$ odd blocks. Therefore, relative to the even coloring, we have

$$\mathrm{AD}(n, S(\bar{T}); x, q) = (1 + xq)^{m+n} \prod_{d \in \bar{T}_{\mathrm{vert}}} x^{v(S(d))} q^{-r(S(d))}.$$

The product in the right-hand side of the above equation can be rewritten as

$$\begin{aligned} \prod_{d \in \bar{T}_{\mathrm{vert}}} x^{v(S(d))} q^{-r(S(d))} &= \prod_{d \in \bar{T}_{\mathrm{vert}}} x^{v(d)} q^{r(d)+1} \\ &= \prod_{d \in \bar{T}_{\mathrm{vert}}} (xq^2)^{v(d)} q^{r(d)} \\ &= \prod_{d \in \bar{T}} (xq^2)^{v(d)} q^{r(d)}. \end{aligned}$$

However, substitution of n for $n - 1$ and xq^2 for x in (8) yields

$$\mathrm{AD}(n, \bar{T}; xq^2, q) = (1 + xq)^m \prod_{d \in \bar{T}} (xq^2)^{v(d)} q^{r(d)}.$$

Hence

$$\mathrm{AD}(n, S(\bar{T}); x, q) = (1 + xq)^n \mathrm{AD}(n - 1, \bar{T}; xq^2, q).$$

Since every odd-deficient tiling of the order-n Aztec diamond with odd coloring is of the form $S(\bar{T})$ for some odd-deficient tiling of the order-$(n - 1)$ Aztec diamond with even coloring, we can sum both sides of the preceding equation over all $\bar{T}$, obtaining

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

The general formula for $\mathrm{AD}(n; x, q)$ follows immediately by induction.

Although this proof made no mention of alternating-sign matrices, they are very much involved in determining the exact locations of the various 2-by-2 blocks. Specifically, let T be a domino tiling of the Aztec diamond of order $n - 1$, and let A be the $(n - 1)$-by-$(n - 1)$ alternating-sign matrix determined by T as in Section 3. Then the locations of the odd blocks in $\bar{T}$ are given by the 1's in A, and the locations of the odd blocks in $S(\bar{T})$ are given by the -1's.

Latent within the proof of the formula for $\mathrm{AD}(n; x, q)$ is an iterative bijection between domino tilings of the order-n Aztec diamond and bit strings of length $n(n + 1)/2$. Say we are given a bit string of length $1 + 2 + \cdots + n$, and suppose we have already used the first $1 + 2 + \cdots + (k - 1)$ bits to construct a domino tiling of the order-$(k-1)$ diamond. Impose the even coloring on this Aztec diamond, and locate the odd blocks, of which there are m. Pick up these odd blocks in some definite order (of which we will say more shortly), and put them elsewhere, retaining their order. Shuffle the dominoes in the remaining partial tiling of the Aztec diamond of order $k-1$. The resulting partial tiling of the order-k Aztec diamond has $m+k$ holes in it; fill these holes (again in some definite order) with the m blocks that were removed before, followed by k other blocks whose orientations (horizontal versus vertical) are determined by the next k bits of the bit string. In this way one obtains a complete tiling of the Aztec diamond of order k. Note that no information has been lost; the procedure is fully reversible. Thus iteration of the process gives a bijection between bit strings of length $n(n+1)/2$ and domino tilings of the order-n Aztec diamond. Moreover, every 0 (respectively, 1) in the bit string leads to the creation of two horizontal (respectively, vertical) dominoes in the tiling, so it is immediate that the number of tilings of the Aztec diamond with $2v$ vertical dominoes is

$$\binom{n(n+1)/2}{v}.$$

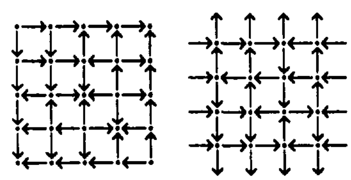
The preceding construction requires a pairing between the m missing odd blocks of an odd-deficient tiling of the order-$(k-1)$ Aztec diamond and m of the $m+k$ missing odd blocks of an odd-deficient tiling of the order-n Aztec diamond. There is a canonical way of doing this pairing. Recall that these two kinds of blocks correspond to the -1's and $+1$'s in an alternating-sign matrix A, so it suffices to decree some sort of pairing between the -1's and a subset of the $+1$'s (which will leave n $+1$'s left over). However, this is easy: simply pair each -1 with the next $+1$ below it in its column. In terms of shuffling, this means that the odd blocks of $\tilde{T}$ drift southeast until they find a hole in $S(\tilde{T})$ into which they can fit; this leaves n holes near the upper-left border of the order-n Aztec diamond, which the n new 2-by-2 blocks exactly fill.

