Another proof of the alternating sign matrix conjecture

Greg Kuperberg

December 10, 1995

Abstract

Mills, Robbins, and Rumsey [8] conjectured, and Zeilberger [13] recently proved, that there are $\frac{1!4!7!\ldots(3n-2)!}{n!(n+1)!\ldots(2n-1)!}$ alternating sign matrices of order n. We give a new proof of this result using an analysis of the six-vertex state model (also called square ice) based on the Yang-Baxter equation.

Mills, Robbins, and Rumsey [8] conjectured that:

Theorem 1 (Zeilberger) *There are*

$$A(n) = \frac{1!4!7!\ldots(3n-2)!}{n!(n+1)!(n+2)!\ldots(2n-1)!}$$

$n \times n$ alternating sign matrices.

Here, an *alternating sign matrix* or *ASM* is a matrix of 0's, 1's, and -1's such that the non-zero elements in each row and column alternate between 1 and -1 and begin and end with 1, for example:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Alternating sign matrices are related to a number of other combinatorial objects that, remarkably, are also enumerated or conjectured to be enumerated by ratios of progressions of factorials or staggered factorials [9, 11].

Zeilberger [13] recently proved Theorem 1 by establishing that ASM's are equinumerous with totally symmetric, self-complementary plane partitions, which were enumerated by Andrews [1]. In this paper, we present a new proof. The most interesting part of the proof is due to Izergin and Korepin [5, 7], who follow Baxter's remarkable use of the Yang-Baxter equation [2].

If x is a number, define the x-enumeration $A(n;x)$ of $n \times n$ ASM's as their total weight, where the weight of an individual matrix is x^k if it has k entries equal to -1. A variation of the proof establishes another conjecture of Mills, Robbins, and Rumsey:

Theorem 2 *ASM's are 3-enumerated by*

$$A(2n+1;3) = \left(3^{n(n+1)/2}\frac{2!5!8!\ldots(3n-1)!}{(n+1)!(n+2)!\ldots(2n)!}\right)^2$$

$$A(2n;3) = 3^{n-1}\frac{(3n-1)!(n-1)!}{(2n-1)!^2}A(2n-1;3)$$

A second variation establishes the well-known 2-enumeration of ASM's [4, 8]:

$$A(n;2) = 2^{n(n-1)/2}.$$

Finally, the following result, also conjectured by Mills, Robbins, and Rumsey, follows easily from the general method:

Theorem 3 *For each n, there exists a polynomial $B(n;x)$ such that*

$$A(n;x) = B(n;x)B(n+1;x)$$

for n odd and

$$A(n;x) = 2B(n;x)B(n+1;x)$$

for n even.

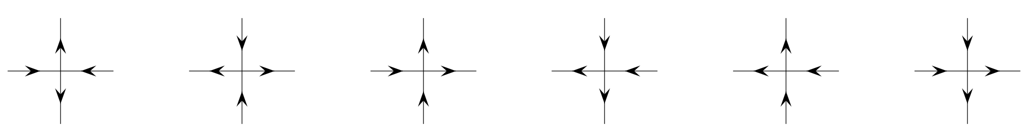
Mills, Robbins, and Rumsey further conjectured that for n odd, $B(n;x)$ is the x-enumeration of vertically symmetric ASM's (where the weight is x^k if there are k ones to the left of the middle column), but this relation remains open.

0.1 Acknowledgements

The author would like to thank David Robbins and Doron Zeilberger for explaining the significance of the alternating sign matrix conjecture, Igor Frenkel and Vaughan Jones for explaining the Yang-Baxter equation, and Jim Propp for introducing the author to enumerative combinatorics. Also, the computer algebra package Maple [3] was indispensible for computations, and the TEX macro package PSTricks [12] was very useful for setting the figures.

