Schur Polynomials and the Yang-Baxter Equation

Ben Brubaker, Daniel Bump and Solomon Friedberg

Department of Mathematics, MIT, Cambridge MA 02139-4307, USA
Department of Mathematics, Stanford University, Stanford CA 94305-2125, USA
Department of Mathematics, Boston College, Chestnut Hill MA 02467-3806, USA

May 25, 2011

Abstract

We describe a parametrized Yang-Baxter equation with nonabelian parameter group. That is, we show that there is an injective map $g \mapsto R(g)$ from $\mathrm{GL}(2,\mathbb{C}) \times \mathrm{GL}(1,\mathbb{C})$ to $\mathrm{End}(V \otimes V)$ where V is a two-dimensional vector space such that if $g, h \in G$ then $R_{12}(g)R_{13}(gh)\,R_{23}(h) = R_{23}(h)\,R_{13}(gh)R_{12}(g)$. Here R_{ij} denotes R applied to the i, j components of $V \otimes V \otimes V$. The image of this map consists of matrices whose nonzero coefficients a_1, a_2, b_1, b_2, c_1, c_2 are the Boltzmann weights for the non-field-free six-vertex model, constrained to satisfy $a_1a_2 + b_1b_2 - c_1c_2 = 0$. This is the exact center of the disordered regime, and is contained within the free fermionic eight-vertex models of Fan and Wu. As an application, we show that with boundary conditions corresponding to integer partitions λ, the six-vertex model is exactly solvable and equal to a Schur polynomial s_λ times a deformation of the Weyl denominator. This generalizes and gives a new proof of results of Tokuyama and Hamel and King.

Baxter's method of solving lattice models in statistical mechanics is based on the *star-triangle relation*, which is the identity

$$R_{12}S_{13}T_{23} = T_{23}S_{13}R_{12}, \tag{1}$$

where R, S, T are endomorphisms of $V \otimes V$ for some vector space V. Here R_{ij} is the endomorphism of $V \otimes V \otimes V$ in which R is applied to the i-th and j-th copies of V

and the identity map to the k-th component, where i, j, k are $1, 2, 3$ in some order. If the endomorphisms R, S, T are all equal, this is the *Yang-Baxter equation* (cf. [15], [25]).

A related construction is the *parametrized Yang-Baxter equation*

$$R_{12}(g)R_{13}(g \cdot h)R_{23}(h) = R_{23}(h)R_{13}(g \cdot h)\, R_{12}(g) \tag{2}$$

where the endomorphism R now depends on a parameter g in a group G and $g, h \in G$ in (2). There are many such examples in the literature in which the group G is an abelian group such as $\mathbb{R}$ or $\mathbb{R}^\times$. In this paper we present an example of (2) having a *non-abelian* parameter group. The example arises from a two-dimensional lattice model—the six-vertex model.

We now briefly review the connection between lattice models and instances of (1) and (2). In statistical mechanics, one attempts to understand global behavior of a system from local interactions. To this end, one defines the partition function of a model to be the sum of certain locally determined Boltzmann weights over all admissible states of the system. Baxter (see [1] and [2], Chapter 9) recognized that instances of the star-triangle relation allowed one to explicitly determine the partition function of a lattice model.

The six-vertex, or 'ice-type,' model is one such example that is much studied in the literature, and we revisit it in detail in the next section. For the moment, we offer a few general remarks needed to describe our results. In our presentation of the six-vertex model, each state is represented by a labeling of the edges of a finite rectangular lattice by $\pm$ signs, called *spins*. If the Boltzmann weights are invariant under sign reversal the system is called *field-free*, corresponding to the physical assumption of the absence of an external field. For field-free weights, the six-vertex model was solved by Lieb [23] and Sutherland [32], meaning that the partition function can be exactly computed. The papers of Lieb, Sutherland and Baxter assume periodic boundary conditions, but non-periodic boundary conditions were treated by Korepin [18] and Izergin [14]. Much of the literature assumes that the model is field-free. In this case, Baxter shows there is one such parametrized Yang-Baxter equation with parameter group $\mathbb{C}^\times$ for each value of a certain real invariant $\triangle$, defined below in (7) in terms of the Boltzmann weights.

One may ask whether the parameter subgroup $\mathbb{C}^\times$ may be enlarged by including endomorphisms whose associated Boltzmann weights lie outside the field-free case. If $\triangle \neq 0$ the group may *not* be so enlarged. However we will show in Theorem 4 that if $\triangle = 0$, then the group $\mathbb{C}^\times$ may be enlarged to $\mathrm{GL}(2, \mathbb{C}) \times \mathrm{GL}(1, \mathbb{C})$ by expanding the set of endomorphisms to include non-field-free ones. In this *expanded* $\triangle = 0$ *regime*, $R(g)$ is not field-free for general g. It is contained within the set of exactly

solvable eight-vertex models called the *free fermionic model* by Fan and Wu [8], [9]. Our calculations suggest that it is not possible to enlarge the group G to the entire free fermionic domain in the eight vertex model.

As an application of these results, we study the partition function for ice-type models having boundary conditions determined by an integer partition λ. More precisely, we give an explicit evaluation of the partition function for any set of Boltzmann weights chosen so that $\triangle = 0$. This leads to an alternate proof of a deformation of the Weyl character formula for GL_n found by Hamel and King [13], [12]. That result was a substantial generalization of an earlier generating function identity found by Tokuyama [33], expressed in the language of Gelfand-Tsetlin patterns.

Our boundary conditions depend on the choice of a partition λ. Once this choice is made, the states of the model are in bijection with strict Gelfand-Tsetlin patterns having a fixed top row. These are triangular arrays of integers with strictly decreasing rows that interleave (Section 3). In its original form, Tokuyama's formula expresses the partition function of certain ice models as a sum over strict Gelfand-Tsetlin patterns.

This connection between states of the ice model and strict Gelfand-Tsetlin patterns has one historical origin in the literature for alternating sign matrices. (An independent historical origin is in the Bethe Ansatz. See Baxter [2] Chapter 8 and Kirillov and Reshetikhin [17].) The bijection between the set of alternating sign matrices and strict Gelfand-Tsetlin patterns having smallest possible top row is in Mills, Robbins and Rumsey [27], while the connection with what are recognizably states of the six-vertex model is in Robbins and Rumsey [29]. This connection was used by Kuperberg [19] who gave a second proof (after Zeilberger) of the alternating sign matrix conjecture of Mills, Robbins and Rumsey.

It was observed by Okada [28] and Stroganov [31] that the number of $n \times n$ alternating sign matrices, that is, the value of Kuperberg's ice (with particular Boltzmann weights involving cube roots of unity) is a special value of the particular Schur function in $2n$ variables with $\lambda = (n, n, n-1, n-1, \cdots, 1, 1)$ divided by a power of 3. Moreover Stroganov gave a proof using the Yang-Baxter equation. This occurrence of Schur polynomials in the six-vertex model is different from the one we discuss, since Baxter's parameter $\triangle$ is nonzero for those investigations.

There are other works relating symmetric function theory to vertex models or spin chains. Lascoux [21], [20] gave six-vertex model representations of Schubert and Grothendieck polynomials of Lascoux and Schützenberger [22] and related these to the Yang-Baxter equation. Fomin and Kirillov [10], [11] also gave theories of the Schubert and Grothendieck polynomials based on the Yang-Baxter equation. Tsilevich [34] gives an interpretation of Schur polynomials and Hall-Littlewood poly-

nomials in terms of a quantum mechanical system. Jimbo and Miwa [16] give an interpretation of Schur polynomials in terms of two-dimensional fermionic systems. (See also Zinn-Justin [35].)

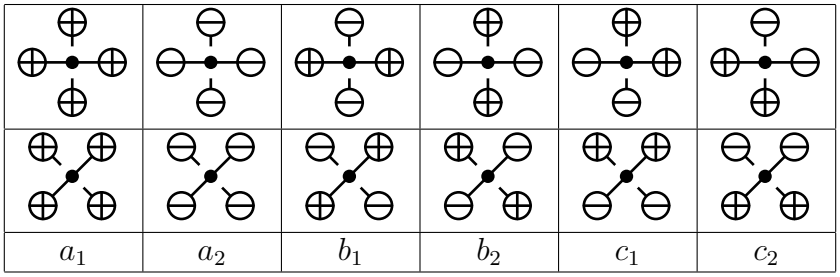
McNamara [26] has clarified that the Lascoux papers are potentially related to ours at least in that the Boltzmann weights [21] belong to the expanded $\triangle = 0$ regime. Moreover, he is able to show based on Lascoux' work how to construct models of factorial Schur functions. Ice models for factorial Schur functions were further investigated by Bump, McNamara and Nakasuji, who found more general constructions.

We are grateful to Gautam Chinta and Tony Licata for stimulating discussions and to the referee for insightful comments. This work was supported by NSF grants DMS-0652609, DMS-0652817, DMS-0652529, DMS-0702438, DMS-1001079, DMS-1001326 and NSA grant H98230-10-1-0183. SAGE [30] was very useful in the preparation of this paper.

1 The six-vertex model

We review the six-vertex model from statistical mechanics. Let us consider a lattice (or sometimes more general graph) in which the edges are labeled with "spins" $\pm$. Each vertex will be assigned a *Boltzmann weight*, which depends on the spins on its adjacent edges.