It would be nice to have a "shuffling" proof of the product formula for W. It would also be nice to have a procedure for randomly generating monotone triangles according to the (uneven) probability distribution given by the weights $W(\cdot)$.

7. Square Ice

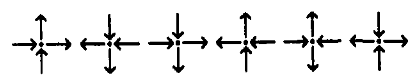
It is worthwhile to point out a connection between the combinatorial objects investigated in this paper and a statistical-mechanical model that has been studied extensively since the 1960's. Recall that an n-by-n alternating-sign matrix can be represented by its skewed summation, as in Figure 17(a). Replace each entry in the matrix by a node, and put a directed edge between every two adjacent entries, pointing from the smaller to the larger. Then one has a directed graph

0	1	2	3	4
1	2	1	2	3
2	1	2	1	2
3	2	1	0	1
4	3	2	1	0



(a) (b) (c)

Figure 17. Alternating-sign matrices to square ice.



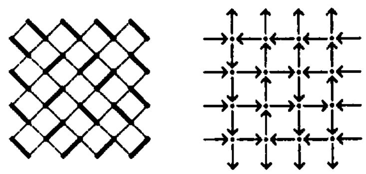
ω_1 ω_2 ω_3 ω_4 ω_5 ω_6

Figure 18. Labeling the vertex configurations.

in which the circulation around every square cell is 0 (that is, each cell has two clockwise edges and two counterclockwise edges); see Figure 17(b). Finally, rotate each of these edges 90° counterclockwise about its midpoint. The result is a configuration like the one shown in Figure 17(c), with divergence 0 at each node (that is, each node has two incoming arrows and two outgoing arrows). This is exactly the square-ice model of statistical mechanics, with the special boundary condition of incoming arrows along the left and right sides and outgoing arrows along the top and bottom. (For discussion of this and related models, see [1] and [14].)

In the general square-ice model, one associates a Boltzmann weight ω_i (i = 1 to 6) with each of the six possible vertex configurations shown in Figure 18; then the weight of a configuration is defined as $\omega_1^{k_1}\omega_2^{k_2}\cdots\omega_6^{k_6}$, where k_i is the number of vertices in the lattice of type i and the *partition function* associated with the model (denoted by Z) is the sum of the weights of all possible configurations. Z has an implicit dependence on the lattice size and the boundary conditions. It is customary to impose periodic boundary conditions, but we instead impose the "in-at-the-sides, out-at-the-top-and-bottom" condition on our n-by-n grid. Call this the *Aztec boundary condition.*

To recast our work on domino tilings of the Aztec diamond in terms of square ice, it is convenient to rephrase domino-tilings as dimer arrangements, or 1-factors. Specifically, we define a graph G' whose vertices correspond to the cells of the order-n Aztec diamond with an edge between two vertices of G' if and only if the corresponding cells are adjacent. Then a domino tiling of the



(a) (b)

Figure 19. A 1-factor and an ice state.

Aztec diamond corresponds to a 1-factor F of G' (a collection of disjoint edges covering all vertices).

There is a general method for writing the number of 1-factors of a planar graph as a Pfaffian [8]. Indeed, if one assigns weight $w(e)$ to each edge e of a planar graph on N vertices and defines the weight of a 1-factor as the product of its constituent weights, then the sum of the weights of all 1-factors of the graph is equal to the Pfaffian of an antisymmetric N-by-N matrix whose (i, j)th entry is $\pm w(e)$ if the graph has an edge e between i and j and 0 otherwise. (The delicate point is the correct choice of signs.) This method has been applied to the problem of counting 1-factors of m-by-n grids (equivalently, domino tiling of m-by-n rectangles); see [7], [2], [10]. The Pfaffian method provides yet another route to our result on tilings of the Aztec diamond, though we have omitted the calculation here; see [23].

It is convenient to rotate the graph G' 45° clockwise, as in Figure 19(a). Call a cell of G' even or odd according to the parity of the corresponding vertex of G (under the standard coloring), so that the four extreme cells of G' are even. Every even cell is bounded by four edges, of which two, one, or none may be present in any particular 1-factor; the seven possibilities appear at the top of Figure 20, where a bold marking indicates the presence of an edge. If we replace each even cell by the corresponding ice junction given at the bottom of Figure 20, it is easy to check that the result is a valid ice configuration satisfying our special boundary conditions and that every such configuration arises in this way. The process is exemplified in Figure 19(b). Note that the transformation from 1-factors to ice configurations is not one-to-one; it is, in fact, 2^{k_5}-to-one, where k_5 is the number of vertices of type 5 in the ice pattern. That this transformation is equivalent to the more roundabout operation of converting the 1-factor to a domino tiling can be checked by using the heights of the even vertices to form an n-by-n alternating-sign matrix and then turning the matrix into an ice pattern as in the first paragraph of this section.