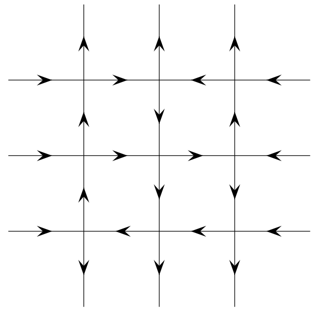
1 State sums

The *six-vertex model* in general refers to the multiplicative weighted enumeration of orientations of a tetravalent planar graph G (called *states*) such that at each vertex, two arrows go in and two go out. Number the six allowed orientations incident to a given vertex (called states of a vertex) 1 through 6:

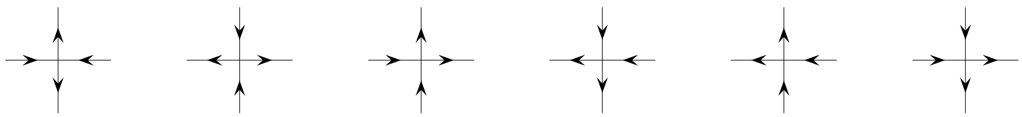


1 2 3 4 5 6

State i at vertex v is given a weight $w(i,v)$. The weight of a state of G is the product of the weights of its vertices, and the *state sum* is the total weight of all states. The six-vertex model may also be considered with boundary conditions, meaning that there may be univalent vertices whose edges have fixed orientations. In particular, consider a six-vertex state of an $n \times n$ square grid with edges pointing inward at the sides and outward at the top and bottom:



The six-vertex model on a square grid is also called *square ice*. A square ice state can be converted to an ASM by the correspondence

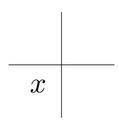


$1 \qquad -1 \qquad 0 \qquad 0 \qquad 0 \qquad 0$

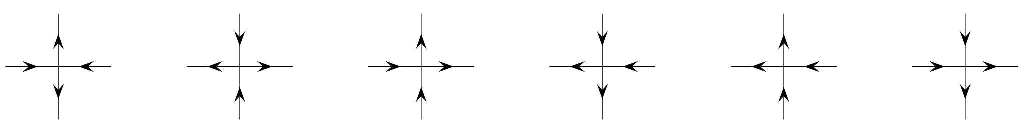
This conversion is bijective [4, 10]. Thus, the enumeration of ASM's is equivalent to a six-vertex state sum in which all weights are 1.

Let h be a complex number or an indeterminate, let q^x denote e^{hx}, and let $[x]$ denote $\frac{q^{x/2}-q^{-x/2}}{q^{1/2}-q^{-1/2}}$. We will consider various *half-integral Laurent polynomials*, meaning polynomials with integral or half-integral exponents of either sign such that the difference between any two exponents is an integer. For example, if q is fixed, $[x]$ is a half-integral Laurent polynomial in q^x. Given two such polynomials $P(t)$ and $Q(t)$ over a ring A, we will say that Q divides P if $P(t)/Q(t) \in A[t^{1/2}, t^{-1/2}]$. For example, t divides 1.

A vertex labelled by x:



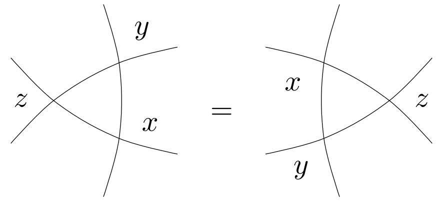
denotes the six weights:



$$-q^{-x/2} \qquad -q^{x/2} \qquad [x-1] \qquad [x-1] \qquad [x] \qquad [x] \tag{1}$$

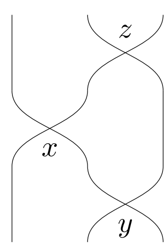
(Since the weights are invariant under rotation by 180 degres, but not 90 degrees, the meaning of a vertex depends on which pair of kitty-corner quadrants contains its label.) Such a vertex is called an R-matrix and is also denoted as $R(x)$.