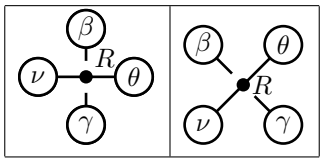
Let us denote the Boltzmann weights as follows:



All remaining Boltzmann weights are taken to be zero; in particular the Boltzmann weight will be zero unless the number of adjacent edges labeled '$-$' is even. We will consider the vertices in two possible orientations, as shown above, and arrange these Boltzmann weights into a matrix as follows:

$$R = \begin{pmatrix} a_1 & & & \\ & b_1 & c_1 & \\ & c_2 & b_2 & \\ & & & a_2 \end{pmatrix} = \begin{pmatrix} a_1(R) & & & \\ & b_1(R) & c_1(R) & \\ & c_2(R) & b_2(R) & \\ & & & a_2(R) \end{pmatrix}. \tag{3}$$

If the edge spins are labeled $\nu, \beta, \gamma, \theta \in \{+, -\}$ as follows:



then we will denote by $R^{\theta\gamma}_{\nu\beta}$ the corresponding Boltzmann weight. Thus $R^{++}_{++} = a_1(R)$, etc. Because we will sometimes use several different systems of Boltzmann weights within a single lattice, we label each vertex with the corresponding matrix from which the weights are taken.

Alternately, R may be thought of as an endomorphism of $V \otimes V$, where V is a two-dimensional vector space with basis v_+ and v_-. Write

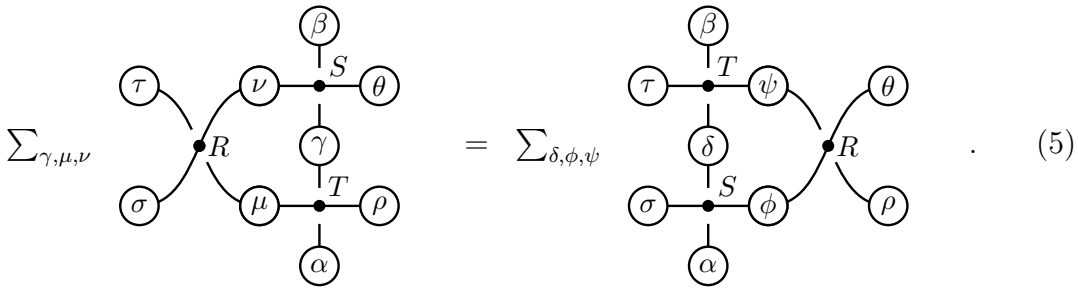
$$R(v_\nu \otimes v_\beta) = \sum_{\theta,\gamma} R^{\theta\gamma}_{\nu\beta}\, v_\theta \otimes v_\gamma. \tag{4}$$

Then the ordering of basis vectors: $v_+ \otimes v_+$, $v_+ \otimes v_-$, $v_- \otimes v_+$, $v_- \otimes v_-$ gives (4) as the matrix (3).

If ϕ is an endomorphism of $V \otimes V$ we will denote by ϕ_{12}, ϕ_{13} and ϕ_{23} the endomorphisms of $V \otimes V \otimes V$ defined as follows. If $\phi = \phi' \otimes \phi''$ where $\phi', \phi'' \in \mathrm{End}(V)$ then $\phi_{12} = \phi' \otimes \phi'' \otimes 1$, $\phi_{13} = \phi' \otimes 1 \otimes \phi''$ and $\phi_{23} = 1 \otimes \phi' \otimes \phi''$. We extend this definition to all ϕ by linearity. Now if ϕ, ψ, χ are three endomorphisms of $V \otimes V$ we define the *Yang-Baxter commutator*

$$[\![\phi, \psi, \chi]\!] = \phi_{12}\psi_{13}\chi_{23} - \chi_{23}\psi_{13}\phi_{12}.$$

Lemma 1. *The vanishing of* $[\![R, S, T]\!]$ *is equivalent to the star-triangle identity*



(5)

for every fixed combination of spins $\sigma, \tau, \alpha, \beta, \rho, \theta$.

The term *star-triangle identity* was used by Baxter. The meaning of equation (5) is as follows. For fixed $\sigma, \tau, \alpha, \beta, \rho, \theta, \mu, \nu, \gamma$, the value or Boltzmann weight of

the left-hand side is by definition the product of the Boltzmann weights at the three vertices, that is, $R^{\nu\mu}_{\sigma\tau}S^{\theta\gamma}_{\nu\beta}T^{\rho\alpha}_{\mu\gamma}$, and similarly for the right-hand side. Hence the meaning of (5) is that for fixed $\sigma, \tau, \alpha, \beta, \rho, \theta$,

$$\sum_{\gamma,\mu,\nu} R^{\nu\mu}_{\sigma\tau}S^{\theta\gamma}_{\nu\beta}T^{\rho\alpha}_{\mu\gamma} = \sum_{\delta,\phi,\psi} T^{\psi\delta}_{\tau\beta}S^{\phi\alpha}_{\sigma\delta}R^{\theta\rho}_{\phi\psi}. \tag{6}$$

It is not hard to see that this is equivalent to the vanishing of $[\![R, S, T]\!]$.

In [2], Chapter 9, Baxter considered conditions for which, given S and T, there exists a matrix R such that $[\![R, S, T]\!] = 0$. We will slightly generalize his analysis. He considered mainly the *field-free case* where $a_1(R) = a_2(R) = a(R)$, $b_1(R) = b_2(R) = b(R)$ and $c_1(R) = c_2(R) = c(R)$. The condition $c_1(R) = c_2(R) = c(R)$ is easily removed, but with no gain in generality. The other two conditions $a_1(R) = a_2(R) = a(R)$, $b_1(R) = b_2(R) = b(R)$ are more serious restrictions.

In the field-free case, let

$$\triangle(R) = \frac{a(R)^2 + b(R)^2 - c(R)^2}{2a(R)\,b(R)}, \qquad a_1(R) = a_2(R) = a(R), \quad \text{etc}\,. \tag{7}$$

Then Baxter showed that given any S and T with $\triangle(S) = \triangle(T)$, there exists an R such that $[\![R, S, T]\!] = 0$.

Generalizing this result to the non-field-free case, we find that there are not one but two parameters

$$\begin{aligned}
\triangle_1(R) &= \frac{a_1(R)a_2(R) + b_1(R)b_2(R) - c_1(R)c_2(R)}{2a_1(R)b_1(R)}, \\
\triangle_2(R) &= \frac{a_1(R)a_2(R) + b_1(R)b_2(R) - c_1(R)c_2(R)}{2a_2(R)b_2(R)}.
\end{aligned}$$

to be considered.

Theorem 2. *Assume that $a_1(S)$, $a_2(S)$, $b_1(S)$, $b_2(S)$, $c_1(S)$, $c_2(S)$, $a_1(T)$, $a_2(T)$, $b_1(T)$, $b_2(T)$, $c_1(T)$ and $c_2(T)$ are nonzero. Then a necessary and sufficient condition for there to exist parameters $a_1(R)$, $a_2(R)$, $b_1(R)$, $b_2(R)$, $c_1(R)$, $c_2(R)$ such that $[\![R, S, T]\!] = 0$ with $c_1(R), c_2(R)$ nonzero is that $\triangle_1(S) = \triangle_1(T)$ and $\triangle_2(S) = \triangle_2(T)$.*

Proof. Suppose that $\triangle_1(S) = \triangle_1(T)$ and $\triangle_2(S) = \triangle_2(T)$. Then we may take

$$\begin{aligned}
a_1(R) &= \frac{b_2(S)a_1(T)b_1(T) - a_1(S)b_1(T)b_2(T) + a_1(S)c_1(T)c_2(T)}{a_1(T)} \\
&= \frac{a_1(S)b_1(S)a_2(T) - a_1(S)a_2(S)b_1(T) + c_1(S)c_2(S)b_1(T)}{b_1(S)},
\end{aligned} \tag{8}$$

$$\begin{aligned} a_2(R) &= \frac{b_1(S)a_2(T)b_2(T) - a_2(S)b_1(T)b_2(T) + a_2(S)c_1(T)c_2(T)}{a_2(T)} \\ &= \frac{a_2(S)b_2(S)a_1(T) - a_1(S)a_2(S)b_2(T) + c_1(S)c_2(S)b_2(T)}{b_2(S)} \end{aligned} \tag{9}$$

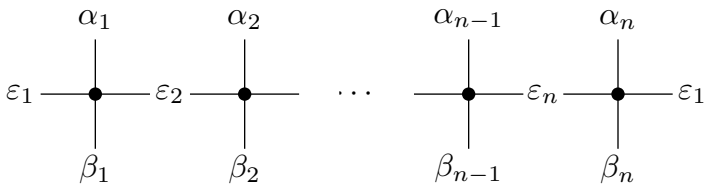
$$b_1(R) = b_1(S)a_2(T) - a_2(S)b_1(T), \qquad b_2(R) = b_2(S)a_1(T) - a_1(S)b_2(T), \tag{10}$$

$$c_1(R) = c_1(S)c_2(T), \qquad c_2(R) = c_2(S)c_1(T). \tag{11}$$

Using $\triangle_1(S) = \triangle_1(T)$ and $\triangle_2(S) = \triangle_2(T)$ it is easy to that the two expressions for $a_1(R)$ agree, and similarly for $a_2(R)$. One may check that $[\![R, S, T]\!] = 0$. On the other hand, it may be checked that the relations required by $[\![R, S, T]\!] = 0$ are contradictory unless $\triangle_1(S) = \triangle_1(T)$ and $\triangle_2(S) = \triangle_2(T)$. □

In the field-free case, these two relations reduce to a single one, $\triangle(S) = \triangle(T)$, and then $\triangle(R)$ has the same value: $\triangle(R) = \triangle(S) = \triangle(T)$.