Let T be a tiling of the Aztec diamond, and let F be the associated 1-factor of G'. Note that every domino in T corresponds to an edge in F and that

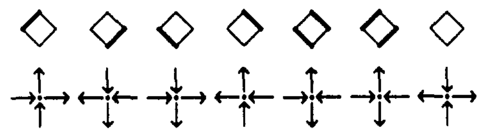


Figure 20. Correspondence between 1-factors and ice states.

this edge belongs to a unique even cell of G'. Hence if we assign the weights x, x, 1, 1, 1, x^2, and 1 to the respective cell figures, the product of the weights of the cell figures appearing in F is equal to $x^{2v(T)}$. Thus if we set

$$\begin{aligned}
\omega_1 &= x, \\
\omega_2 &= x, \\
\omega_3 &= 1, \\
\omega_4 &= 1, \\
\omega_5 &= 1 + x^2, \\
\omega_6 &= 1,
\end{aligned}$$

then the partition function Z coincides with the generating function

$$\mathrm{AD}(n; x^2) = (1 + x^2)^{n(n+1)/2}.$$

Note that $k_5 - k_6 = n$ for all order-n ice configurations with Aztec boundary condition (corresponding to the fact that the number of 1's in an n-by-n alternating-sign matrix must be n more than the number of -1's). Hence replacing ω_5 and ω_6 by $\sqrt{1 + x^2}$ merely divides the partition function by $(1 + x^2)^{n/2}$. Furthermore, $k_1 + k_2 + \cdots + k_6 = n^2$, so multiplying all the Boltzmann weights by a factor b merely multiplies the partition function by b^{n^2}. Writing $a = bx$ and $c = b\sqrt{x^2 + 1} = \sqrt{a^2 + b^2}$, we see (after an easy calculation) that for the square-ice model with Aztec boundary condition and with Boltzmann weights

$$\begin{aligned}
\omega_1 &= a, \\
\omega_2 &= a, \\
\omega_3 &= b, \\
\omega_4 &= b, \\
\omega_5 &= c, \\
\omega_6 &= c
\end{aligned}$$

satisfying $a^2 + b^2 = c^2$ the partition function is given by $Z = c^{n^2}$.

It should be noted that this family of special cases of the ice model (given by a, b, c satisfying $a^2 + b^2 = c^2$) is also the family that corresponds to the free-fermion case and is precisely the case in which the model has been solved by the method of Pfaffians. (See [1, pp. 151, 270–271] and [4]). This leads us to suspect that domino shuffling may, in fact, arise from some combinatorial interpretation of the Pfaffian solution.

We must emphasize the role played by the Aztec boundary conditions in the foregoing analysis, since it adds an element essentially foreign to the physical significance of the ice model. In particular, Lieb's solution of the ice model in the case $\omega_1 = \omega_2 = \cdots = \omega_6 = 1$ [9] tells us that there are asymptotically

$$\left(\sqrt{64/27}\right)^{n^2}$$

order-n ice configurations with periodic boundary conditions; on the other hand, if the conjecture of Mills *et al.* [12], [13] is correct, the number of order-n ice configurations with Aztec boundary conditions should asymptotically be only

$$\left(\sqrt{27/16}\right)^{n^2}.$$

Clearly, there are more constraints on a domino tiling near the boundary of an Aztec diamond than there are near the middle; this accounts for at least some of the drop in entropy. It would be interesting to know in a more quantitative way how the entropy of a random tiling is spatially distributed throughout a large Aztec diamond.

8. Epilogue

There have been many combinatorial transformations in this article, so it may be useful to review them. First, we have

(i) tilings,
(ii) height functions associated with tilings, and
(iii) the order ideals associated with those height functions.

We saw how to go from (i) to (ii) (Thurston's marking scheme), from (ii) to (iii) (see the construction of the poset P in Section 3), and from (iii) back to (i) (the stacked cubes).

Then we have

(iv) alternating-sign matrices,
(v) height functions associated with alternating-sign matrices,
(vi) the order ideals associated with those height functions,
(vii) monotone triangles, and
(viii) states of the square-ice model (or, equivalently, its dual).