Theorem 4 (Baxter) *If $x = y + z$, the R-matrices $R(x)$, $R(y)$, and $R(z)$ satisfy the equation*

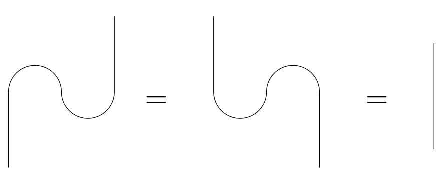


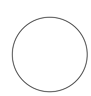
This remarkable identity is known as the *star-triangle* relation or the *Yang-Baxter* equation[2]. Specifically, $R(x)$ is said to parameterize the *trigonometric solutions* to the Yang-Baxter equation. Before proving it, we discuss exactly what the equation means. Each of the two graphs in the equation has six external edges, meaning edges with a univalent vertex. For each external edge on the left, there is a corresponding external edge on the right whose univalent vertex is in the same position; for example, on both sides there is a lowest univalent endpoint, and the two edges with this endpoint correspond to each other. For each of the 64 orientations of the external edges on the left, one can form a state sum Z by summing over admissible orientations of the three internal edges, and one can consider the same orientation on the right and form another state sum Z'. The equation then says that the $Z = Z'$ in all 64 cases. In order for the state sum to be non-zero, three edges must point in and three must point out, so the identity is trivial in 44 of the 64 cases. Note further that the equation simply says that the left side is invariant under rotation by 180 degrees, so the 20 non-trivial numerical identities reduce to 10 identities repeated twice. The argument that follows uses other tricks to further reduce the number of numerical identities to one which can checked be checked easily:

Proof: We first rearrange the left side of the Yang-Baxter equation:



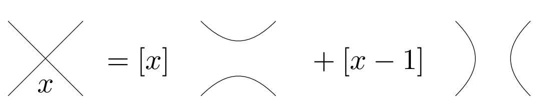
Consider the following augmentation of the six-vertex model: Suppose that a graph has a curved edge with a horizontal tangent at a point p and which is concave down at p. If the edge is oriented to the left in some six-vertex state, p is assigned a multiplicative weight of $-q^{1/2}$, but if it points to the right, it is assigned a multiplicative weight of 1. Contrariwise, if the tangent is horizontal but the curve is concave up, p has weight $-q^{-1/2}$ when the edge points to the left and weight 1 when it points to the right. With this convention, the following simple identities hold:





$$= -q^{1/2} - q^{-1/2} = -[2] \qquad (2)$$

Moreover, $R(x)$ can be expressed as



Thus, a six-vertex state sum involving R-matrices can be expanded as a sum of curves in a calculus in which each closed loop contributes a factor of $-[2]$. (This calculus is called the *Temperley-Lieb category* and is closely related to the quantum group $U_q(\mathrm{sl}(2))$ [6].) The calculus is invariant under isotopy of curves by equation (2). The left side of the Yang-Baxter equation then expands to eight terms, which may collected into five terms corresponding to the five crossingless matchings of six points on a circle. Three of the matchings are invariant under rotation by 180 degrees. The coefficients of the other two are

$$[z-1][x][y-1]$$

and

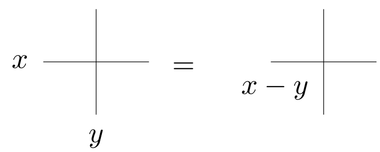
$$[z][x-1][y-1]+[z][x][y]+[z-1][x-1][y]-[2][z][x-1][y].$$

These two quantities are rendered equal by the identities $x=y+z$, $[-a]=-[a]$, and