The equality (5) has important implications for the study of *row-transfer matrices*, one of Baxter's original motivations for introducing the star-triangle relation. Given Boltzmann weights $a_1(R), a_2(R), \cdots$, we associate a $2^n \times 2^n$ matrix $V(R)$. The entries in this matrix are indexed by pairs $\alpha = (\alpha_1, \cdots, \alpha_n)$, $\beta = (\beta_1, \cdots, \beta_n)$, where $\alpha_i, \beta_i \in \{\pm\}$. The coefficient $V(R)_{\alpha,\beta}$ is computed by first calculating the products of the Boltzmann weights of the vertices of the configuration



for each choice of spins $\varepsilon_1, \cdots, \varepsilon_n \in \{\pm\}$, and then summing these products over all possible states (that is, all assignments of the ε_i). Note that in this configuration the right-hand spin is denoted ε_1 rather than ε_{n+1}; that is, the boundary conditions are periodic.

It follows from Baxter's argument that if R can be found such that $[\![R, S, T]\!] = 0$ then $V(S)$ and $V(T)$ commute, and can be simultaneously diagonalized. We will not review Baxter's argument here, but variants of it with non-periodic boundary conditions will appear later in this paper.

In the field-free case when $[\![R,S,T]\!]=0$, $V(R)$ belongs to the same commuting family as $V(S)$ and $V(T)$. This gives a great simplification of the analysis in Chapter 9 of Baxter [2] over the analysis in Chapter 8 using different methods based on the Bethe Ansatz.

In the non-field-free case, however, the situation is different. If $\triangle_1(S)=\triangle_1(T)$ and $\triangle_2(S)=\triangle_2(T)$ then by Theorem 2 there exists R such that $[\![R,S,T]\!]=0$, and so one may use Baxter's method to prove the commutativity of $V(S)$ and $V(T)$. However $\triangle_1(R)$ and $\triangle_2(R)$ are not necessarily the same as $\triangle_1(S)=\triangle_1(T)$ and $\triangle_2(S)=\triangle_2(T)$, respectively, and so $V(R)$ may not commute with $V(S)$ and $V(T)$.

In addition to the field-free case, however, there is *another* case where $V(R)$ necessarily does commute with $V(S)$ and $V(T)$, and it is that case which we turn to next. This is when $a_1a_2+b_1b_2-c_1c_2=0$. The next theorem will show that if the weights of S and T satisfy this condition, then R exists such that $[\![R,S,T]\!]=0$, and moreover the weights of R also satisfy the same condition. Thus not only $V(S)$ and $V(T)$ but also $V(R)$ lie in the same space of commuting transfer matrices.

In this case, with $a_1=a_1(R)$, etc., we define

$$\pi(R)=\pi\begin{pmatrix} a_1 & & & \\ & b_1 & c_1 & \\ & c_2 & b_2 & \\ & & & a_2 \end{pmatrix}=\begin{pmatrix} c_1 & & & \\ & a_1 & b_2 & \\ & -b_1 & a_2 & \\ & & & c_2 \end{pmatrix}. \tag{12}$$

Theorem 3. *Suppose that $c_1(S)\,c_2(S)$ and $c_1(T)\,c_2(T)$ are nonzero and*

$$a_1(S)a_2(S)+b_1(S)b_2(S)-c_1(S)c_2(S)=a_1(T)a_2(T)+b_1(T)b_2(T)-c_1(T)c_2(T)=0. \tag{13}$$

Then the $R\in End(V\otimes V)$ defined by $\pi(R)=\pi(S)\,\pi(T)^{-1}$ satisfies $[\![R,S,T]\!]=0$. Moreover,

$$a_1(R)\,a_2(R)+b_1(R)\,b_2(R)-c_1(R)\,c_2(R)=0. \tag{14}$$

Proof. We will use Theorem 2, where it was assumed that $a_1(S)$, $a_2(S)$, $b_1(S)$, $b_2(S)$, $c_1(S)$, $c_2(S)$, $a_1(T)$, $a_2(T)$, $b_1(T)$, $b_2(T)$, $c_1(T)$ are all nonzero. Now we are only assuming that the c_i are nonzero. It is enough to prove Theorem 3 assuming also that the a_i and b_i are nonzero, so that Theorem 2 applies, since the case where the a_i and b_i are possibly zero will then follow by continuity. The matrix R will not be the matrix in Theorem 2, but will rather be a constant multiple of it. We have

$$\pi(T)^{-1}=\frac{1}{D}\begin{pmatrix} c_2(T) & & & \\ & a_2(T) & -b_2(T) & \\ & b_1(T) & a_1(T) & \\ & & & c_1(T) \end{pmatrix}$$

where $D = a_1(T)a_2(T) + b_1(T)b_2(T) = c_1(T)c_2(T)$. With notation as in Theorem 2, equations (8) and (9) may be rewritten, using (13), as the equations

$$\begin{aligned} a_1(R) &= a_1(S)a_2(T) + b_2(S)b_1(T), \\ a_2(R) &= a_2(S)a_1(T) + b_1(S)b_2(T). \end{aligned}$$

Combined with (10) and (11) these imply that $\pi(R) = \pi(S)\, D\, \pi(T)^{-1}$. We are free to multiply R by a constant without changing the validity of $[\![R, S, T]\!] = 0$, so we divide it by D. □

We started with S and T and produced R such that $[\![R, S, T]\!] = 0$ because this is the construction motivated by Baxter's method of proving that transfer matrices commute. However it is perhaps more elegant to start with R and T and produce S as a function of these. Thus let $\mathcal{R}$ be the set of endomorphisms R of $V \otimes V$ of the form (3) where $a_1a_2 + b_1b_2 = c_1c_2$. Let $\mathcal{R}^*$ be the subset consisting of such R such that $c_1c_2 \neq 0$.

Theorem 4. *There exists a composition law on $\mathcal{R}^*$ such that if $R, T \in \mathcal{R}^*$, and if $S = R \circ T$ is the composition then $[\![R, S, T]\!] = 0$. This composition law is determined by the condition that $\pi(S) = \pi(R)\pi(T)$ where $\pi : \mathcal{R}^* \longrightarrow \mathrm{GL}(4, \mathbb{C})$ is the map (12). Then $\mathcal{R}^*$ is a group, isomorphic to $\mathrm{GL}(2, \mathbb{C}) \times \mathrm{GL}(1, \mathbb{C})$.*

Proof. This is a formal consequence of Theorem 3. □

It is interesting that, in the non-field-free case, the group law occurs when $\triangle_1 = \triangle_2 = 0$. In the application to statistical physics for field-free weights, phase transitions occur when $\triangle = \pm 1$. If $|\triangle| > 1$ the system is "frozen" in the sense that there are correlations between distant vertices. By contrast $-1 < \triangle < 1$ is the disordered range where no such correlations occur, so our group law occurs in the analog of the middle of the disordered range.

2 Boundary conditions and partition functions

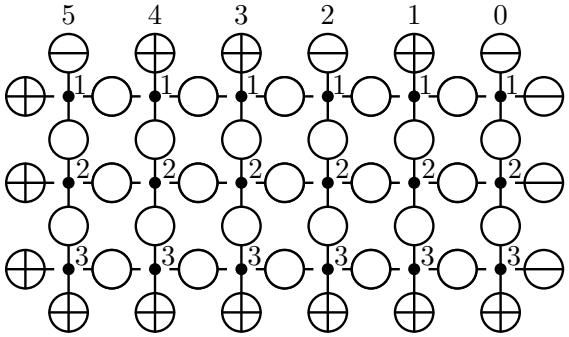
In this section, we describe the global model to be studied using the local Yang-Baxter relation from the previous section.

Let $\lambda = (\lambda_1, \cdots, \lambda_{r+1})$ be a fixed integer partition with $\lambda_{r+1} = 0$ and let $\rho = (r, r-1, \cdots, 0)$. Consider a rectangular lattice with $r+1$ rows and $\lambda_1 + r + 1$ columns. Number the columns of the lattice in descending order from left to right, $\lambda_1 + r$ to 0.

We attach boundary conditions to this lattice according to the choice of $\lambda + \rho$. This amounts to a choice of spin $\pm$ to every edge along the boundary prescribed as follows.

Boundary Conditions Determined by λ. *On the left and bottom boundary edges, assign spin +; on the right edges assign spin −. On the top, assign spin − at every column labeled $\lambda_i + n - i$ $(1 \leqslant i \leqslant n)$, that is, for the columns labeled with values in $\lambda + \rho$; assign spin + at every column not labeled by $\lambda_i + n - i$ for any i.*

For example, suppose that $n = 3$ and $\lambda = (3, 1, 0)$, so that $\lambda + \rho = (5, 2, 0)$. Then the spins on the boundary are as in the following figure.



(15)

The column labels have been written along the top. (The use of the row labelling will be explained shortly.) The choice of $\lambda = (0, \ldots, 0)$ would give all − signs across the top row, and is referred to in the literature as "domain wall boundary conditions."

A *state* of the model will consist of an assignment of spin $\pm$ to each internal edge, pictured above as open circles. The interior spins are not entirely arbitrary, since we require that every vertex "•" in the configuration has adjacent edges whose spins match one of the six admissible configurations in Table 1 under "Square ice" in the table below. The set of all such states with boundary conditions corresponding to λ as above will be called $\mathfrak{S}_\lambda$.