We saw the correspondence between (iv) and (v) and between (v) and (vi) in Section 3, between (iv) and (vii) in Section 4, and between (v) and (viii) in Section 7. Further correspondences can be made. For instance, to get from (iv) to (viii) directly, in a given alternating-sign matrix we replace a 1 by a vertex configuration of type 5, we replace a -1 by a vertex configuration of type 6, and we replace each 0 by the unique vertex configuration of type 1, 2, 3, or 4 which fits in the pattern (note that arrows "go straight through" configurations of types 1–4 without reversing).

Then there are the mappings between (i)–(iii) and (iv)–(viii) under the correspondence between domino tilings and compatible pairs of alternating-sign matrices. We saw in Section 3 how to pass between (ii) and (v) and between (iii) and (vi). Other connections can be made, and the reader might find it instructive to try to establish them.

There are actually even more incarnations of alternating-sign matrices than have been discussed here: 3-colorings of certain graphs (subject to boundary constraints), 2-factors of some related graphs, and tilings of various regions in the plane by shapes of two kinds. These other structures may be discussed in a future paper. Then there are other combinatorial objects which seem (but have not been proved) to be equinumerous with the alternating-sign matrices, namely, descending plane partitions and self-complementary totally symmetric plane partitions. See [16] for details.

$\mathrm{AD}(n)$ is a perfect square when $n(n+1)/2$ (one-fourth of the number of vertices) is even, and it is twice a perfect square when $n(n+1)/2$ is odd. More generally, Jockusch [6] has shown that if G is any bipartite graph in the plane with no crossing edges, such that G has 4-fold rotational symmetry and rotation by 90^o exchanges the two parts of G, then the number of matchings of G is a square or twice a square, according to whether the number of vertices of G (necessarily divisible by 4) is divisible by 8 or is not.

Stanley (private communication) has discovered that our two-variable generating function for tilings of the order-n Aztec diamond is actually a specialization of a $2n$-variable generating function. A proof of this identity by means of the shuffling method of Section 6 is described in [23].

Acknowledgments

We wish to thank Carol Sandstrom and Chris Small for suggesting some of the terminology used in this article. Noam Elkies was supported by the Harvard Society of Fellows and by National Science Foundation grant DMS 8718965. Greg Kuperberg was supported by a National Science Foundation Graduate Research Fellowship. Michael Larsen was supported by National Science Foundation grants DMS 8610730 and DMS 8807203. James Propp was supported by a National Science Foundation Postdoctoral Research Fellowship. This is the second part of a two-part article; for terminology and notation see N. Elkies, G. Kuperberg,

M. Larsen, and J. Propp, "Alternating-sign matrices and domino tilings (part I)," *J. Alg. Combin.* **1** (1992), 111–132.

References

1. R.J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, San Diego, CA, 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 A* **53** (1990), 183–208.
4. C. Fan and F.Y. Wu, General lattice model of phase transitions, *Phys. Rev. B* **2** (1970), 723–733.
5. J.A. Green, *Polynomial Representations of GL_n*, Springer Lecture Notes in Mathematics, Vol. 830, Springer-Verlag, Berlin, 1980.
6. W. Jockusch, "Perfect matchings and perfect squares," *J. Combin. Theory A*, to appear.
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 Phsyics*, F. Harary, ed., Academic Press, San Diego, CA, 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, problem 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 A* **34** (1983), 340–359.
14. J.K. Percus, *Combinatorial Methods*, Courant Institute of Mathematical Sciences, 1969.
15. G. Pólya and S. Szegö, *Problems and Theorems in Analysis, Vol. II*, Springer, New York, 1976, p. 134, problem 132.
16. D.P. Robbins, "The story of 1, 2, 7, 42, 429, 7436, ...," *Math. Intelligencer* **13**(2) (1991), 12–19.
17. D.P. Robbins and H. Rumsey, Jr., "Determinants and alternating sign matrices," *Adv. Math.* **62** (1986), 169–184.
18. A.E. Spencer, Problem E 2637, *Amer. Math. Monthly* **84** (1977), 134–135; solution published in **85** (1978), 386–387.
19. R. Stanley, *Enumerative Combinatorics, Vol. I.* Brooks-Cole, Belmont, MA, 1986.
20. R. Stanley, "A baker's dozen of conjectures concerning plane partitions," in *Combinatoire Énumérative*, Springer Lecture Notes in Mathematics, Vol. 1234, (G. Labelle and P. Leroux, eds.), Springer-Verlag, Berlin, 1986, pp. 285–293.
21. W. Thurston, "Conway's tiling groups," *Amer. Math. Monthly* **97** (1990), 757–773.
22. 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," doctoral thesis, Massachusetts Institute of Technology, Cambridge, MA, submitted.