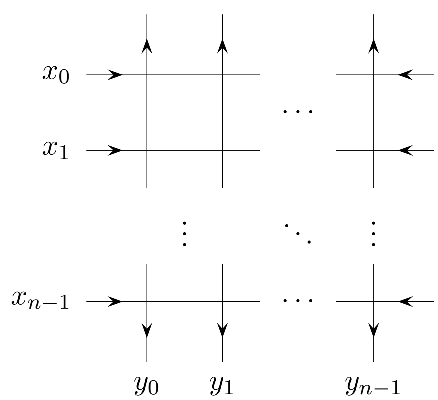
$$[a][b]-[a+1][b-1]=[a-b+1].$$

Thus, the left side is invariant under rotation by 180 degrees. $\square$

As a final notational convenience, define



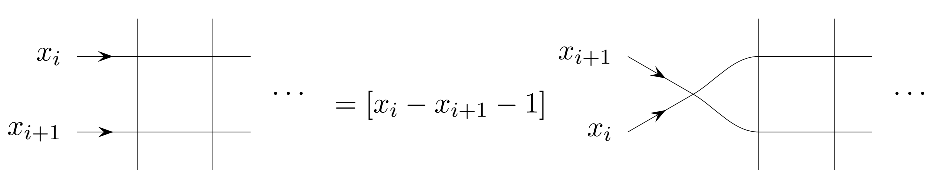
when the lines rather than the vertices of a tetravalent graph are labelled. Following Izergin and Korepin [5, 7], consider $n\times n$ square ice with arbitrary parameters $X=x_0,\ldots,x_{n-1}$ and $Y=y_0,\ldots,y_{n-1}$ for the horizontal and vertical lines:



Let $Z(n;X,Y)$ be the resulting state sum.

Lemma 5 (Baxter) *The function $Z(n;X,Y)$ is symmetric in the x_i's and in the y_i's.*

PROOF: Consider the ith and $i+1$st horizontal lines. An extra vertex (implicitly labelled by x_i-x_{i+1}) may be introduced on the left at the expense of a generically non-zero multiplicative factor:



This relation holds because in an allowed state, all four edges of the new vertex must point to the right. By the Yang-Baxter equation, the vertex can be moved from the left side to the right, whereupon it can be removed, which recovers the multiplicative factor. This operation switches the labels x_i and x_{i+1}. Therefore $Z(n;X,Y)$ is symmetric in x_i and x_{i+1} for each i, which renders it symmetric in all x_i's. The same argument applies to the y_i's. □

Lemma 6 *If $x_i = y_j + 1$, then*

$$Z(n;X,Y) = -q^{-1/2}\left(\prod_{k\neq i}[x_i - y_k]\right)\left(\prod_{k\neq j}[x_k - y_j]\right) Z(n-1; X \setminus x_i, Y \setminus y_j).$$

Proof: Assume first that $i = j = 0$. By Figure (1), the upper left vertex must have state 1 in a non-zero state of the grid. This forces the rest of the top row to have state 5 and the rest of the left column to have state 6, which yields the given multiplicative factor. (In terms of ASM's, only those matrices with a 1 in the top left corner contribute.) The remainder of the grid is an $n-1 \times n-1$ square ice state.

The general case follows from Lemma 5. □

Lemma 7 *The quantity $q^{nx_0/2}Z(n;X,Y)$ is a polynomial in q^{x_0} of degree at most $n-1$.*

Proof: If we multiply all weights of vertices in the first row by $q^{x_0/2}$, then q^{x_0} appears linearly in those weights in which it appears at all. Therefore the modified state sum

$$Z'(n) = q^{nx_0/2}Z(n;X,Y)$$

is a polynomial in q^{x_0}. The first row is the only row in which x_0 appears. In this row, there must be one vertex in state 1, whose modified weight does not involve x_0, and $n-1$ vertices in state 5 or 3. (In terms of ASM's, there must be a 1 in the top row.) Therefore $Z'(n)$ has degree at most $n-1$. □

Theorem 8 (Izergin,Korepin) *The state sum $Z(n;X,Y)$ is given by*

$$Z(n;X,Y) = \frac{(-1)^n\left(\prod_{i=0}^{n-1} q^{(y_i-x_i)/2}\right)\prod_{0\leq i,j<n}[x_i-y_j][x_i-y_j-1]}{\left(\prod_{0\leq j<i<n}[x_i-x_j]\right)\left(\prod_{0\leq i<j<n}[y_i-y_j]\right)}\det M,$$