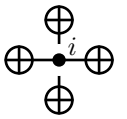
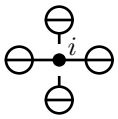
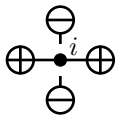
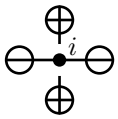
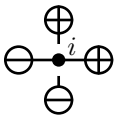
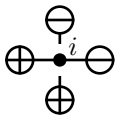
Square Ice						
Boltzmann weight	$a_1^{(i)}$	$a_2^{(i)}$	$b_1^{(i)}$	$b_2^{(i)}$	$c_1^{(i)}$	$c_2^{(i)}$

Table 1: Square ice and their associated Boltzmann weights.

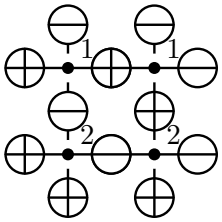
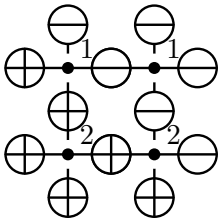
Each of the six types of vertex is assigned a Boltzmann weight, which is allowed to depend on the row i in which it occurs. We have emphasized this dependence in the notation of Table 1. To each state $x \in \mathfrak{S}_\lambda$, the Boltzmann weight $w(x)$ of the

state x is then the product of the Boltzmann weights of all vertices in the state. The *partition function* $Z(\mathfrak{S}_\lambda)$ is defined to be the sum of the Boltzmann weights over all states:

$$Z(\mathfrak{S}_\lambda) = \sum_{x \in \mathfrak{S}_\lambda} w(x).$$

Note: The word "partition" occurs in two different senses in this paper. The partition function in statistical physics is different from partitions in the combinatorial sense. So for us a reference to a "partition" without "function" refers to an integer partition.

As an example, suppose that $r = 1$ and $\lambda = (0, 0)$ so $\lambda + \rho = (1, 0)$. In this case $\mathfrak{S}_\lambda$ has cardinality two. The states and their associated Boltzmann weights are:

state in $\mathfrak{S}_{(0,0)}$		
Boltzmann weight	$b_1^{(1)} c_2^{(1)} c_2^{(2)} b_2^{(2)}$	$c_2^{(1)} a_2^{(1)} a_1^{(2)} c_2^{(2)}$

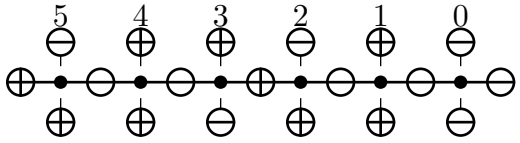
Hence

$$Z(\mathfrak{S}_\lambda) = c_1^{(1)} c_2^{(1)} \left(a_1^{(2)} a_2^{(1)} + b_1^{(1)} b_2^{(2)} \right).$$

The partition function for general λ of arbitrary rank r will be evaluated in Theorem 9, assuming the free-fermionic condition (19).

3 Tokuyama's deformation of the Weyl character formula

Let us momentarily consider a piece of square ice with just one layer of vertices. Let $\alpha_1, \cdots, \alpha_m$ be the column numbers (from left to right) of $-$'s along the top boundary and let $\beta_1, \cdots, \beta_{m'}$ be the column numbers of $-$'s along the bottom boundary. For example, in the ice

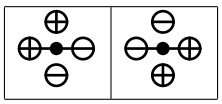


we have $m = 3$, $m' = 2$, $(\alpha_1, \alpha_2, \alpha_3) = (5, 2, 0)$ and $(\beta_1, \beta_2) = (3, 0)$. Since the columns are labeled in decreasing order, we have $\alpha_1 > \alpha_2 > \cdots$ and $\beta_1 > \beta_2 > \cdots$.

Lemma 5. *Suppose that the spin at the left edge is +. Then $m = m'$ or $m' + 1$ and $\alpha_1 \geqslant \beta_1 \geqslant \alpha_2 \geqslant \ldots$. If $m = m'$ then the spin at the right edge is +, while if $m = m' + 1$ it is −.*

We express the condition that $\alpha_1 \geqslant \beta_1 \geqslant \alpha_2 \geqslant \ldots$ by saying that the sequences $\alpha_1, \ldots, \alpha_m$, and $\beta_1, \ldots, \beta_{m'}$, *interleave*. This lemma is essentially the *line-conservation* principle in Baxter [2], Section 8.3.

Proof. The spins along horizontal edges are determined by a choice of spins along the top and bottom boundary and a choice of spin at the left boundary edge (which is assumed to be +). This is clear since, according to the six vertices appearing in Table 1, the edges at each vertex have an even number of + spins. If the rows do not interleave then one of the illegal configurations (i.e., not one of the six in Table 1)

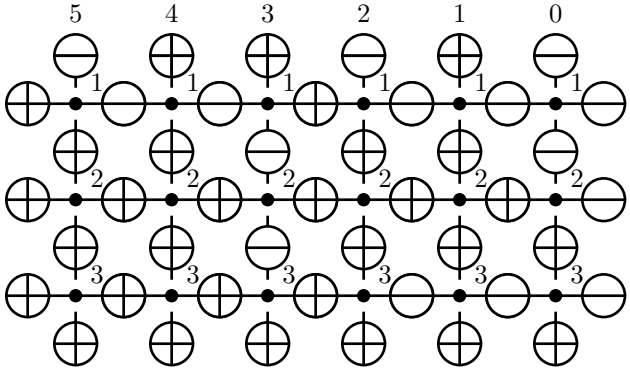


will occur. It follows that $\alpha_1 \geqslant \beta_1$ since if not the vertex in the β_1 column would be surrounded by spins in the first illegal configuration. Similarly $\beta_1 \geqslant \alpha_2$ since otherwise the vertex in the α_2 column would be surrounded by spins in the second above illegal configuration, and so forth. The last statement is a consequence of the observation that the total number of spins must be even. □

A *Gelfand-Tsetlin pattern* is a triangular array of dominant weights (or equivalently integer vectors whose entries are weakly decreasing), in which each row has length one less than the one above it, and the rows interleave. The pattern is called *strict* if the rows are strictly dominant (i.e., the integer components are strictly decreasing).

It follows from Lemma 5 that taking the column indices of − spins along vertical edges gives a sequence of strictly dominant weights forming a strict Gelfand-Tsetlin

pattern. For example, given the state



the corresponding pattern is

$$\mathfrak{T} = \begin{Bmatrix} 5 & & 2 & & 0 \\ & 3 & & 0 & \\ & & 3 & & \end{Bmatrix}. \tag{16}$$

It is not hard to see that this gives a bijection between strict Gelfand-Tsetlin patterns having fixed top row $\lambda + \rho$ and states with boundary conditions determined by λ.

We recall some further definitions from Tokuyama [33]. An entry of a Gelfand-Tsetlin pattern (not in the top row) is classified as *left-leaning* if it equals the entry above it and to the left. It is *right-leaning* if it equals the entry above it and to the right. We will call a pattern *leaning* if all entries below the top row are right- or left-leaning. A pattern is *special* if it is neither left- nor right-leaning. Thus in (16), the 3 in the bottom row is left-leaning, the 0 in the second row is right-leaning and the 3 in the middle row is special. If $\mathfrak{T}$ is a Gelfand-Tsetlin pattern, let $l(\mathfrak{T})$ be the number of left-leaning entries. Let $d_k(\mathfrak{T})$ be the sum of the k-th row of $\mathfrak{T}$, and $d_{r+2}(\mathfrak{T}) = 0$.

Theorem 6. (Tokuyama) *We have*

$$\sum_{\mathfrak{T}} \left(\prod_{k=1}^{r+1} z_k^{d_k(\mathfrak{T}) - d_{k+1}(\mathfrak{T})} \right) t^{l(\mathfrak{T})} (t+1)^{s(\mathfrak{T})} = \prod_{i<j} (z_i + t z_j) s_\lambda(z_1, \cdots, z_{r+1}), \tag{17}$$

where the sum is over all strict Gelfand-Tsetlin patterns with top row $\lambda + \rho$.

As Tokuyama [33] explains, if $t = -1$, this reduces to the Weyl character formula, while if $t = 0$ it reduces to the combinatorial description of Schur polynomials. See

also [3], Chapter 5 for further discussion of Tokuyama's formula. Later in this paper we will give a new proof of this theorem and of a generalization of it by Hamel-King, and we will generalize yet further by evaluating the partition function $Z(\mathfrak{S}_\lambda)$ for Boltzmann weights in the free-fermionic regime.

4 Evaluation of the partition function $Z(\mathfrak{S}_\lambda)$

We begin by recording a version of the parametrized Yang-Baxter equation using the notation from Section 2.

Lemma 7. *Let $S(i) = (a_1^{(i)}, \ldots, c_2^{(i)})$ and $T(j) = (a_1^{(j)}, \ldots, c_2^{(j)})$ be sets of Boltzmann weights corresponding to rows i and j, respectively, satisfying (13). If we choose Boltzmann weights for $R(i,j)$ as follows:*

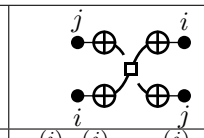
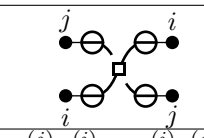
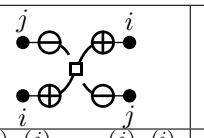
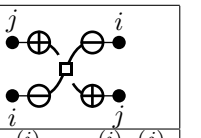
$$R(i,j) = \begin{pmatrix} a_1^{(i)}a_2^{(j)} + b_1^{(j)}b_2^{(i)} & & \\ & a_2^{(i)}b_1^{(j)} - a_2^{(j)}b_1^{(i)} & c_1^{(i)}c_2^{(j)} \\ & c_1^{(j)}c_2^{(i)} & a_1^{(j)}b_2^{(i)} - a_1^{(i)}b_2^{(j)} \\ & & a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)} \end{pmatrix},$$

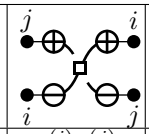
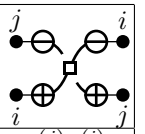
then the star-triangle identity holds:

$$[\![R(i,j), S(i), T(j)]\!] = 0.$$

Proof. This is just a restatement of Theorem 3 using notation for the Boltzmann weights to reflect the dependence on rows. □

Let us record this in tabular form for later reference.

R-ice vertex				
$R(i,j)$	$a_1^{(i)}a_2^{(j)} + b_1^{(j)}b_2^{(i)}$	$a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)}$	$a_2^{(i)}b_1^{(j)} - a_2^{(j)}b_1^{(i)}$	$a_1^{(j)}b_2^{(i)} - a_1^{(i)}b_2^{(j)}$

R-ice vertex		
$R(i,j)$	$c_1^{(i)}c_2^{(j)}$	$c_1^{(j)}c_2^{(i)}$

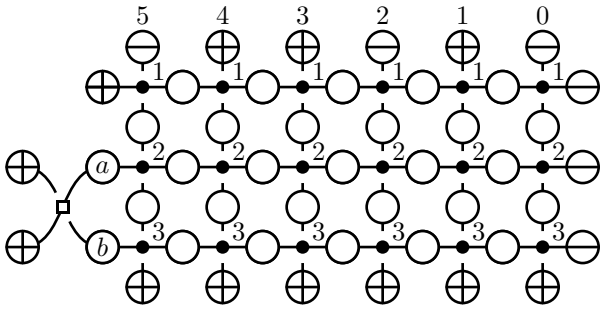
(18)

Lemma 8. *Let $\mathfrak{S}_\lambda$ be an ensemble with boundary conditions corresponding to λ and Boltzmann weights satisfying (13). Then for $i < j$, the expression*

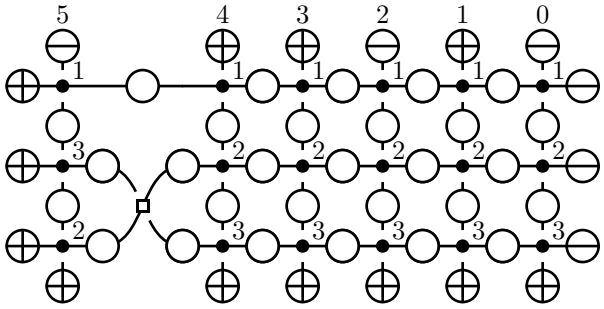
$$(a_1^{(i)}a_2^{(j)} + b_1^{(j)}b_2^{(i)})Z(\mathfrak{S}_\lambda)$$

is invariant under the interchange of spectral parameters i and j.

Proof. We modify the boundary conditions by introducing a single R vertex at the left edge of the ice connecting rows i and j. For simplicity, we illustrate with an ensemble $\mathfrak{S}_\lambda$ with $\lambda = (3, 1, 0)$ and $i = 2, j = 3$:



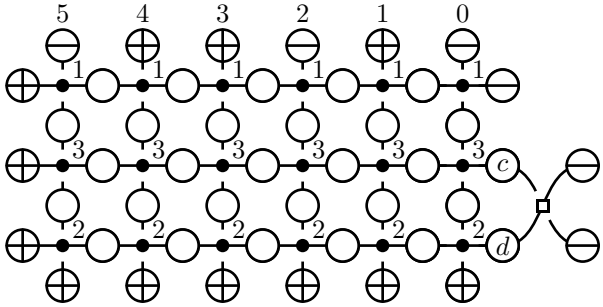
Comparing with the admissible configurations for the R vertex given in the table in (18), the only possible values for a and b are $+$. Thus every state of this new boundary value problem determines a unique state of the original problem, and the partition function for each such state is the original partition function multiplied by the Boltzmann weight of the R-vertex, which is $(a_1^{(i)}a_2^{(j)} + b_1^{(j)}b_2^{(i)})$. Now we apply the star-triangle identity, and obtain equality with the the following configuration (again pictured for our special case):



Thus if $\mathfrak{S}'$ denotes this ensemble then its partition function is

$$Z(\mathfrak{S}') = (a_1^{(i)}a_2^{(j)} + b_1^{(j)}b_2^{(i)})Z(\mathfrak{S}_\lambda).$$

Repeatedly applying the star-triangle identity, we eventually obtain the configuration in which the R-vertex is moved entirely to the right:



Now there is only one admissible configuration for the R-vertex on the right-hand side, namely $c = d = -$. The Boltzmann weight for this R-vertex is $a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)}$. Note that upon moving through the ice, the roles of i and j have been interchanged. Comparing partition functions, this proves that $(a_1^{(i)}a_2^{(j)}+b_1^{(j)}b_2^{(i)})Z(\mathfrak{S}_\lambda)$ is unchanged by switching the spectral parameters i and j. □

Theorem 9. *Let λ be a partition with $r+1$ parts, largest part λ_1 and smallest part 0. Let $\mathfrak{S}_\lambda$ be the corresponding ensemble, with $r+1$ rows. Suppose that*

$$a_1^{(i)}a_2^{(i)} + b_1^{(i)}b_2^{(i)} = c_1^{(i)}c_2^{(i)}. \tag{19}$$

Then

$$Z(\mathfrak{S}_\lambda) = \left[\prod_{k=1}^{r+1}(a_1^{(k)})^{\lambda_1}c_2^{(k)}\prod_{i<j}(a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)})\right] \mathbf{s}_\lambda\left(\frac{b_2^{(1)}}{a_1^{(1)}}, \frac{b_2^{(2)}}{a_1^{(2)}}, \ldots, \frac{b_2^{(r+1)}}{a_1^{(r+1)}}\right) \tag{20}$$

where $\mathbf{s}_\lambda$ is the Schur polynomial corresponding to λ.

Remark 4.1. The equation (19) describes the *free-fermionic regime* in the six-vertex model. More generally, see Fan and Wu [8], [9] for the free-fermionic eight-vertex model.

The identity (20) is an equality of homogeneous polynomials of degree λ_1+r+1 in $a_1^{(i)}, \ldots, c_2^{(i)}$ for each i. The Schur polynomial is expressed in terms of the variables $b_2^{(i)}/a_1^{(i)}$, but if $a_1^{(i)} = 0$ for some i then, due to this homogeneity, one may clear denominators before evaluating and the equation (20) still makes sense.

Turning to the proof of Theorem 9, also due to the homogeneity of (20), we may assume $c_2^{(i)} = 1$ for all i.

Lemma 10. *Given any partition λ, the expression*

$$s_{\mathfrak{S}_\lambda} \stackrel{def}{=} \left[\prod_{i<j}(a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)})\right]^{-1} Z(\mathfrak{S}_\lambda)$$

is symmetric with respect to the spectral parameters and expressible as a polynomial in the variables $a_1^{(i)}, a_2^{(i)}, b_1^{(i)}, b_2^{(i)}$ with integer coefficients.

Proof. It suffices to show this function is invariant under transpositions (i.e. interchanging k and $k+1$). By Lemma 8, the function $(a_1^{(k)}a_2^{(k+1)} + b_1^{(k+1)}b_2^{(k)})Z(\mathfrak{S}_\lambda)$ is invariant under the interchange $k \leftrightarrow k+1$. It follows that

$$\left[\prod_{i<j}(a_1^{(i)}a_2^{(j)} + b_1^{(j)}b_2^{(i)})\right] Z(\mathfrak{S}_\lambda) = \left[\prod_{i\neq j}(a_1^{(i)}a_2^{(j)} + b_1^{(j)}b_2^{(i)})\right] s_{\mathfrak{S}_\lambda} \tag{21}$$

is invariant under $k \leftrightarrow k+1$ since the left-hand side of (21) is a product of $(a_1^{(k)}a_2^{(k+1)} + b_1^{(k+1)}b_2^{(k)})Z(\mathfrak{S}_\lambda)$ and factors $(a_1^{(i)}a_2^{(j)} + b_1^{(j)}b_2^{(i)})$ that are permuted under $k \leftrightarrow k+1$. Thus $s_{\mathfrak{S},\lambda}$ must also be symmetric.

The identity (19) with $c_2^{(i)} = 1$ becomes $a_1^{(i)}a_2^{(i)} + b_1^{(i)}b_2^{(i)} = c_1^{(i)}$. This allows one to eliminate $c_1^{(i)}$ $(i = 1, \ldots, r+1)$ from $Z(\mathfrak{S}_\lambda)$, regarding it as a polynomial in the ring

$$\mathcal{R} = \mathbb{Z}[a_1^{(1)}, \ldots, b_2^{(1)}, \ldots, a_1^{(r+1)}, \ldots, b_2^{(r+1)}].$$

The left-hand side of (21) is divisible by $(a_1^{(i)}a_2^{(j)} + b_1^{(j)}b_2^{(i)})$ with $i < j$ and may be regarded as an element of the unique factorization domain $\mathcal{R}$. As the left-hand side of (21) is symmetric, we conclude that it is also divisible by $(a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)})$ with $i < j$. This shows $s_{\mathfrak{S}_\lambda}$ is a polynomial in the $a_1^{(i)}, a_2^{(i)}, b_1^{(i)}, b_2^{(i)}$. □

Proof of Theorem 9. Since $s_{\mathfrak{S}_\lambda}$ defined in Lemma 10 is independent of $c_1^{(i)}$ for all i, we may take $c_1^{(i)} = 0$ for all i in computing the partition function $Z(\mathfrak{S}_\lambda)$. Upon doing this, the remaining states of ice with non-zero Boltzmann weights (i.e., those without any $c_1^{(i)}$) are in bijection with leaning Gelfand-Tsetlin patterns.