where

$$M_{i,j} = \frac{1}{[x_i-y_j][x_i-y_j-1]}.$$

Proof: Lemmas 6 and 7, together with $Z(0) = 1$, inductively determine $Z(n)$ by Lagrange interpolation. It is routine to check that the right side satisfies Lemma 6. To check that it also satisfies Lemma 7, Let P be the numerator, let Q be the denominator, let D be the determinant, and let D' be a term in the expansion of the determinant. The product PD is a half-integral Laurent polynomial because PD' is for any choice of D'. Moreover, Q divides PD, because D is antisymmetric in the x_i's and in the y_j's and therefore in the q^{x_i}'s and in the q^{y_j}'s. Thus, $\frac{P}{Q}D$ is a half-integral Laurent polynomial polynomial in q^{x_0}. Finally, the leading term (expanded as a Laurent polynomial in q^{x_0}) of any PD' has exponent $(2n-3)/2$, while the trailing term has exponent $(1-2n)/2$. Therefore the same is true of PD, and $\frac{P}{Q}D$ has leading exponent at most $(n-2)/2$ and trailing exponent at least $-n/2$. In conclusion, $q^{nx_0/2}\frac{P}{Q}D$ is a polynomial in q^{x_0} and has degree at most $n-1$. □

2 Determinants

Consider the state-sum value

$$Z_{\frac{1}{2}}(n) = Z(n; \frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2}, 0, 0, \ldots, 0).$$

In any $n \times n$ square ice state, there are n more vertices in state 1 than state 2, equal numbers in states 3 and 4, and equal numbers in states 5 and 6. Since the weights of these states in $R(2)$ are $-q^{-1/4}$, $-q^{1/4}$, $-[\frac{1}{2}]$, $-[\frac{1}{2}]$, $[\frac{1}{2}]$, and $[\frac{1}{2}]$, respectively, it follows that

$$A(n; x) = [\frac{1}{2}]^{n-n^2}(-1)^n q^{n/4} Z_{\frac{1}{2}}(n), \tag{3}$$

where $x = 1/[\frac{1}{2}]^2 = [2] + 2$. Unfortunately, the determinant in Theorem 8 is singular for $Z_{1/2}(n)$. Therefore, we will instead evaluate

$$Z_{\frac{1}{2}}(n;\epsilon) = Z(n; \frac{1}{2} + \epsilon, \frac{1}{2} + 2\epsilon, \ldots, \frac{1}{2} + n\epsilon, 0, -\epsilon, -2\epsilon \ldots, (1-n)\epsilon)$$

when $h = \frac{4\pi\sqrt{-1}}{3}$, which implies that $x = 1$.

Let $s = q^\epsilon$. Firstly,

$$[k\epsilon + \frac{1}{2}][k\epsilon - \frac{1}{2}] = \frac{s^k + 1 + s^{-k}}{-3}$$

and

$$[k\epsilon] = \frac{s^{k/2} - s^{-k/2}}{\sqrt{-3}}.$$

The matrix M of Theorem 8 becomes

$$M_{i,j} = \frac{-3}{s^{i+j+1} + 1 + s^{-(i+j+1)}}.$$

The state sum becomes

$$Z_{\frac{1}{2}}(n;\epsilon) = \frac{q^{-n/4} s^{-n^2/2} 3^{-n(n+1)/2} \prod_{0 \le i,j < n}(s^{i+j+1} + 1 + s^{-(i+j+1)})}{\prod_{0 \le j < i < n}(s^{(i-j)/2} - s^{(j-i)/2})^2} \det M.$$

The determinant of M can be computed using the following two lemmas, for which we extend the bracket notation by defining $[x]_t = \frac{t^{x/2} - t^{-x/2}}{t - t^{-1}}$ for any t.

Lemma 9 (Cauchy) *Let $X = x_0, \ldots, x_{n-1}$ and $Y = y_0, \ldots, y_{n-1}$ be variables, and let*

$$T(n; X, Y)_{i,j} = \frac{1}{[x_i - y_j]_t}$$

for $0 \le i, j < n$. Then

$$\det T(n, k; t) = \frac{\left(\prod_{0 \le j < i < n}[x_i - x_j]_t\right)\left(\prod_{0 \le i < j < n}[y_i - y_j]_t\right)}{\prod_{0 \le i,j < n}[x_i - y_j]_t}.$$

PROOF: Let D be the determinant, let P be the denominator, and let Q be the numerator. Then the arguments of Theorem 8 apply, but with the conclusion that $\frac{P}{Q}D$ is a degree 0 polynomial in all variables, *i.e.*, a constant. Let D' be the diagonal term in the determinant; D' is the only term such that PD' is not divisible by any $[x_i - y_i]_t$. All factors of $\frac{P}{Q}D'$ cancel at the specialization $x_i = y_i$ and all other terms of $\frac{P}{Q}D$ vanish; therefore $\frac{P}{Q}D = 1$. □

Lemma 9 can also be proved by induction using Dodgson's condensation method [9, 10].