We will show that in this case, the function $s_{\mathfrak{S}_\lambda}$ is, up to a constant multiple, a Schur polynomial in the variables $b_2^{(i)}/a_1^{(i)} = -a_2^{(i)}/b_1^{(i)}$ by comparing our expression with the Weyl character formula. First, we demonstrate that the product

$$\left[\prod_{i<j}(a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)})\right]$$

can be expressed in terms of the Weyl denominator in the $b_2^{(i)}/a_1^{(i)}$. Indeed, for any pair i, j with $i < j$, write

$$a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)} = a_1^{(j)}\left(\frac{-b_1^{(i)}b_2^{(i)}}{a_1^{(i)}}\right) + b_1^{(i)}b_2^{(j)} = a_1^{(j)}b_1^{(i)}\left(\frac{b_2^{(j)}}{a_1^{(j)}} - \frac{b_2^{(i)}}{a_1^{(i)}}\right),$$

where we have used the identity (19) in the form $a_1^{(i)}a_2^{(i)} + b_1^{(i)}b_2^{(i)} = 0$ since $c_1^{(i)} = 0$. Performing this for all such pairs i, j, we have

$$\begin{aligned}\prod_{i<j}(a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)}) &= \prod_{i<j} a_1^{(j)}b_1^{(i)}\left(\frac{b_2^{(j)}}{a_1^{(j)}} - \frac{b_2^{(i)}}{a_1^{(i)}}\right) \\ &= \prod_{k=1}^{r+1}(a_1^{(k)})^{k-1}(b_1^{(k)})^{r+1-k}\prod_{i<j}\left(\frac{b_2^{(j)}}{a_1^{(j)}} - \frac{b_2^{(i)}}{a_1^{(i)}}\right). \qquad (22)\end{aligned}$$

As we argued in Lemma 10, after setting $c_1^{(i)} = 0$, the function $Z(\mathfrak{S}_\lambda)$ is a polynomial in the variables $a_1^{(i)}, a_2^{(i)}, b_1^{(i)}, b_2^{(i)}$. We make the substitution $a_2^{(i)} = -b_1^{(i)}z_i$ for all i where the z_i are (for the moment) just formal parameters. Call the resulting function $N_{\mathfrak{S}_\lambda} = N_{\mathfrak{S}_\lambda}(z_1, \ldots, z_{r+1})$, which is a polynomial in the z_i whose coefficients are polynomial expressions in the $a_1^{(i)}, b_1^{(i)}, b_2^{(i)}$. We claim that the power of $b_1^{(i)}$ appearing in each coefficient of $N_{\mathfrak{S}_\lambda}$ is equal to $r+1-i$. Indeed, this is the total number of $a_2^{(i)}$ and $b_1^{(i)}$ appearing in the Boltzmann weight of each state of $Z(\mathfrak{S}_\lambda)$. These weights are contributed by the two vertices having north and south spins both equal to $-$. In leaning patterns, the number of such vertices in row i is always $r+1-i$. Thus

$$N'_{\mathfrak{S}_\lambda} \stackrel{def}{=} N_{\mathfrak{S}_\lambda}\prod_{k=1}^{r+1}(b_1^{(k)})^{-(r+1-i)}$$

is independent of $b_1^{(k)}$ for all k. Hence $N'_{\mathfrak{S}_\lambda}$ is a polynomial in the z_i with coefficients that are polynomials in the $a_1^{(i)}, b_2^{(i)}$. These two Boltzmann weights are the unique pair with north and south spins both equal to $+$. In states of ice corresponding to leaning Gelfand-Tsetlin patterns, the total number of such vertices in row i is equal to $\ell_1 + i - (r+1)$ where $\ell_1 = \lambda_1 + \cdots + \lambda_r + r$ is the index of the left-most column in ice in $\mathfrak{S}_\lambda$. Thus we may write

$$N'_{\mathfrak{S}_\lambda} = \prod_{i=1}^{r+1}(a_1^{(i)})^{\ell_1+i-(r+1)}N''_{\mathfrak{S}_\lambda}$$

where $N''_{\mathfrak{S}_\lambda}$ is a polynomial in the z_i and $b_2^{(i)}/a_1^{(i)}$ with integer coefficients.

Initially we set $z_i = -a_2^{(i)}/b_1^{(i)}$. However, in light of the relation $a_1^{(i)}a_2^{(i)} + b_1^{(i)}b_2^{(i)} = c_1^{(i)} = 0$, we also have $z_i = b_2^{(i)}/a_1^{(i)}$. Thus $N''_{\mathfrak{S}_\lambda}$ is a polynomial in the z_i with integer coefficients. Combining with (22), we have

$$s_{\mathfrak{S}_\lambda} = \left[\prod_{i=1}^{r+1}\prod_{i<j}(a_1^{(j)}a_2^{(i)} + b_1^{(i)}b_2^{(j)})\right]^{-1} Z(\mathfrak{S}_\lambda) = \left[\prod_{i=1}^{r+1} a_1^{(i)}\right]^{\lambda_1} \frac{N''_{\mathfrak{S}_\lambda}(z_1,\ldots,z_{r+1})}{\prod_{i<j}(z_j - z_i)}, \quad (23)$$

where λ_1 is the largest part of the partition.

The weight μ of a state, i.e., its degree of as a monomial in z_i, is given by counting the number of vertices having Boltzmann weight a_2 or b_2 in row i. These are the unique pair of vertices having west spin equal to $-$ (since $c_1^{(i)} = 0$). It is easy to see that the weight μ of any non-zero state of ice is a permutation σ of the top row of $\mathfrak{T}$, that is, of $\lambda + \rho$. These weights are all distinct since $\lambda + \rho$ is strongly dominant, i.e. without repeated entries, so in fact the coefficients of $N''_{\mathfrak{S}_\lambda}$ are all ± 1. Since $N''_{\mathfrak{S}_\lambda}$ is skew-symmetric (because the denominator is skew-symmetric while $s_{\mathfrak{S}_\lambda}$ is symmetric), $\pm N''_{\mathfrak{S}_\lambda}$ is equal to the sum over permutations σ of terms of the form $\mathrm{sgn}(\sigma)\prod z_j^{\ell_{\sigma(j)}}$, where $\ell_i = \lambda_i + \rho_i$ for all i. To determine the sign, we may take the state whose corresponding Gelfand-Tsetlin pattern consists entirely of right-leaning entries:

$$\mathfrak{T} = \left\{\begin{array}{ccccccccc} \ell_1 & & \ell_2 & \cdots & & & \ell_r & & 0 \\ & \ell_2 & \cdots & & & \ell_r & & 0 & \\ & & \ddots & & & \iddots & & & \\ & & & & 0 & & & & \end{array}\right\}.$$

This has Boltzmann weight $\prod z_j^{\ell_j}$ and so

$$N''_{\mathfrak{S}_\lambda}(z_1,\cdots,z_{r+1}) = \sum_{\sigma\in S_{r+1}} \mathrm{sgn}(\sigma)\prod z_j^{\ell_{\sigma(j)}}.$$

The theorem then follows by combining the above with (23) and invoking the Weyl character formula. □

5 Another proof of Tokuyama-Hamel-King

This section gives a new proof of results of Tokuyama and Hamel-King.

Proof of Theorem 6. Using the specialization of the weights in Theorem 9 as follows (for all rows $1 \le i \le r+1$):

$$a_1^{(i)} = 1, a_2^{(i)} = z_i, b_1^{(i)} = t_i, b_2^{(i)} = z_i, c_1^{(i)} = z_i(t_i+1), c_2^{(i)} = 1. \tag{24}$$

(These weights are called $S^\Gamma(i)$ in Table 2 below.) If all $t_i = t$ then the resulting partition function simplifies to the right-hand side of Tokuyama's theorem (17), and in general

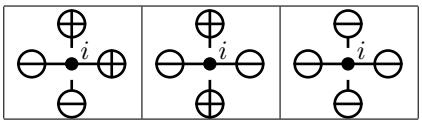
$$*Z(\mathfrak{S}_\lambda) = \prod_{i<j} (t_i z_j + z_i)\, s_\lambda(z_1, \ldots, z_{r+1}). \tag{25}$$

It remains to use the bijection between states in $\mathfrak{S}_\lambda$ and strict Gelfand-Tsetlin patterns with top row $\lambda + \rho$ to show that the summands on the left-hand side of (17) are equal to the Boltzmann weights of the corresponding states.

Given Boltzmann weights as in (24), we say that the *z-weight* of a state is $(\mu_1, \cdots, \mu_n)$ if the Boltzmann weight is the monomial $\boldsymbol{z}^\mu = \prod z_i^{\mu_i}$ times a polynomial in the variables t_i. Recall that if $\mathfrak{T}$ is a Gelfand-Tsetlin pattern, we set $d_k(\mathfrak{T})$ to be the sum of the k-th row and $d_{r+2}(\mathfrak{T}) = 0$. The following lemma shows that the powers of z_i appearing in the Boltzmann weight of a state agree with the corresponding summand in Tokuyama's theorem.

Lemma 11. *If $\mathfrak{T}$ is the Gelfand-Tsetlin pattern corresponding to a state of z-weight μ, then $\mu_k = d_k(\mathfrak{T}) - d_{k+1}(\mathfrak{T})$.*

Proof. From Table 1, μ_k is the number of vertices in the k-th row that have an edge configuration of one of the three forms:



Let α_i's (respectively β_i's) be the column numbers for which the top edge spin (respectively, the bottom edge spin) of vertices in the k-th row is $-$ (with columns numbered in descending order, as always). By Lemma 5 we have $\alpha_1 \geqslant \beta_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_{r+2-k}$. It is easy to see that the vertex in the j-column has one of the above configurations if and only if its column number j satisfies $\alpha_i > j \geqslant \beta_i$ for some i. Therefore the number of such j is $\sum \alpha_i - \sum \beta_i = d_k(\mathfrak{T}) - d_{k+1}(\mathfrak{T})$. □

Finally, it is easy to see that if an entry in the k-th row of $\mathfrak{T}$ is left leaning (respectively special), and that entry is j, then the configuration in the j-column

and the k-th row of the ice is

so from our specialization of weights above, it follows that the powers of t_i in the Boltzmann weight of a state match the powers of t_i in the corresponding summands on the left-hand side of (17). Setting all $t_i = t$ gives the result of Tokuyama. □

It is clear from the proof that we've shown something a bit stronger, in that the t_i's are allowed to be independent variables. Indeed, the case of t_i independent is easily seen to be equivalent to Hamel and King's generalization of Tokuyama's theorem (cf. Proposition 1.1 of [13].)

6 A relation between partition functions

It is natural to ask whether other models and choices of Boltzmann weights exist for which the resulting partition function matches the right-hand side of (17). Let us denote by $\mathfrak{S}_\lambda^\Gamma$ the ice model with boundary conditions as in Section 2 and assignment of Boltzmann weights as in (24). So we may write

$$Z(\mathfrak{S}_\lambda^\Gamma) = \prod_{i<j}(z_i + t_i z_j) s_\lambda(z_1, \cdots, z_{r+1}).$$

If we simply replace the choices of $a_2^{(i)}$ and $b_1^{(i)}$ appearing in (24) with

$$a_2^{(i)} = t_i z_i, \quad b_1^{(i)} = 1,$$

then by Theorem 9, the resulting partition function will be

$$Z(\mathfrak{S}_\lambda) = \prod_{i<j}(t_i z_i + z_j) s_\lambda(z_1, \ldots, z_{r+1}).$$

However, if we also change the model slightly by renumbering the rows in *descending* order from top to bottom and call the resulting model $\mathfrak{S}_\lambda^\Delta$ then Theorem 9 implies that

$$Z(\mathfrak{S}_\lambda^\Delta) \prod_{i<j}(t_i z_j + z_i) = Z(\mathfrak{S}_\lambda^\Gamma) \prod_{i<j}(t_j z_j + z_i), \tag{26}$$

and setting all $t_i = t$, we may cancel the products on either side giving a second expression for Tokuyama's result. The equality of partition functions in (26) is

deceptively subtle as there is provably no bijective proof which matches Boltzmann weights of states in $\mathfrak{S}_\lambda^\Gamma$ and $\mathfrak{S}_\lambda^\Delta$.

Four proofs of this identity are known. The first has already been given—we may exactly solve both models using (9) and then conclude they are equal.

Second, we may prove (20) by realizing each side as the same Whittaker coefficient of a minimal parabolic Eisenstein series on a covering group—a certain integral over a unipotent group inductively calculated in two different ways. This method will not be described in detail but see [4] for one of the two inductive calculations. (The other is not written down but similar.) The identity corresponds to the special case $n = 1$ where n is the degree of the cover that occurs in [4].

Third, this is the special case of Statement B in [3] in which the degree n of the Gauss sums that appear in that statement equals 1. There, sums over lattice points in two different polytopes are compared by a combinatorial procedure related to the Schützenberger involution of a crystal graph.

We conclude this section with a fourth proof of (26) using another instance of the parametrized Yang-Baxter equation. This allows us to compare two partition functions without having to explicitly evaluate either. As mentioned above, a generalization of (26) is known using n-th order Gauss sums which specialize, when $n = 1$, to the Boltzmann weights above. For general n, ice models exist for both sides of the identity in Statement B in [3] but no corresponding Yang-Baxter equation is known. See [5] and Chapter 19 of [3] for further discussion of this.

We turn to the proof of (26). Recall that the Boltzmann weights $S^\Gamma(i)$ and $S^\Delta(i)$ used in the systems $\mathfrak{S}_\lambda^\Gamma$ and $\mathfrak{S}_\lambda^\Delta$ are as in Table 2. This table also defines a third type of Boltzmann weight $R^{\Gamma\Delta}(i,j)$ that we will require for the proof of (26).

As in Table 2, vertices in a given state having Boltzmann weights corresponding to Γ will be labeled with a black dot ($\bullet$) and those corresponding to Δ Boltzmann weights will be labeled with an open dot ($\circ$).

Lemma 12. *Consider the Boltzmann weights in Table 2. Then the following star-triangle identity holds:*

$$\left[\!\left[R^{\Gamma\Delta}(i,j), S^\Gamma(i), T^\Delta(j)\right]\!\right] = 0.$$

Proof. This is just a special case of Lemma 7. We have

$$R^{\Gamma\Delta}(i,j) = \begin{pmatrix} t_j z_j + z_i & & & \\ & z_i - t_i t_j z_j & z_i(t_i+1) & \\ & z_j(t_j+1) & z_i - z_j & \\ & & & z_i + t_i z_j \end{pmatrix}.$$

□

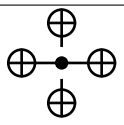
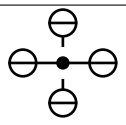
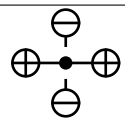
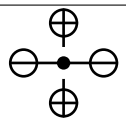
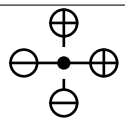
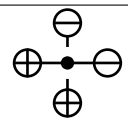
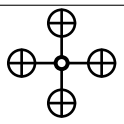
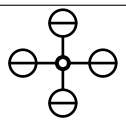
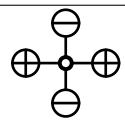
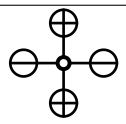
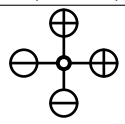
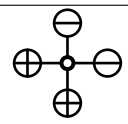
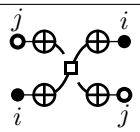
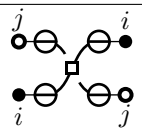
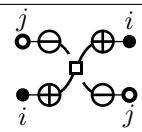
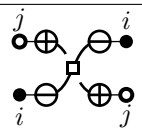
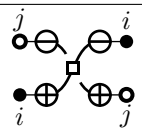
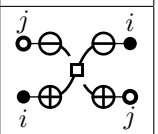
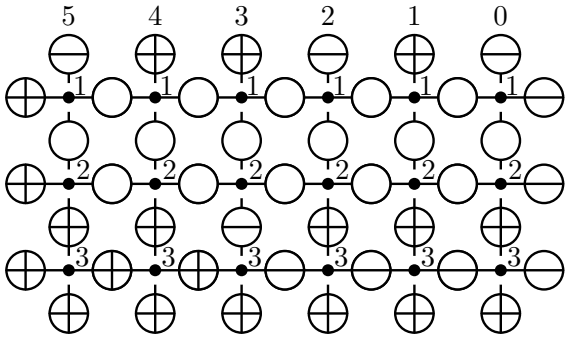
	a_1	a_2	b_1	b_2	c_1	c_2
$S^\Gamma(i)$						
	1	z_i	t_i	z_i	$z_i(t_i+1)$	1
$S^\Delta(i)$						
	1	z_i	t_i	z_i	$z_i(t_i+1)$	1
$R_{\Gamma\Delta}(i,j)$						
	$t_jz_j+z_i$	$t_iz_j+z_i$	$t_it_jz_j-z_i$	z_i-z_j	$(t_i+1)z_i$	$(t_j+1)z_j$

Table 2: Boltzmann weights for (26) and its proof.

Proposition 13. *Let Boltzmann weights for two ice models $\mathfrak{S}^\Gamma$ and $\mathfrak{S}^\Delta$ be chosen as in Table 2, both having boundary conditions corresponding to a partition λ as in Section 2, but with rows in states of $\mathfrak{S}^\Gamma$ labeled in ascending order from top to bottom, and rows in states of $\mathfrak{S}^\Delta$ in descending order. Then*

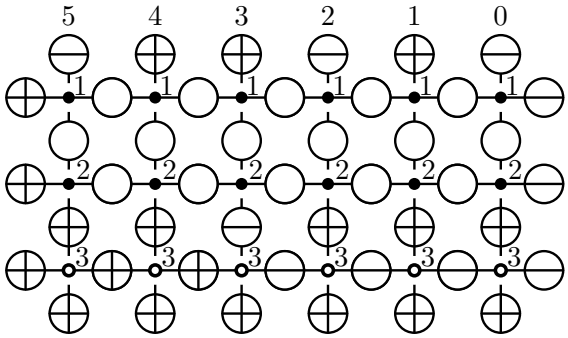
$$Z(\mathfrak{S}^\Delta_\lambda)\prod_{i<j}(t_iz_j+z_i) = Z(\mathfrak{S}^\Gamma_\lambda)\prod_{i<j}(t_jz_j+z_i).$$

Proof. Begin with an state x of $\mathfrak{S}^\Gamma_\lambda$, say (for example with $\lambda=(3,1,0)$):