Let $T(n) = T(n; 1, 2, \ldots, n, 0, -1, -2, \ldots, 1-n)$. Then

$$T(n)_{i,j} = \frac{1}{[i+j+1]_t}$$

and $\det T(n)$ is given by Lemma 9.

Lemma 10 *Let*

$$S(n; s, t)_{i,j} = \frac{s^{(i+j+1)/2} - s^{-(i+j+1)/2}}{t^{(i+j+1)/2} - t^{-(i+j+1)/2}}$$

for $0 \le i, j < n$. *Then*

$$\det S(n; s, t) = \frac{(-1)^{n(n-1)/2}}{(t^{1/2} - t{-}1/2)^n} \left(\prod_{0 \le j < i < n} [i-j]_t^2 \right) \left(\prod_{0 \le i,j < n} \frac{s^{1/2} t^{(i-j)/2} - s^{1/2} t^{(j-i)/2}}{[i+j+1]_t} \right).$$

PROOF: The quantity $s^{n^2/2}(\det S(n; s, t)$ has degree n^2 as a polynomial in s. Moreover, for $0 \le k < n$,

$$S(n; t^k, t)_{i,j} = [k]_{t^{i+j+1}} = \sum_{\ell=0}^{k-1} A(t^{\ell - (k-1)/2})_{i,j},$$

where the matrix $A(z)$ given by

$$A(z)_{i,j} = z^{i+j+1}$$

has rank 1. Thus, the rank of $S(n; t^k, t)$ is at most k, and it follows that $(s - t^k)^{n-k}$ divides $\det S(n; s, t)$. Similarly, $(s - t^{-k})^{n-k}$ divides the determinant for $1 \le k < n$. These divisibilities determine $\det S(n; s, t)$ up to a factor which is a function of t. The leading coefficient is then $(\det T(n))/(t^{1/2} - t^{-1/2})^n$. □

The determinant of $-M/3 = S(n; s, s^3)$ is given by Lemma 10. Collecting factors yields

$$\begin{aligned} Z_{1/2}(n; \epsilon) &= \frac{q^{-n/4} s^{-n^2/2} 3^{-n(n+1)/2} \left(\prod_{0 \le i,j < n} [3(i+j+1)]_s \right)}{\left(\prod_{0 \le i,j < n} [i+j+1]_s \right) \left(\prod_{0 \le j < i < n} [i-j]_s^2 \right)} \\ &\quad \frac{(-3)^n (-1)^{n(n-1)/2}}{[3]_s^n} \left(\prod_{0 \le j < i < n} \frac{[3(i-j)]_s^2}{[3]_s^2} \right) \left(\prod_{0 \le i,j < n} \frac{[3]_s [3(i-j)+1]_s}{[3(i+j+1)]_s} \right) \\ &= q^{-n/4} s^{-n^2/2} (-1)^n \left(\prod_{0 \le j < i < n} \frac{[3(i-j)]_s}{3[i-j]_s} \right) \left(\prod_{i=0}^{n-1} \frac{\prod_{j=1}^{3i+1} [j]_s}{\prod_{j=1}^{n+i} [j]_s} \right) \end{aligned}$$

Note that the second factor is Andrews' q-enumeration of descending plane partitions [9] with q replaced by s. Taking the limit as $\epsilon \to 0$ and combining with equation (3), the factors of q and -1 cancel, the factors of s become factors of 1, and the brackets disappear. The result is

$$\frac{1!4!7!\ldots(3n-2)!}{n!(n+1)!(n+2)!\ldots(2n-1)!.}$$

This completes the proof of Theorem 1.