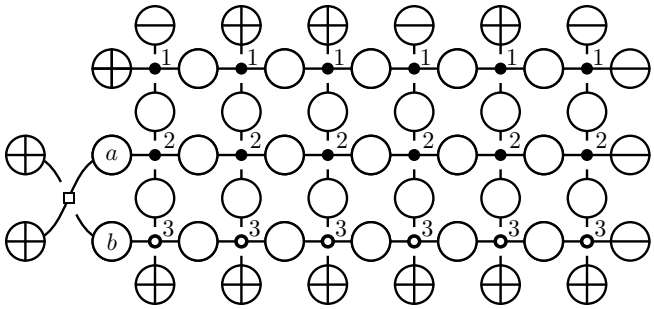


(We're focusing on the bottom row for the moment, so the unlabeled edges can be filled in arbitrarily.) We wish to transform this into a state having a row of Delta

ice so that we may use the star-triangle relation in Lemma 12. The vertices in the bottom row all have south spin equal to + and the Boltzmann weights for Γ and Δ given in Table 2 differ only for a_2 and b_1 vertices, which both have south spin equal to $-$. Hence we may simply consider the bottom vertices to be Δ Boltzmann weights without affecting the Boltzmann weight of the entire state:

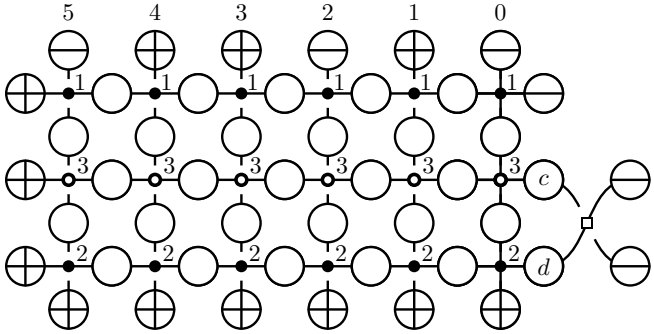


Note that this would not work in any row but the last because it is essential that there be no $-$ on the bottom edge spins. Now we add a Gamma-Delta R-vertex.

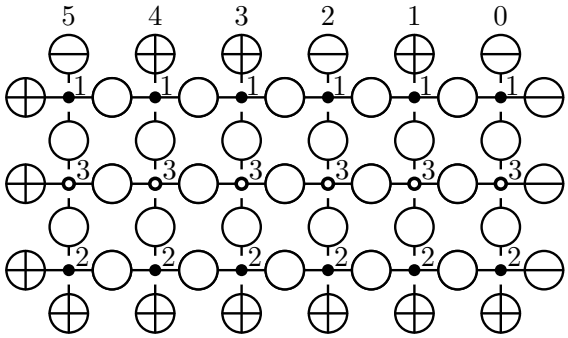


If we call this model $\mathfrak{S}'$, then we claim that $Z(\mathfrak{S}') = (t_3 z_3 + z_2) Z(\mathfrak{S}^{\Gamma}_{\lambda})$. Indeed, from Lemma 12, the values of spins a and b indicated in the figure must both be + and so the value of the R-vertex is $t_3 z_3 + z_2$ for every state in the model. Now repeatedly using the star-triangle relation, $Z(\mathfrak{S}') = Z(\mathfrak{S}'')$ where $\mathfrak{S}''$ is the model

with boundary:



Here we must have spins c, d in the figure above both equal to $-$ which implies that $(t_3z_3+z_2)Z(\mathfrak{S}^\Gamma_\lambda) = Z(\mathfrak{S}'') = (t_2z_3+z_2)Z(\mathfrak{S}''')$ where $\mathfrak{S}'''$ is the model with boundary of form



We repeat the process, first moving the row of Delta ice up to the top, then introducing another row of Delta ice by simple replacement at the bottom, etc., until we arrive at the model $\mathfrak{S}^\Delta_\lambda$ and obtain (26). □

References

[1] R. J. Baxter. The inversion relation method for some two-dimensional exactly solved models in lattice statistics. *J. Statist. Phys.*, 28(1):1–41, 1982.

[2] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1982.

[3] B. Brubaker, D. Bump, and S. Friedberg. *Weyl Group Multiple Dirichlet Series: Type A Combinatorial Theory*, Annals of Mathematics Studies Vol. 175, Princeton University Press, Princeton, NJ, 2011.

[4] B. Brubaker, D. Bump, and S. Friedberg. Weyl group multiple Dirichlet series, Eisenstein series and crystal bases, *Ann. of Math.* 173, 1081–1120, 2011.

[5] B. Brubaker, D. Bump, G. Chinta, S. Friedberg, and P. Gunnells. Metaplectic ice, preprint, 2010. `http://arxiv.org/abs/1009.1741`.

[6] V. G. Drinfeld. Quantum groups. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986)*, pages 798–820, Amer. Math. Soc., Providence, RI, 1987.

[7] L. D. Faddeev, N. Yu. Reshetikhin, and L. A. Takhtajan. Quantization of Lie groups and Lie algebras. In *Algebraic Analysis, Vol. I*, pages 129–139. Academic Press, Boston, MA, 1988.

[8] C. Fan and F. Y. Wu. Ising model with next-neighbor interactions. I. Some exact results and an approximate solution, *Phys. Rev.* 179, 560-570 (1969).

[9] C. Fan and F. Y. Wu. General lattice model of phase transitions. *Physical Review B*, 2(3):723–733, 1970.

[10] S. Fomin and A. N. Kirillov. The Yang-Baxter equation, symmetric functions, and Schubert polynomials. *Proc. of the 5th Conference on Formal Power Series and Algebraic Combinatorics (Florence, 1993)*, vol. 153, pages 123–143, 1996.

[11] S. Fomin and A. N. Kirillov. Grothendieck polynomials and the Yang-Baxter equation. In *Formal power series and algebraic combinatorics/Séries formelles et combinatoire algébrique*, pages 183–189. DIMACS, Piscataway, NJ, 1994.

[12] A. M. Hamel and R. C. King. U-turn alternating sign matrices, symplectic shifted tableaux and their weighted enumeration. *J. Algebraic Combin.*, 21(4):395–421, 2005.

[13] A. M. Hamel and R. C. King. Bijective proofs of shifted tableau and alternating sign matrix identities. *J. Algebraic Combin.*, 25(4):417–458, 2007.

[14] A. G. Izergin. Partition function of a six-vertex model in a finite volume. *Dokl. Akad. Nauk SSSR*, 297(2):331–333, 1987.

[15] M. Jimbo. Introduction to the Yang-Baxter equation. *Internat. J. Modern Phys. A*, 4(15):3759–3777, 1989.

[16] M. Jimbo and T. Miwa. Solitons and infinite-dimensional Lie algebras. *Publ. Res. Inst. Math. Sci.* 19 (1983), no. 3, 943–1001.

[17] A. N. Kirillov and N. Yu. Reshetikhin, N. Yu. The Bethe ansatz and the combinatorics of Young tableaux. *J. Soviet Math.* 41 (1988), no. 2,

[18] V. E. Korepin. Calculation of norms of Bethe wave functions. *Comm. Math. Phys.*, 86(3):391–418, 1982.

[19] G. Kuperberg. Another proof of the alternating-sign matrix conjecture. *Internat. Math. Res. Notices*, (3):139–150, 1996.

[20] A. Lascoux. Chern and Yang through ice. *Preprint* (2002).

[21] A. Lascoux. The 6 vertex model and Schubert polynomials. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 3:Paper 029, 12 pp. (electronic), 2007.

[22] A. Lascoux and M.-P. Schützenberger. Symmetry and flag manifolds. In *Invariant theory (Montecatini, 1982)*, volume 996 of *Lecture Notes in Math.*, pages 118–144. Springer, Berlin, 1983.

[23] E. Lieb. Exact solution of the problem of entropy in two-dimensional ice. *Phys. Rev. Lett.*, 18:692–694, 1967.

[24] I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky.

[25] S. Majid. Quasitriangular Hopf algebras and Yang-Baxter equations. *Internat. J. Modern Phys. A*, 5(1):1–91, 1990.

[26] P. J. McNamara. Factorial Schur functions via the six-vertex model. *Preprint* (2009).

[27] W. H. Mills, D. P. Robbins, and H. Rumsey, Jr. Alternating sign matrices and descending plane partitions. *J. Combin. Theory Ser. A*, 34(3):340–359, 1983.

[28] S. Okada. Alternating sign matrices and some deformations of Weyl's denominator formulas. *J. Algebraic Combin.*, 2(2):155–176, 1993.

[29] D. P. Robbins and H. Rumsey, Jr. Determinants and alternating sign matrices. *Adv. in Math.*, 62(2):169–184, 1986.

[30] W. Stein et. al. SAGE Mathematical Software, Version 4.1. http://www.sagemath.org, 2009.

[31] Yu. G. Stroganov. The Izergin-Korepin determinant at a cube root of unity. *Teoret. Mat. Fiz.*, 146(1):65–76, 2006.

[32] B. Sutherland. Exact solution for a model for hydrogen-bonded crystals. *Phys. Rev. Lett.*, 19(3):103–4, 1967.

[33] T. Tokuyama. A generating function of strict Gelfand patterns and some formulas on characters of general linear groups. *J. Math. Soc. Japan*, 40(4):671–685, 1988.

[34] N. V. Tsilevich. The quantum inverse scattering problem method for the q-boson model, and symmetric functions. *Funktsional. Anal. i Prilozhen.*, 40(3):53–65, 96, 2006.

[35] P. Zinn-Justin. Six-vertex loop and tiling models: Integrability and combinatorics. *Habilitation thesis* (arXiv:0901.0665), 2009.