For general x, the matrix M becomes

$$M_{i,j} = \frac{x^2 - 4x}{s^{i+j+1} + 2 - x + s^{-(i+j+1)}}.$$

There are two other values of x when the denominator is a cyclotomic (Laurent) polynomial in some power of s, namely $x = 2$ and $x = 3$. In the former case, $-M/4 = S(n; s^2, s^4)$, whose determinant is given by Lemma 10; alternatively, the determinant may also be derived from Lemma 9. In the latter case, $-M/3 = S'(n; s, s^3)$, where

$$S'(n; s, t)_{i,j} = \frac{s^{i+j+1} + 1}{t^{i+j+1} + 1}.$$

A variation of Lemma 10 establishes the determinant of $S'(n; s, t)$; the leading coefficient is simply the determinant of $S(n; t, t^2)$. These manipulations clearly lead to product formulas for $A(n; 2)$ and $A(n; 3)$, and in particular, to a proof of Theorem 2. We omit the details of rearranging and cancelling factors to put the product formulas in their standard form.

Finally, we use Theorem 8 to prove Theorem 3. Recall the variables x_i and y_i in the definition of Z, which are not to be confused with the x of $A(n; x)$. If we set $x_i = \frac{1}{2} + f_i\epsilon$ and $y_j = f_j\epsilon$ for some f_i's such that $f_{n-1-i} = -f_i$, then $Z(n; X, Y)$ again converges to $A(n; x)$ up to normalization as $\epsilon \to 0$. In this case, the corresponding matrix M is given by

$$M_{i,j} = \frac{x^2 - 4x}{s^{f_i - f_j} + 2 - x + s^{f_j - f_i}}.$$

This matrix M possesses the symmetry $(i, j) \mapsto (n - 1 - i, n - 1 - j)$, *i.e.*, it commutes with the antidiagonal permutation matrix P. Therefore, a change of basis divides M into blocks corresponding to the eigenspaces of P. Therefore the determinant of M is the product of the determinants of the blocks. This is the origin of the factorization of the $A(n; x)$'s into the $B(n; x)$'s.

References

[1] G. Andrews. Plane partitions V: The T. S. S. C. P. P. conjecture. *J. Combin. Theory Ser. A*, 66:28–39, 1994.

[2] R. J. Baxter. *Exactly solved models in statistical mechanics.* Academic Press, London, 1982.

[3] B. W. Char, K. O. Geddes, G. H. Gonnet, B. L. Leong, M. B. Monagan, and S. M. Watt. *Maple V.* Waterloo Maple Software, Waterloo, Ontario, 1991.

[4] N. Elkies, G. Kuperberg, M. Larsen, and J. Propp. Alternating sign matrices and domino tilings, I. *J. Algebraic Combin.*, 1:111–132, 1992.

[5] A. G. Izergin. Partition function of a six-vertex model in a finite volume. *Soviet Phys. Dokl.*, 32:878–879, 1987.

[6] L. H. Kauffman and S. L. Lins. *Temperley-Lieb recoupling theory and invariants of 3-manifolds.* Annals of Mathematics Studies. Princeton University Press, Princeton, N. J., 1994.

[7] V. E. Korepin, N. M. Bogoliubov, and A. G. Izergin. *Quantum inverse scattering method and correlation functions.* Cambridge University Press, Cambridge, England, 1993.

[8] W. H. Mills, D. P. Robbins, and H. Rumsey. Alternating-sign matrices and descending plane partitions. *J. Combin. Theory Ser. A*, 34:340–359, 1983.

[9] D. P. Robbins. The story of 1,2,7,42,429,7436,.... *Math. Intelligencer*, 13(2):12–19, 1991.

[10] D. P. Robbins and H. Rumsey. Determinants and alternating-sign matrices. *Advances in Math.*, 62:169–184, 1986.

[11] R. P. Stanley. Symmetries of plane partitions. *J. Combin. Theory Ser. A*, 43:103–113, 1986.

[12] T. Van Zandt. Pstricks: Postscript macros for generic TEX. Available at `ftp://ftp.princeton.edu/pub/tvz/`.

[13] D. Zeilberger. Proof of the alternating-sign matrix conjecture. To appear in *Electronic J. Combin.*

[14] D. Zeilberger. Proof of the refined alternating sign matrix conjecture. Preprint.